AN ECONOMIC EVALUATION OF MASS POPULATION SCREENING FOR COLORECTAL CANCER USING A FAECAL OCCULT BLOOD TEST

Andrew Walker

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

December 1993

Acknowledgements

Financial support
Most of the work for this thesis was carried out while I was employed as part of the Medical Research Council trial of colorectal cancer screening in Nottingham; the financial support of that organisation is gratefully acknowledged. After leaving Nottingham I was allowed time to finish this work while in the employment of Northern Regional Health Authority and the Health Economics Research Unit, University of Aberdeen. Thank you to those responsible for giving me the chance to tie up the lose ends.

Personal acknowledgements
The advice and support of my supervisors David Whynes and Professor Hardcastle has been invaluable; many of their thoughts will have crept (unaccredited) into the following pages. Thanks also to Professor Chamberlain of the Cancer Screening Evaluation Unit.

The task of data collection was considerably simplified by the availability of thorough trial records, for which I am indebted to Chris Mangham and her staff (June, Liz and Pat) as well as to Neil Baker.

Many other people in the Department of Surgery at Nottingham provided me with advice, particularly Mike Thomas, Mike Robinson, Geoff Pye and Richard Charnley.

Finally, for on-going support over the past five years above and beyond the calls of friendship, I owe a debt of gratitude to Sara and Steve Twaddle.

Thank you all.

Contents

Abstract

Cancer of the colon and rectum is a major cause of ill-health. Options for reducing the burden of the disease include primary prevention, screening for early stage asymptomatic disease and improvements in the treatment of symptomatic disease. If the policy objective is to make a major impact on mortality from the disease then screening appears to be the only technically feasible option.

One indication of asymptomatic colorectal cancer is small quantities of blood mixed with faeces. Screening tests capable of detecting bleeding are currently being evaluated in clinical trials. Interim measures of the costs and disease yield of a screening programme using a faecal occult blood test imply that screening may offer good value for money but only if the intended mortality reduction from the disease is realised.

There are various ways of 'fine-tuning' the screening programme to improve the balance of costs and benefits; information for making choices regarding important parameters such as the age range of the population to be offered screening are presented. Alternative screening tests are also evaluated and the results presented in terms of the cost-yield trade-off.

The policy implications of the evaluation must be qualified at this stage since no proof of mortality reduction will be available until the conclusion of the on-going trial. Nevertheless, under various assumptions about the impact of screening, the option appears to be an efficient way of reducing the health 'costs' of colorectal cancer.

INTRODUCTION

The evaluation of health services

In 1990 an internal market was introduced into the provision of publicly-funded health services in the United Kingdom. Functions were allocated to the charge of various management groups. Provider units took on the task of delivering services, either from a hospital or in the community. A separate function was established to consider the health needs of local people and to commission the services they required from the providers; the bodies carrying out this task are the purchasers*. The potential role of health economics is thus much more clear-cut: purchasers are given a fixed budget and attempt to achieve their objectives to the greatest extent possible. These objectives will include health gain (in terms of length or quality of life), as well as other factors. Thus, services which are commissioned must be clinically effective (i.e. there is evidence that patients' health will improve).

The budget constraint also requires that efficiency be considered in terms of health benefits per pound spent. It is widely accepted that the cost of providing all of the services from which people could potentially benefit outstrips the resources available, thus there is a situation of scarcity. As a result, opportunity costs become relevant. This, in turn, has created a demand for information on the relative costs and benefits of alternative health care treatments in order to assist decision-making. Economic evaluation is a systematic means of evaluating two or more options in terms of their relative costs and benefits from a specified perspective. While easily stated, these ideas are not always straightforward to apply in practice.

Health gain as a strategic goal of cancer services

The internal market is still regulated, however: for example, targets have been set for purchasers in terms of reducing the health 'costs' of particular diseases in the "Health of the Nation" (Department of Health (1991)). The first set of targets concentrated on diseases with the biggest burden in terms of years of life lost prematurely (defined as before the age of 65); these do not appear to have been set with economic criteria in mind. Purchasers had already recognised that cancer care was one of their top priorities (Klein and Redmayne (1992)). Some strategies are already available, such as smoking cessation programmes to tackle lung cancer and screening for breast cancer. However, the second biggest cause of cancer deaths, cancer of the large bowel (i.e. the colon and rectum), is not being tackled apart from providing prompt surgical attention for cases diagnosed. Some advances in treatment are being examined but their impact will be limited at best. Advances in genetics research may one day allow all of those at risk of the disease to be identified at an early stage of their lives and treated appropriately but this is not yet feasible. Thus attention has turned to screening as the only option

* Increasingly, budget-holding General Practitioners (GPs) are carrying out this role but the bulk of the work is still the province of the health authorities.

which has the potential to substantially reduce the health 'costs' of the disease.

The problem is that there is little evidence that such a policy will extend lives. Some evidence from America is available but relates to protocols which are not generally considered feasible in this country on grounds of expense. In addition the quality of some studies suggests that the data should be taken to be indicative rather than conclusive. While breast cancer screening was introduced in this country in advance of the results of the on-going British trial, today's purchasers may well be more wary of commissioning services for which there is inadequate proof of benefit. The economic efficiency of the programme will also be carefully scrutinised as part of the decision; it is the evidence that will be needed that forms the subject matter of this evaluation.

The issues are considered under four headings:

i) what are the costs and benefits to NHS purchasers of screening for colorectal cancer in an asymptomatic population aged 50-74 by an offer of Haemoccult II testing every two years?

ii) is this the most efficient way to screen for colorectal cancer?

iii) is this the most efficient way to reduce the health loss of colorectal cancer?

iv) can such a policy be justified in comparison with other uses of health service resources?

To address these questions the structure of the evaluation is as follows:

Chapter One sets out the health and economic costs of colorectal cancer. The poor prognosis for advanced stage cases is noted. The causes of colorectal cancer are discussed and current treatment practices described. Neither offers immediate prospects for improving the prognosis for the majority of patients. In conclusion, attention switches to screening for early stage disease.

Chapter Two picks up this thread by considering the conditions for an effective screening programme. This is shown to consist of a suitable disease, a suitable test and a suitable programme or protocol. The evidence relating to colorectal cancer in each respect is presented. Colorectal cancer is a good candidate for screening, although currently medical knowledge is deficient in a number of key areas. Suitable tests are available although none is particularly satisfactory. Suitable programmes can be devised on the basis of clinical trials.

Chapter Three considers some of the problem of evaluating any screening programme by assessing its costs and benefits. Some methodological issues are discussed in detail and the lessons drawn out are applied to a literature review of previous studies of colorectal cancer screening.

Chapter Four introduces data from the Medical Research Council clinical trial. On the basis of the protocol specified a simple numerical model of the costs and disease yield is constructed. This allows the cost of detecting a case of asymptomatic disease to be calculated. The model is then used to identify the most important variables in determining costs and benefits; these are then given more detailed attention in the subsequent chapters.

Chapter Five looks at the effects of screening using a variety of different methods. Data are gathered from a variety of sources in order to compare all of these methods on a common basis using the model constructed in Chapter Four. A sub-set of options which appear more efficient than the rest are identified and the data with which to choose between them presented. The chapter then considers the various possible criteria for a positive screening test and how this will affect the results.

Chapter Six concentrates on the economic aspects of attempts to increase the participation rate in screening for colorectal cancer. A model of individual behaviour using perceptions of costs and benefits is set up and lessons for health education drawn out. Different ways of inviting people to participate in screening are compared. The costs and benefits of pursuing people who do not complete a test are then considered; on the basis of the costs and benefits an optimal participation rate is identified which is shown to differ from the commonly-perceived optimum of 100%.

Chapter Seven extends the idea of benefits beyond the detection of early stage cancer alone. Screening also detects a suspected pre-cursor of cancer, the adenoma, although the precise benefits of detecting and excising these lesions is uncertain. A review of the medical literature reveals a number of estimates and these are compared by applying them to the screening trial data. The significance of variations in the results is discussed.

Chapter Eight considers the question of which age group of the population to screen on economic grounds. Several of the key variables in the screening programme are shown to be related to age, including the proportion returning a test, the proportion of returned tests which prove to be positive, and the number of cases of disease detected (the yield). These factors have various implications for costs and benefits which are explored. The possibility of recommending different protocols for different age groups is considered. The costs and benefits of extending screening to very young and to very old ages is also evaluated. The former may well benefit from detection of adenomas but these are rare. On the other hand elderly people are much more likely to have asymptomatic cancer but are less likely to be able to undergo treatment.

Chapter Nine evaluates the options for diagnostic investigation of positive screening test results. Several options are considered including radiology, endoscopy (a fibreoptic telescope) and combinations of both together. Data are presented for symptomatic patients which allows the various aspects of a model to be constructed. Screening trial data are then used to consider which strategy is most appropriate in a screening programme. Finally the role of retesting of positive results before a diagnostic investigation is performed is considered in terms of costs and benefits.

Chapter Ten considers the impact of screening on the costs of treating colorectal cancer. First of all, the methodological problems in costing hospital care are considered; particular attention is paid to the level of detail required in this type of work using a case study based on allocating nursing costs. Data from a large sample of cases from the MRC trial are then presented and the cost per case and the total costs for each option compared. Explanations are sought for the results.

Chapter Eleven uses the results to date to compare colorectal cancer screening with other ways of treating the disease. A number of alternatives are identified including screening high-risk groups, radiotherapy and chemotherapy following surgery, treatment of advanced disease and better follow-up of people who have previously had an adenoma. The costs and benefits of each are calculated and the results are compared in terms of the costs per life saved. While some of the assumptions are questionable, some lessons can be learnt by purchasers of care.

In the conclusion the policy implications are spelt out. The methodological lessons learnt are also discussed in more detail. The need for further research in particular areas is identified.

Several sections of this work have already been published in peer-reviewed journals. The published references are as follows:

Chapter Four appeared as:
"The costs of screening for colorectal cancer" (with D. Whynes, J. Hardcastle and J. Chamberlain) Journal of Epidemiology and Community Health 1991 45 220-224.
and
"The hospital costs of diagnostic investigations for colorectal cancer" (with D. Whynes, J. Hardcastle and J. Chamberlain) Journal of Clinical Epidemiology 1991 44 907-914.

Chapter Five appeared as
"Filtering strategies in mass population screening for colorectal cancer: an economic evaluation" (with D. Whynes and J. Hardcastle) Medical Decision Making 1992 12 2-7.

and
"Cost-effective screening strategies for colorectal cancer" (with D. Whynes and J. Hardcastle) Journal of Public Health Medicine 1992 43-49.
and
"Rehydration of Haemoccult tests in mass screening for colorectal cancer: an economic perspective" (with D. Whynes and J. Hardcastle) Scandinavian Journal of Gastroenterology 1991 26 215-218.
and
"Hemoccult testing and colorectal cancer" (with D. Whynes and J. Hardcastle) Gastroenterology 1990 99 608.

Chapter Six appeared as
"Participation and screening programmes for colorectal cancer: more would be better?" (with D. Whynes) Journal of Health Economics 1991 10 207-225.

Chapter Seven appeared as
"Cost savings in mass population screening for colorectal cancer resulting from the early detection and excision of adenomas" (with D. Whynes and J. Hardcastle) Health Economics 1992 1 53-60.

Chapter Eight appeared as
"The effect of subject age on the cost-effectiveness of mass population screening for colorectal cancer" (with D. Whynes and J. Hardcastle) Journal of Epidemiology and Community Health 1992 46 577-581.

Chapter Nine appeared as
"Retesting positive results in screening for colorectal cancer: a marginal analysis" (with D. Whynes, J. Hardcastle, M. Thomas and J. Chamberlain) Applied Economics 1991 23 1015-1018.

Chapter Ten appeared as
"Screening and the costs of treating colorectal cancer" (with D. Whynes and J. Hardcastle) British Journal of Cancer, accepted for publication.
and
"The costing of nursing care: a study of 65 patients with colorectal cancer" (with D. Whynes) Journal of Advanced Nursing 1990 15 1305-1309.

I confirm that the following document is my own work. While two of the above papers appeared in print with David Whynes as the first author these pieces have either been extensively reworked (in the case of the Health Economics paper (chapter 7)) or were based entirely on my original work (British Journal of Cancer, chapter 10).

Chapter One

COLORECTAL CANCER

INTRODUCTION

The purpose of this chapter is to introduce the most important features of colorectal cancer. Initially, the costs of cancer and of colorectal cancer are described relative to the costs of other diseases. The causes of colorectal cancer are then explored to assess the chances of finding some means of eradicating the disease (although the economic desirability of pursuing such a policy is not considered at this stage). Factors affecting the prognosis of patients is described. Finally, current treatment practice is assessed and the prospects for improvements in prognosis (reducing the social cost of the disease) are discussed.

THE COST OF CANCER

It is common practice to commence the discussion of screening for a disease by estimating the so-called disease burden in terms of lives lost, working-days lost or resources used in treatment. The figures for cancer are daunting in this respect:

- In 1990, the annual economic burden of cancer in America was $94,400 million or $120 per head of population (Brown (1990)). Premature mortality, valued using average lifetime earnings, contributed 58% of this figure, with 31% resulting from health care costs. Since 1985 the total has risen by 45% in real terms.
- In the UK cancer contributed 6% of the total economic burden of all diseases in the 1970s (Black and Pole (1975)) and by the late 1980s cancer treatment accounted for 7% of all NHS spending, or £1,000 million annually (1986-7 prices) (OPCS (1993)).

These figures should not guide health policy on their own, however, since there is no indication of either the costs or the benefits of action to reduce this burden. Table One shows the number of life-years lost to various conditions in England and Wales in 1990, based on age at death and life-expectancy at that age:

Table One

Cause	Life-years lost (000)
All causes	5,979
All neoplasia	1,803
Lung cancer	401
Breast cancer	217
Colorectal cancer	190
Stomach cancer	88
Pancreatic cancer	68
Prostate cancer	66

Cervical cancer	35
Ischaemic heart disease	873
Cerebro-vascular disease	549

(Source: own calculations based on OPCS (1987), OPCS (1993)).

These data are also presented in Figure One. It is not evident that a major cause of lost life-years, such as lung cancer, is more deserving of attention, since apart from primary prevention there is little effective care available.

Cancer of the colon and rectum

Cancer of the large bowel (or colorectal cancer) has as its primary site the large intestine, a fleshy tube approximately 1.5m long extending from the margin of the anus to the junction with the small intestine.* The burden of colorectal cancer is also heavy:

- 570,000 new cases are reported world-wide each year (Shike et al. (1990)) representing 9% of all diagnosed cancers.
- By the year 2000 it will account for 755,000,000 lost years of life world-wide (Eddy (1986)).
- As noted above, it accounts for 11% of life-years lost to cancer in Britain and 3% of all life-years lost, shown in Figure Two.

The disease has a poor prognosis at present with only about 35% of patients living to five years after diagnosis (Stower and Hardcastle (1985)). Once account is taken of cases never admitted to hospital or never considered for operation the rate may be as low as 27%, however (Slaney (1991)).

The following sections consider the causes of the disease, hoping to establish some means of prevention, and the treatment, to search for advances which will improve cure rates.

CAUSES

The causes of many types of cancer are either unknown or unquantified. Research into the causes of colorectal cancer has concentrated on the influence of inherited susceptibility through genetic defects and the role of environmental factors such as diet.

Genetic factors

Support for a genetic origin to colorectal cancer has come from the study of two groups of patients (Canon-Albright et al. (1988)):

i) those with inherited gastrointestinal syndromes such as familial adenomatous polyposis, a condition producing the growth of thousands of

* The sites within the large bowel include: rectum, recto-sigmoid junction, sigmoid colon, descending colon, splenic flexure, transverse colon, hepatic flexure, ascending colon and the caecum.

Figure One - Life-years lost by major cause

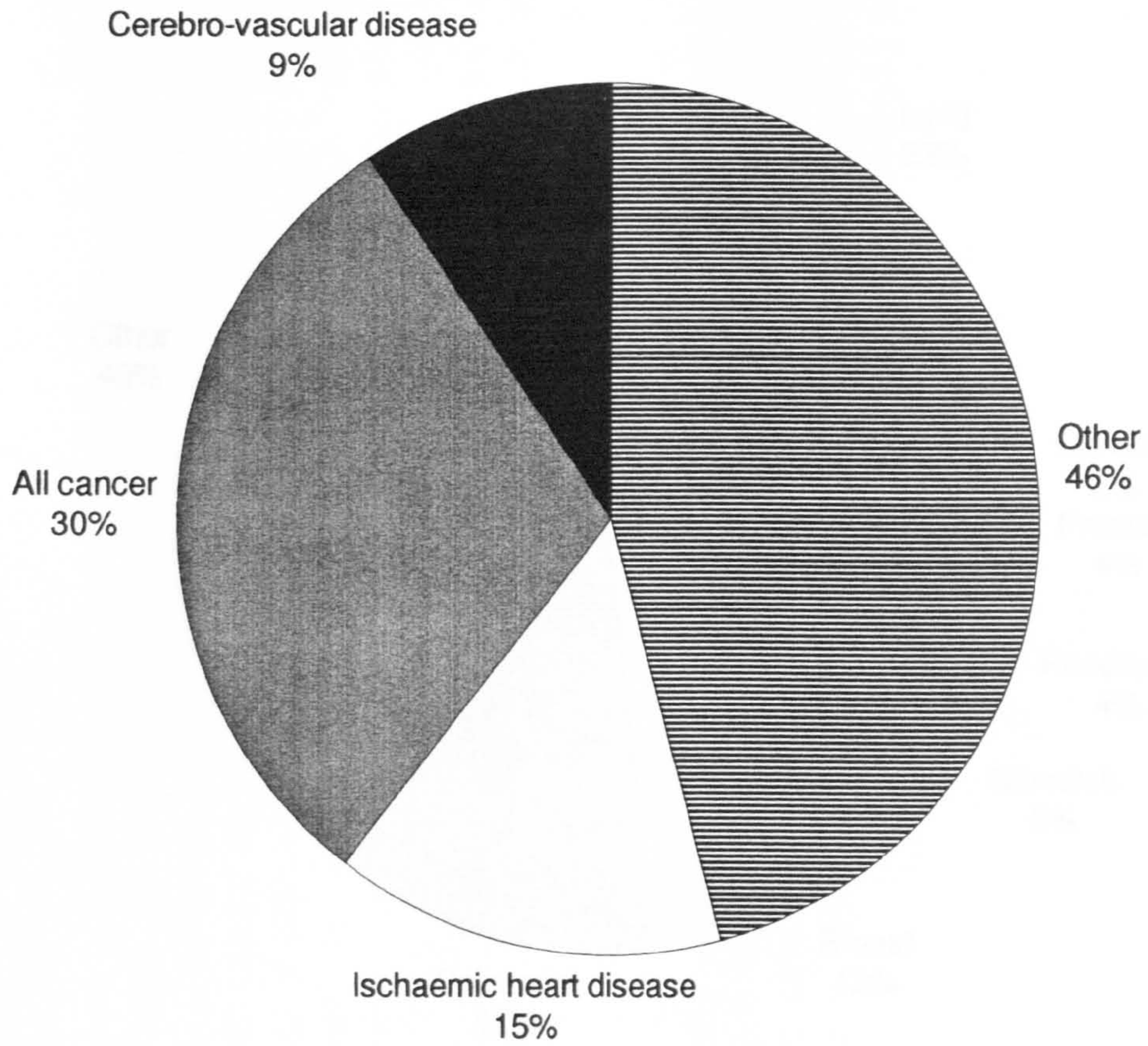
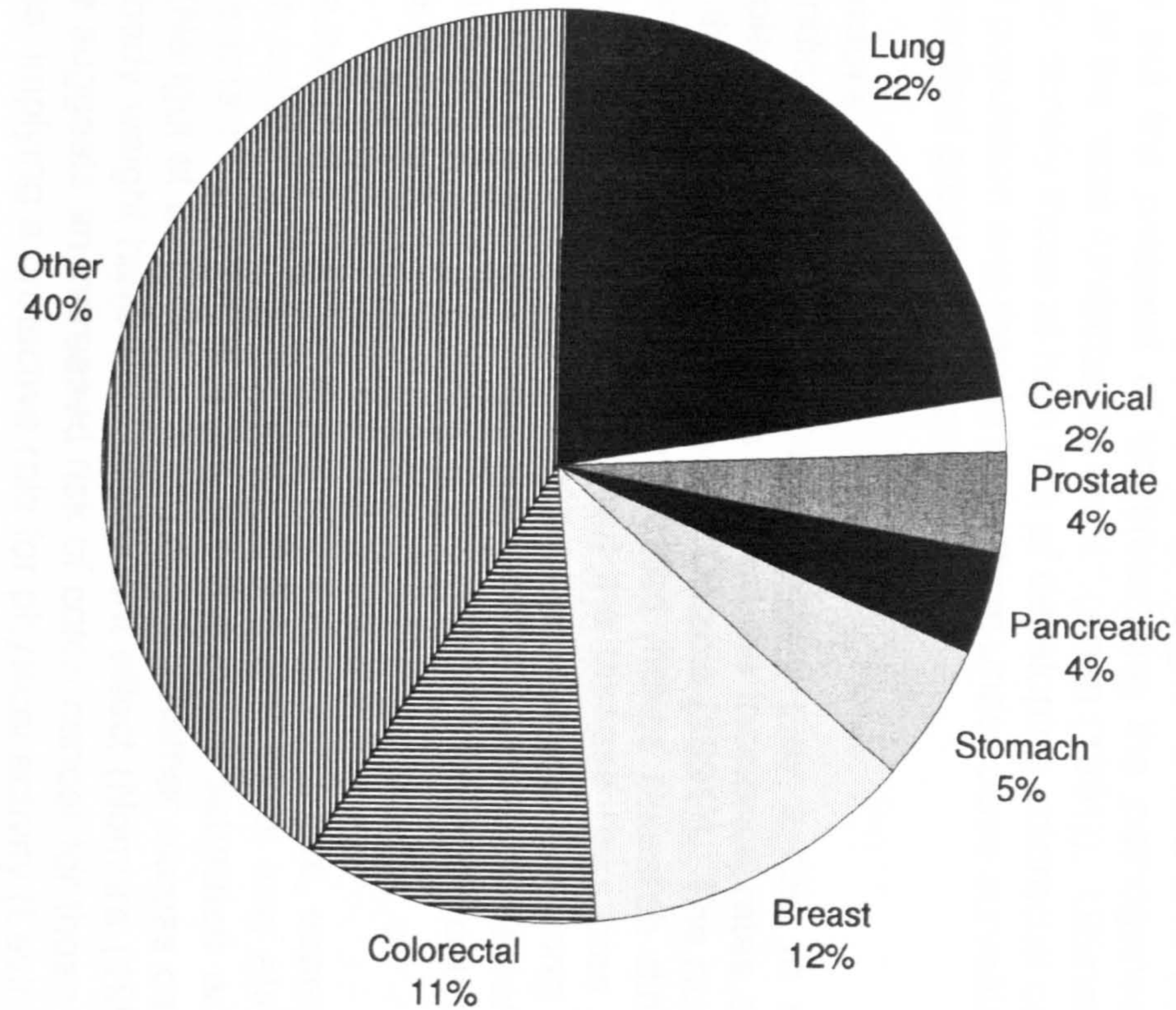


Figure Two - Life-years lost to cancer



small polyps thought to be the pre-cursor of cancer in the bowel, have a very high risk of developing colorectal cancer, and

ii) those with a first-degree relative (i.e. parent, sibling or child) who have had colorectal cancer support are themselves at increased future risk.

The higher risk in hereditary non-polyposis syndromes does not affect spouses implying that environmental factors are not decisive in these cases. A whole range of genetic defects are implicated: based on the hypothesis that cancer arises as the result of up to seven accumulated events, a possible model for mutation has implicated a total of 11 genetic defects in the evolution of metastatic colorectal cancer (Scott and Quirke (1993)). These advances have been described as, "...the beginning of a new and exciting phase that holds out the prospect of understanding the pathogenesis of colorectal cancer at the most fundamental level." (Anon (1991)). Ultimately it may be possible to identify those at high risk of developing colorectal cancer within the general population and then to keep them under close surveillance. This is not yet a practical proposition, however.

Environmental factors

Other evidence indicates that environmental factors play a role in many cases. For example, migrant populations take on the incidence rates of the host population within a generation (Levin and Doizois (1991)). The number of different elements of diet which are implicated makes research difficult. Simple questioning of people with and without the disease risks bias as a result of differences in recall, while prospective studies are long and expensive. The role of a 'marker' in colorectal cancer (akin to cholesterol in ischaemic heart disease) is not understood; this makes identifying causative and protective agents difficult.

Increased fat consumption is associated with increased incidence, especially in combination with calcium deficiency, although the type of fat may also be relevant: vegetable fats may have a protective role while saturated animal fats increase risk (Neugut et al. (1993)). It is unclear whether excess calorie consumption and body weight have an independent effect (Nomura (1990)), but some evidence suggests an increased risk of colon cancer for those with sedentary life-styles, implying a protective role for physical activity (Levin and Dozois (1991)).

Fibre and starch are also regarded as protective agents, but fibre is not a single entity; again, this makes research more complicated. There is limited evidence that fibre consumption is associated with less aggressive disease and a less advanced pre-malignant stage (Armitage (1991)). Overall, though, the link between fibre intake and colorectal cancer has been described as "at best, an oversimplification." (Willett (1989)).

A recent consensus conference concluded, "...the evidence is not strong enough to recommend specific dietary changes." (King's Fund (1990)). However a 'common-sense' approach of increasing the intake of vegetables

and starch while reducing fat and alcohol consumption "will not increase the risk of large bowel cancer and may be of benefit." (Bingham (1990)). Dietary and genetic factors play a part but their role is complex, with many possible interactions.

PROGNOSTIC VARIABLES

Staging

The spread (or stage) of cancer at diagnosis is the most significant variable affecting the prognosis. The pathological classification for colorectal cancer was originally devised by Dukes in 1932 to stage rectal cancer. Cancers limited to the bowel wall are at stage A, those spreading into the surrounding muscle but not involving any lymph nodes* are at stage B, those which have affected local lymph nodes are at stage C, and (a modification to the original proposals) those which have metastasised are stage D. A review of patients presenting to a British teaching hospital during the 1970s gives a typical distribution of stages at presentation, together with the survival rate for each stage (Stower and Hardcastle (1985)):

Table Two

Stage	% diagnosed at each stage	% surviving to five years
A	6	77
B	35	58
C	31	22
D	27	5

Thus, many patients present with disease at such an advanced stage that the prospects of cure are poor. Note, however, that the assumed effectiveness of the treatment of early stage disease is assumed rather than proven: surgical treatment of early stage prostate cancer, for example, appears to offer little survival gain over a policy of monitoring alone (Dearnaley (1994)).

Dukes' stage is an imperfect predictor of prognosis, however (Deans et al. (1992)). While the classification system takes account of *whether* lymph nodes are involved it does not use information on the *number* affected, which may affect prognosis.

Other factors

The size and bowel site of the tumour appear to have little independent influence on survival. Clinical features of the presentation, such as whether or not the tumour is adhering to other organs, bowel obstruction and the age of the patient all have some independent effect (Chapius et al. (1985), Fielding et al. (1986)). In addition, the pathological histology or differentiation

* Lymph nodes can be thought of as access points to the body's lymphatic system, a route which is commonly seen as the means by which malignant disease spreads (or metastasises) to sites distant from the primary growth, such as the liver.

of the disease is significant: this acts as a proxy for the 'aggressiveness' of the disease with poorly differentiated cases being the least amenable to treatment.

In individual cases the surgeon performing the operation may also exert some influence on the outcome: differences in survival rates following a supposedly curative resection have been observed (McArdle and Holt (1991)). This has led to calls for specialists colorectal surgeons to carry out all operations of this type.

TREATMENT

Surgery is the only widely accepted treatment for the disease, although the use of radiotherapy and chemotherapy is being evaluated (and is discussed further in Chapter Eleven). Surgery involves a resection of the bowel, with the affected tissue removed together with a 'safety margin' of healthy tissue (about 2cm. in rectal cancer and 5cm. in colon cancer); the healthy bowel is sewn or stapled together in an anastamosis. The procedure is judged to be potentially curative if there is no evidence of residual disease. There is a risk of operative complications or mortality, commonly due to cardio-respiratory problems or wound infection. Operations for rectal cancer may also damage nerves in the genito-urinary system of male patients causing subsequent problems with urination and even impotence.

Endoscopic excision of the cancer by means of an endoscope (a rigid or flexible fibreoptic telescope) is possible if the tumour is small (less than 3cm. diameter) and has been accurately staged. If the excision is judged to be complete on pathological examination of the specimen surgery may be judged unnecessary.

A different technique is required for rectal cancers very close to the anal verge (usually within 4cm. of the anus). Since there is no safety margin of healthy tissue an anastamosis between two pieces of unaffected tissue is not possible, hence the anal sphincter cannot be safely preserved. The colon is diverted to a stoma (mouth) in the patient's abdomen. Recent advances in technique and relaxation of acceptable safety margins have allowed more rectal cancers to be treated by anterior resection, preserving the patient's quality of life wherever possible. A series of operations for rectal cancer found 40% of procedures between 1978 and 1982 involved a permanent colostomy compared to 21% between 1983 and 1988 (Dixon et al. (1991)).

Resection is also commonly performed as a means of palliating symptoms and avoiding obstruction of the colon where the disease is too advanced to achieve a cure. Operations involving a colostomy are still a common form of treatment in emergencies or for palliative procedures in cases of advanced disease: this prevents tumour growth from obstructing the colon or perforating

the bowel wall. Wedge resection of the liver is used in some centres to remove hepatic metastases but the benefits of this have not yet been proven.

A number of reports have been made on series of patients admitted to hospital with a diagnosis of colorectal cancer. Differences between units, between the way results are reported and possible time trends make some of the data difficult to compare; the following sections present a summary of the main findings from the major British series (Anderson et al. (1992); Brown et al. (1991); Dixon et al. (1991); McArdle and Hole (1991); Stower and Hardcastle (1985); Umpleby et al. (1984)). While the summary inevitably generalises from a variety of studies, the potential for large variations between individual surgeons should be borne in mind (McArdle and Hole (1991)).

Patient characteristics at presentation

The average age of the series of patients is commonly between 65 and 70: one study reports an age range for resections of 25 to 101. The split by gender is fairly even.

The primary site is the rectum (including the rectosigmoid junction) in about 44% of cases, in the sigmoid or descending colon for 30%, in the transverse colon or flexures for 11% and in the ascending colon or caecum for 15%. About 15% of cases are poorly differentiated. Of those with metastases the liver is affected in 74%, the peritoneal cavity in 33%, and the lungs in 5%; the brain, skeleton, adrenal gland and ovaries can also be affected but these are rarer.

Type of admission

Between 30% and 40% of admissions are emergencies. These patients tend to be older than average with advanced cancer of the colon; surgery is also more likely to be undertaken by a junior doctor. Very few rectal cancers require emergency admissions.

Treatment received

Historically, over 20% of patients were never operated upon but recent trends are to perform more palliative resections particularly in elderly people with advanced disease. Between 5% and 8% of admissions do not undergo an operation either because of obvious terminal disease or because they die before any procedure is possible; the surgery rate is age-related. A similar number are found to have unresectable tumours at laparotomy, while a further 10% have only a partial resection (again, owing to obvious advanced disease). About one-third of patients have operations requiring a colostomy, either temporarily or permanently.

Between 50% and 60% of completed resections are regarded as curative: the figure is related to the age of the patient, whether the admission was an emergency, and the differentiation of the tumour. The mobility of the tumour

plays a part here: half the operations performed are on cancers which are fixed to other structures in the abdomen, implying spread to other organs and a reduced chance of a cure.

Operative mortality

About 20% of patients die in hospital following their first admission with the disease; again this is age-related. For those who have an operation the rate is between 10% and 15%; while most of these cases are due to pre-existing conditions (particularly cardiac and respiratory disease), 27% of deaths in one series were attributable to surgical complications.

Other complications

The complications of major surgery can be very serious, both for the health of the patient and in terms of prolonged hospital stay. The main problem in comparing reports of complication rates is variation in the definition of a complication and the thoroughness with which such data (which is not routinely available) are collected.

Anastamotic dehiscence (defined as clinical evidence of a leak with resultant fistula formation, sepsis, or both) occurs in between 6% and 12% of cases, with wound dehiscence in less than 1%. The prognosis in such cases is poor with 59% dying before discharge in one series. Wound infection rates vary from 8% to 22%. Following operations for rectal cancer 2% of patients require an operation on their prostate and 1% suffer from impotence (Dixon et al. (1991)). Other complications include chest infection (14%), pneumonia (5%), deep venous thrombosis (3%) and urinary tract infection (11%).

Recurrence

Even following resections which are judged to be curative disease can recur, due to small amounts of residual disease being missed at the operation. "Local recurrences are an important reflection of surgical technique in cancer treatment and produce the worst form of death for patients, producing pain, incontinence, discharge, etc." (Dixon et al. (1991)). The proportion of curative resections who develop recurrence varies between series but one review suggests bounds for the rate of 15% to 30% (Pollard et al. (1989)), although figures of up to 47% have been reported (Umpleby et al. (1984)). Most recurrences present within two years of surgery, at a median of 22 months in one series (Devesa et al. (1988)).

Recurrence is found in 4% of stage As, 13% of stage Bs and 18% of Cs (Phillips et al. (1984)); 11% of well differentiated tumours recur, compared to 14% and 21% of those that are moderately and poorly differentiated. Stage and differentiation interact, hence a poorly differentiated stage C tumour is at particular risk. 'Cured' rectal cancers appear to be at slightly increased risk of recurrence compared to cancers of the sigmoid colon (Umpleby et al. (1984)).

CONCLUSION

Colorectal cancer is a major cause of ill-health and a significant cause of premature mortality. The ability to cure disease in its earlier stages implies that the potential gains from earlier detection are large. Treatment advances do not solve the problem of late-stage presentation. Prevention of the disease may be possible at some point in the future but is not yet a realistic option. This leaves screening as the only immediate prospect of reducing the social cost of colorectal cancer.

Chapter Two

SCREENING FOR COLORECTAL CANCER

Introduction

Screening is defined as, "[T]he identification, among apparently healthy individuals, of those who are sufficiently at risk of a specific disorder to justify a subsequent diagnostic test or procedure, or in certain circumstances direct preventive action." (Cuckle and Wald, quoted in Holland and Stewart (1990)). Mass screening relates to the screening of entire groups of the population as opposed to high-risk groups.

The purpose of this chapter is to introduce the principles of screening for disease and the practical application of these principles to programmes for the detection of colorectal cancer. The conditions for effective screening are described and the evidence regarding colorectal cancer presented.

Screening and effectiveness

An effective screening programme will result in the detection of early stage disease, improving the prognosis and survival rate. Effectiveness can sometimes be taken for granted, however: "[T]he act of screening runs the risk of acquiring respectability almost by virtue of its existence." (Holland and Stewart (1990)). This has led to an unwarranted assumption in favour of preventative medicine in advance of proper evaluation. A leading breast cancer screening researcher, for example, expressed her concern that doctors were brainwashing themselves into thinking that they were having a major impact upon the disease before they brainwashed the public (Roberts (1989)). Such assumptions have worked their way into medical training according to the experience of one American doctor: "I recall as a medical resident asking whether a patient needed sigmoidoscopy and being asked in return, 'Does the patient have a rectum?' " (Grey (1991)).

Others have commented upon the ethical imperative to prove effectiveness before implementing screening (McKeown in Gyde (1990)). It is argued that the doctor cannot be held to blame for the state of medical knowledge when the patient approaches him with symptoms of disease requiring treatment: it is only incumbent upon the doctor to do as well as they can, subject to constraints. In screening, however, doctors approach healthy members of the public offering a test which (implicitly) has some value. The onus should thus be on the medical profession to be able to quantify that value.

Conditions for effective screening

Public health physicians have described a number of criteria for a disease screening programme. These are not always clearly defined, nor are they based on economic principles. The following list covers the most commonly

mentioned factors (from Wilson and Junger plus Cuckle and Wald quoted in Holland and Stewart (1990)):

i) factors relating to the disease
   - the disease should represent an important health problem
   - the natural history of the disease should be understood
   - the disease should have a recognisable latent or early symptomatic stage

ii) factors relating to the screening test
   - the test should be acceptable
   - the test should be accurate
   - the test should be inexpensive
   - the distribution of test values in healthy and in ill people should be known and the cut-off for a positive test defined so that the extent to which people with and without the disease are misclassified is minimised

iii) factors relating to the screening programme
   - sufficient diagnostic and therapeutic facilities should be available
   - there should be an effective and widely accepted treatment protocol
   - there should be an agreed policy on who to treat

Despite doubts about the value of these criteria, they form a useful checklist for discussing aspects of colorectal cancer with regard to the possibility of screening.

1. IS COLORECTAL CANCER A SUITABLE DISEASE FOR SCREENING?

The burden of colorectal cancer makes it a major source of ill health and reduced life expectancy. The natural history is suitable for intervention: disease develops in a recognised pre-malignant condition (the adenoma) and is initially localised before spreading to other sites. While this broad description is widely accepted, many of the parameters in the process are the subject of debate. The role of the adenoma is a good example.

Adenomas

Polyps are small, fleshy growths protruding from the bowel wall. Pathologists classify polyps according to their cell structure: those suggesting neoplastic alteration of the epithelium (dysplasia) are known as adenomas. The histology of an adenoma relates to its cell structure: tubular, villous and tubulovillous have varying proportions of cells which are either finger-like or tubular when viewed through a microscope. The histology and dysplasia are significant because they are thought to influence the malignant potential of an adenoma.

In support of the so-called adenoma-cancer sequence there is genetic evidence of common chromosome defects in each case (Anon (1992)). Familial adenomatous polyposis, a rare inherited disorder which causes

countless adenomas to grow in the large intestine, almost invariably results in cancer. Follow-up studies of patients with a previously excised adenoma imply that there is an increased future risk of cancer. Epidemiological support is also available: the distribution of adenomas and cancers is correlated both internationally and across ages within a population. In addition, there is a tendency for both types of neoplasia in young people to be mainly in the left-side of the bowel while older people have a higher prevalence of each on the right side. Clinical evidence points to a similar conclusion: more than half of resected cancer specimens contain residual adenomatous tissue, while observation of unexcised adenomas has revealed cancer subsequently developing at the same site (Morson (1984), Rawlinson et al. 1989), Stryker et al. (1987)). In one large study the removal of rectal adenomas was claimed to reduce the future incidence of rectal cancer below that expected (Gilbertsen and Nelms (1978)), but this finding is controversial (and is discussed further below).

Thus, while there is no conclusive proof that cancers arise in adenomas, a large body of circumstantial evidence exists. No other theory has gained any widespread acceptance. However, although cancers arising directly from the mucosa (de novo) are rare there are some reports of small (5-15mm. diameter) 'button' cancers with raised edges and a depressed centre, with no evidence of a pre-malignant stage. Pathological evidence that more than half of cancer specimens contain residual adenomatous tissue implies that, at most, only a minority of cases arise in this way. However, some researchers have claimed that these could represent a more aggressive, clinically significant, disease while polypoid cancers, which grow into the lumen (the actual space within the walls of the intestine) rather than into the bowel wall, are less aggressive. One study has estimated that 15% of cancers miss out the adenoma stage of progression (Koretz (1993)) although these findings have been questioned (Simon (1993)).

It is widely agreed that many adenomas remain benign and pose no health risk. While cancer prevalence (the proportion of the population with a condition) is thought to be about 3 per 1,000 in people aged over 50, adenoma prevalence is as follows:

18% in under 55s
29% in those aged between 55 and 64
40% for those aged 65 to 74
41% for those aged 75 and above (Williams et al. (1982)).

Estimates of the exact proportion of adenomas destined to become malignant vary considerably from 1% up to 10% (Hoff (1987)). The length of the sequence is also uncertain with estimates ranging from 'at least five years' (Simon (1985)) to 'up to 18 years' (Frank (1985)). The variation between individual cases is still less clear.

This brief review of the uncertainty surrounding the important aspects of the disease process. Similar doubts surround the progression through the stages

of malignancy. Colorectal cancer conforms to the broad description of a suitable test but several important pieces of data are missing.

2. IS THERE A SUITABLE SCREENING TEST FOR COLRECTAL CANCER?

The accuracy of a screening test is described by its sensitivity and specificity. The former is the proportion of people who have the disease who test positive; the latter is the proportion of people without the disease who test negative. The two types of information provided by a screening test may ultimately prove to either have been true or false, although this is not observable at the time. There are thus four outcomes to screening: diseased and positive ('true positive'); healthy and negative ('true negative'); healthy and positive ('false positive'); and diseased and negative ('false negative'). These are considered in turn below.

The vast majority of people completing a screening test will be negative for the disease. This may offer them some reassurance, although it has been pointed out that an individual has only a small chance of developing the disease so realistically this is unlikely to be very important (Frank (1985)). For a small number the reassurance will be false, either because the screening test has failed to detect the disease or because disease develops in the interval between screens. The negative test may cause them to ignore symptoms of the disease if and when these occur, although it could be argued that the screening process will make them more aware of the disease and hence cause them to react more quickly.

A minority of people have positive test results, which are likely to cause anxiety. For many diagnostic investigation will confirm that the cause of the positive test is benign or that there is no evidence of any bowel disease at all. While this may reassure some, others will have persistent anxiety and possibly depression: this effect is also known as labelling (Marteau (1989) and (1990)). Only those who have disease detected and treated at an earlier stage than if they presented symptomatically can derive a health benefit; however, this group may also suffer psychological sequelae.

Sensitivity and specificity can be expressed algebraically and the relationship of each to variables such as the positive rate explored. Supposing that there is a target population T of whom A accept and X is the rate of participation i.e. A=TX. The prevalence rate in the general population is R and the rate among acceptors is R_A. The total number of cases prevalent in the population is C which equals TR while the number in the accepting population, CA, is TXR_A.

Screening yields P positive results of which P_T are true positives and P_F are false positives; similarly of the N negative tests N_T are true negatives and N_F are false negatives. By definition $A=P_T+P_F+N_T+N_F$ and $C_A=P_T+N_F$.

The sensitivity rate of the test, S_N is defined as P_T/C_A, thus $P_T=C_A*S_N$. The specificity rate S_P is defined as $N_T/(A-C_A)$ so $N_T=(A-C_A)*S_P$.

Expressions for N_F and P_F can also be derived:

$$N_F=C_A-P_T \text{ or } C_A-(C_A*S_N)$$

$$=C_A(1-S_N)$$

and

$$P_F=A-P_T-N_T-N_F$$

$$=A-(C_A*S_N)-S_P(A-C_A)-C_A(1-S_N)$$

$$=A(1-S_P)-C_A(1+S_P)$$

Thus, while trial results can only be used to estimate sensitivity and specificity if a definitive test has also been used, estimates of these parameters can be converted into predicted trial outcomes if prevalence estimates are available.

Available tests for colorectal cancer

Three means of screening for colorectal cancer have attracted more attention than any other:

- direct visualisation of the large bowel (endoscopy);
- surveying the population for bowel symptoms; and
- testing for blood in samples of faeces.

Other methods such as testing for raised levels of carcinogenic antigens in blood samples has also been proposed but has been found more useful in surveillance after treatment. Digital rectal examination extends up to 10cm. from the anal verge and can detect approximately 10% of all colorectal cancers, as well as prostate cancer; however, it is part of a fuller health check-up rather than a screening test in its own right. One possible screening test for the future has been described as follows: "In Imax's showrooms in Ginza, Tokyo, a prototype 'intelligent toilet' is on display, which can measure the user's weight, temperature and blood pressure (don't ask me how) and carry out up to eight different tests on urine and faeces; an electronic link to the nearest hospital will transmit all abnormal findings...By the year 2000, we shall all strain on lavatory bowls more intelligent than ourselves." (Skrabanek (1991)). In the absence of further data this option has not been evaluated but the results of trials are eagerly awaited.

Sigmoidoscopy screening

Endoscopy is very accurate since direct viewing of the bowel combined with biopsying of suspicious lesions means that the examination will be sensitive

and specific. The extent of the examination is often limited, however, and only a portion of the bowel can be viewed. There is also a small health risk involved through post-investigation haemorrhaging, perforation of the bowel wall and other complications. Since it is a hospital-based investigation there are also concerns about its cost.

The two main types of endoscope considered are the short rigid instrument and the longer flexible instrument.

i) Rigid sigmoidoscopy screening

Rigid sigmoidoscopy can visualise the rectum up to the junction with the sigmoid colon some 25 cm. from the anal verge, covering the primary site of 40% of colorectal cancers (Stower and Hardcastle (1985), Umpleby et al. (1984)). However, the instrument is only fully inserted in 50% of cases and the average view is to about 17cm. In common with other types of endoscopy, the procedure is uncomfortable for the patient and perforates the rectum in 1.4 per 10,000 investigations (Selby and Friedman (1989)) and results in the death of the patient in about 1 in 10,000 (GOICC (1990)).

Some uncontrolled trials have claimed that regular rigid sigmoidoscopy can reduce the incidence of rectal cancer (Gilbertsen and Nelms (1978); Dales et al. (1979)). Doubts have been expressed about the calculations made, the thoroughness of follow-up, and the high proportion of the control group undergoing screening.

A case-control study of the effects of rigid sigmoidoscopy screening suggested a 59% reduction in the risk of dying of cancer sited in the area of the colon covered by the instrument (with a confidence interval of 31% to 75%) (Selby et al. (1992)). The number of new cases is low following a sigmoidoscopic examination; the reduction may last for up to ten years.

ii) Flexible sigmoidoscopy screening

Up to 70% of cases of colorectal cancer are found within 60cm. of the anal verge. The instrument is fully inserted and the view adequate in 80-90% of cases; perforation occurs in less than 1 in 2,000 investigations (Selby and Friedman (1989)). It is superior to rigid sigmoidoscopy in many respects. However, it requires fuller bowel preparation in advance, and is more expensive because of the more complex equipment required.

iii) Conclusion

Sigmoidoscopy has proven value in detecting neoplasia. However, claims relating to mortality reductions are based on clinical trials which are usually regarded as suggestive rather than conclusive owing to the possibility of biases in the results. American and Canadian working parties on screening both felt unable to recommend screening by this means in the absence of more proof of effectiveness. However, other American organisations, such as

the American Cancer Society, feature sigmoidoscopy in their recommendations to physicians on screening.

Screening using the symptom questionnaire

Typical symptoms leading to a diagnosis of colorectal cancer include rectal bleeding, change in bowel habit, weight loss and abdominal pain. One option is to improve public awareness of the potential significance of such symptoms. However, overcoming the social mores surrounding bowel function and problems is difficult, particularly in the elderly population. Another problem is the frequency of the symptoms in the general population after middle age: about 1-in-6 in these age groups experience some rectal bleeding each year and 1-in-24 GP consultations is for bowel disease (Hennigan et al. (1990), Jones and Dudgeon (1992)). It is possible that significant anxiety about such symptoms could be aroused for little reason.

Screening based on these factors by postal questionnaire has been attempted and found to be severely hampered by the high prevalence of such symptoms in the target age groups. One study had a 23% positive rate and a positive predictive value* for any neoplasia of only 5% (Pye et al. (1988)). Some symptoms have proven more useful than others: the best positive predictive power is from a combination of dark red rectal bleeding and diarrhoea (Silman et al. (1983)). In the light of the evidence, however, this option is rarely considered for mass screening.

Faecal occult blood testing

As noted above, many middle-aged and elderly people experience small amounts of bleeding from the rectum. Some blood is deposited on to stools as they pass through the bowel; the median quantity of haemoglobin (or Hb.) found in healthy individuals has been measured at 0.72mg. of Hb. per gramme of stool (Ahlquist et al. (1985)). Cancers of the bowel also deposit blood onto stools, but in greater quantities; the median blood loss in patients diagnosed as having cancer is 6mg. Hb. per gramme of stool (Ahlquist et al. (1985)). By the time bleeding occurs in visible quantities the disease is often advanced. One means of detecting cancer at an early stage is to test the stools of asymptomatic subjects for bleeding in excess of normal 'background' levels. The mean blood loss in subjects with asymptomatic cancer has been measured at 3.3 mg Hb per gramme of stool (Ahlquist et al. (1989)), implying that this is technically feasible.

However, even a perfect test for faecal occult blood (FOB) would be an imperfect test for colorectal cancer:

* The positive rate is defined as the number of people testing positive divided by the number completing the test. The positive predictive value is defined as the proportion of those with a positive test who are shown to have the disease on the basis of diagnostic investigation.

i) a positive test may have a cause of little or no clinical significance e.g. haemorrhoids. Elements of a normal diet can also cause a positive test such as red meat and some types of vegetables and fruit.

ii) not all cancers (and few adenomas) bleed in excess of the level commonly regarded as 'normal' in their asymptomatic stage to test positive, and even those that do so may bleed intermittently over time and/or deposit blood only in a few areas of the stool.

This implies test sensitivity and specificity of less than 100%.

FOB testing as a means of mass population screening was first seriously contemplated in the early 1970s when the Haemoccult test became available. It consists of a series of guaiac-impregnated squares of filter paper mounted on a single piece of cardboard. It is completed by the patient at home by taking a pea-sized sample of faeces and smearing it on a square. The test has six squares and is completed over three days to allow for intermittent bleeding. The completed test is returned for laboratory development, where two drops of hydrogen peroxide, a reagent which reacts with small quantities of haemoglobin, are added. If the sample develops a blue tinge within thirty seconds then it is judged positive; unfortunately this fades within a few seconds so that 'on-the-spot' interpretation is required and a second opinion is ruled out.

i) Sensitivity to cancer

When a new screening test is developed, a simple means of assessing its sensitivity is to test it in symptomatic cancers: for three-day testing 70% of such cases were positive (Farrands and Hardcastle (1983)), although this figure is site-dependent (Leicester et al. (1983)). The problem with this method is that people who have symptoms have higher levels of rectal bleeding than asymptomatic cases, hence the figures are over-estimates of the figure for an asymptomatic population.

It has been found that, for a given haemoglobin concentration, wet stools are six times more likely to test positive than a dry one (Ahlquist et al. (1985)). As mass population screening relies on the postal return of completed tests, delays may mean that weak positive reactions could give negative results (Macrae and St John (1982)). One technique is to rehydrate the stool sample with a drop of water prior to development. This makes the test very sensitive when applied to symptomatic cancer (Macrae and St John (1982)).

ii) Sensitivity to adenomas

The overall sensitivity for adenomas has been estimated at 28% (Macrae and St John 1982); Rex et al. (1991)). Other characteristics are also relevant: large, pear-shaped, villous, severely dysplastic adenomas on the left side of the colon are most likely to be detected (Gabrielsson et al. (1985)).

Evaluation of screening tests

Clinical trials are the traditional source of data on the effectiveness of screening. In one form of trial, the uncontrolled study, a sample is recruited, offered screening and followed over time. Results in terms of stage at detection and 5-year survival rates can be compared with data on groups of patients presenting with symptoms. This evidence allows conclusions to be drawn about the technical feasibility of screening; however, it is insufficient proof of the effectiveness of screening because of the problems in uniquely attributing the observed effects to screening, owing to other sources of bias in the results.

Lead-time bias describes the effect of detecting a disease at an earlier stage in its progression but not affecting the prognosis. It has been likened to getting on a train one station earlier to get off at the same destination. Consider two alternative scenarios. Under one, the disease causes symptoms when the patient is age 67, it is diagnosed and treated, and the patient dies at 70. In the other, the disease is detected on screening at age 65, it is treated and the patient dies at age 70. A measure of outcome like the five-year survival rate will be improved by screening but the effect does not benefit the patient, and may even cause harm.

Length bias relates to variations in the growth rate and aggression of cancers of a single site. Slow-growing cancers will have a longer asymptomatic stage than aggressive cancers. Screening will, therefore, detect a high proportion of slow-growing cancers, since these are present for so much longer. If these cancers would have remained asymptomatic until the individual dies of some other cause then there is no health benefit to diagnosis. Early detection is of little value either in cases with very aggressive, fast-growing cancers (likely to recur even if treated) or in cases with very slow-growing cancers.

Bias may also be introduced by using healthy volunteer populations in trials. This group may take more care of their health than the general population, making screening appear more beneficial and making generalisation of the results to the general population difficult.

Some of the problems in interpreting the evidence from uncontrolled studies can be illustrated by the German national programme. 92% of the population of that country is covered by a statutory health insurance scheme, including a cancer screening programme which has offered digital rectal examination for people aged 45 and above since 1971 and annual Haemoccult screening since 1977. Data protection legislation makes linkage of screening and cancer incidence records impossible; there is incomplete reporting of screening and diagnostic test use; and adenoma excision is not recorded. This makes the effectiveness of the programme very hard to evaluate. The incidence of unresectable rectal cancers has been falling since the early 1970s and colorectal cancer mortality has been falling since 1979, but a more health-conscious population and improvements in diagnosis and treatment

could account for this. Contrary trends in neighbouring countries indicate that screening has some value but the decline in mortality (starting only two years after introduction of Haemoccult screening) implies other factors are also at work (Robra (1986)).

The flaws in evidence from non-randomised studies has not deterred sweeping recommendations on the basis of potentially misleading results: "Evidence of a decline in the proportion of advanced-stage cases should be viewed as an acceptable analogue for mortality reduction. If screening leads to a higher frequency of earlier stage cases or reciprocally fewer late-stage cases, a good argument for the efficacy of screening can be made." (de Cosse (1988)). The biases in the evidence listed above imply a more cautious view: "The current knowledge base, including open questions of efficacy can support a range of recommendations." (Knight et al. (1989)). A similarly misleading claim is the following: "...within a potential group of patients offered regular colonoscopic screening there may be, say, 10 patients who were otherwise going to develop bowel cancer and die of it. One would strongly suspect that with regular colonoscopic surveillance at least three of these would be diagnosed at either the adenoma or early, curable carcinoma stage and thus result in a 30% decrease in mortality." (Meagher and Stuart (1992)). This assumes: full participation; no complications or mortality as a result of screening; adenoma excision prevents cancers occurring in the future; asymptomatic disease can be detected at an early stage; advancing the stage of detection increases the cure rate; and that people live longer as a result.

Randomised controlled trials (RCTs) minimise the bias in results by randomly allocating a population to two groups, one of which will be offered screening and one which will not. These populations should then be identical in all respects other than the offer of screening and any observed differences in mortality rates can be attributed to screening with some confidence. However, such studies are also time-consuming and expensive. Evidence with regard to colorectal cancer screening is considered in the following section.

3. CAN A SUITABLE PROGRAMME FOR COLORECTAL CANCER SCREENING BE DEVISED?

Two aspects of a suitable programme are considered in this section. The first problem is to consider the relative risk of the disease in various sub-groups of the population in order to decide who will be offered screening. The second problem is to define a suitable screening protocol. To address this issue some of the options used in the on-going clinical trials of screening are described.

Risk factors and screening
One of the main requirements for a screening programme is to identify a suitable group for screening. Ideally this will be on the basis of a set of factors which identify a high-risk group leaving the unscreened population at as low a risk as possible.

Sub-groups of the population can be defined as being at increased risk of developing the disease. The following classification of risks has been proposed (Fleischer et al. (1989)):

Markedly increased risk
- Cancer families - inherited susceptibility to particular types of cancer
- Familial adenomatous polyposis - hereditary condition producing thousands of adenomas in the large bowel
- Extensive ulcerative colitis - much of the bowel wall shows some form of dysplasia

Moderately increased risk
- First degree relative with colorectal cancer
- Previous neoplasm in large bowel
- Limited ulcerative colitis
- Women undergoing irradiation for gynaecological cancer

Probable increased risk
- Previous cancer of either the breast or of a gynaecological site
- Previous ureterosigmoidostomy

Nevertheless, these groups cover only about 20% of cases of colorectal cancer (Jarvinen and Mecklin (1989)). To make a major impact on mortality from the disease, a more general risk factor must be considered.

Screening by age group
Cancer is essentially a disease of old age: even if it were to be completely eradicated average life expectancy would only rise by a little over two years (Eddy (1981)). Table One shows the number of new cases in England and Wales in 1987 and the number of deaths from colorectal cancer in 1991 (these are the most recent figures in both cases). Also included is a breakdown of the life-years lost as a result of deaths in each age group, calculated according to the life expectancy at the midpoint of each five-year age range (LYs denotes life-years lost):

Table One

Age	Cases	Deaths	LYs lost	% of total
<40	310	93	3,952	2
40-4	354	166	5,599	3
45-9	576	344	10,080	5
50-4	977	584	14,362	8
55-9	1,739	922	18,906	10
60-4	2,695	1,600	26,795	14
65-9	3,512	2,351	31,573	17
70-4	3,930	2,710	28,266	15
75-9	4,432	2,876	22,891	12
80+	6,340	4,481	27,154	15

(Source: own calculations based on OPCS (1987), OPCS (1993)).

Of all life-years lost to colorectal cancer, 18% are before the age of 55, but each five-year age-group carries an equal share of the total thereafter as rising incidence is offset by declining life-expectancy.

Little is known about the prevalence of the disease in the United Kingdom, but, on the basis of American estimates (Feldman et al. (1986)), there are over 246,000 asymptomatic cases in the population of England and Wales, slightly more than half of these being in the 50 to 75 year old age group. The prevalence is ten times the incidence, implying a long pre-symptomatic stage to the disease. This is confirmed by estimates of colorectal cancer growth rates, which suggests that a 70-year old presenting with symptomatic disease may have developed the first mutation while aged about 40 (Spratt and Spratt (1985)).

On this basis most protocols begin screening at around 50 years of age. The upper limit depends upon ability to undergo treatment if asymptomatic disease is detected; this implies stopping screening at about 75 years of age.

Suitable screening protocol

Much of this evaluation is concerned with defining the optimal screening protocol for colorectal cancer. However, in order to consider what alternatives are used in practice, the various clinical trials are described in more detail. The description of the MRC trial also sets the scene for the following chapters.

MRC trial, Nottingham (Hardcastle et al. (1980), (1983), (1986), (1989))

As the present economic evaluation concentrates on data from the Medical Research Council (MRC) screening trial many results are discussed in the following chapters. Nevertheless, some indication of the history and structure of the trial and a review of the main results will serve as an introduction.

The MRC trial is based in the Department of Surgery at the University Hospital in Nottingham. The grant is held jointly with the Department of Health Cancer Screening Evaluation Unit. A pilot study commenced in 1981 randomising 20,525 people aged 45-74, with a Haemoccult test posted to the study group. The early participation rate was only 38%, but measures were introduced to improve this such as reminder letters to non-responders. In 1983, an extension compared Haemoccult with another FOB test, Fecatwin/Feca-EIA*, in a further 6,450 subjects. However, the positive rate of the latter test was 8.1% which was judged to be unacceptably high given the limited diagnostic resources available.

The improvement in the stage distribution of cancers detected by screening was adequate to justify funding for a full trial which commenced in the second half of 1984. The evaluation is of an offer of screening to the asymptomatic population between the ages of 50 and 75 by the Haemoccult II test. The aim is to reduce mortality from colorectal cancer in the group offered screening by 30% compared with that in the control group. The offer of screening is repeated to those who participated in the previous round of screening only and this is at an intervals of two years after the previous round of screening.

It has been calculated that 156,000 people will have to be recruited and followed-up for a minimum of seven years in order to show a significant gain on the basis of the assumptions used. The main data source is general practitioner records. Consent is sought from each practice to identify all those in the age limits specified, with participating GPs sent the list of eligible patients to exclude ineligible patients (e.g. those with a previous bowel cancer). Those remaining are then randomised and an offer of screening is made to the study group; only 2.3% of invitations are returned unopened by the GPO.

One peculiarity of this trial is that the sample is recruited over time, thus spreading the workload involved but raising the possibility of time trends affecting the results. The length of follow-up is also difficult to calculate without using statistical techniques.

The control group are not contacted in any way and do not appreciate that they are part of the trial. The study group are sent an offer of screening which consists of a standard letter signed by the individual's GP (letters directly from the screening unit had a lower participation rate), a Haemoccult test, instructions on how to complete it, and an SAE to return completed tests to the Department of Surgery. The letter explains that, while the individual has no current bowel disease, screening is advisable since early treatment has a better chance of success than waiting for symptoms to appear. The GP

* The characteristics and performance of this test are discussed in more detail in Chapter Five.

therefore recommends that the individual complete and return the enclosed test in the envelope provided.

Development of the test does not include the rehydration method mentioned above to counter any delays in the postal system. The criterion for a positive test is that any one of the squares on the Haemoccult test is positive when developed, although the exact number of positive squares is also recorded. An invitation is sent to those who have returned a positive test to attend a special out-patient clinic. This is an opportunity for the individual to discuss the test results and their potential implications. The clinical aim is to assess their suitability for further investigation; this includes taking a full medical history and any family history of colorectal disease. An abdominal investigation is then made followed by rigid sigmoidoscopy; flexible sigmoidoscopy is not used owing to the increased bowel preparation required. A blood sample is taken for routine tests. The patient is then offered a full diagnostic investigation.

At the start of the trial diagnostic investigation was by means of the 60cm. flexible sigmoidoscope in combination with the double-contrast barium enema X-ray (BEXR). The excision of adenomas seen at radiology beyond 60cm. was by means of the limited colonoscopy facilities available at that time. The performance of BEXR was poor, however, despite the expertise of the consultant radiologist. Of seventeen cancers detected in the early stages of the trial following a positive screening test six were negative on radiology including four early stage cases. As more facilities have become available colonoscopy has become the favoured means of investigation with radiology reserved for the small numbers of patients who are unfit for sedation.

Adenomas and polypoid cancers are excised at colonoscopy using snare diathermy (or polypectomy) wherever possible, with surgery reserved for larger tumours. Follow-up of all those recruited uses several routinely available data sources, including the records of the hospital's pathology department, the regional cancer registry and the NHS Central Registry in Southport. In addition, many local GPs are aware of the trial and notify the trial administrator of recent deaths from colorectal cancer among their patients.

An interim report on the results of the first 107,344 patients recruited was made in 1989, although it is clear that the screening protocol has been subject to 'fine-tuning' and thus the results are not representative of any one means of testing. In total 53% of people offered a test complete one with 2.3 cases detected per 1,000 acceptors. Cancers detected on screening are at an earlier stage than those presenting in the control group, are more clearly differentiated, more mobile and more amenable to endoscopic polypectomy. Participation with rescreening is high: 77% of initial acceptors accept the offer of rescreening, while 80% who have completed two tests complete a third.

The impact of dietary factors on the positive rate of the test has been discussed but including a restricted diet in the instructions for completing the test has been shown to adversely affect the participation rate. To tackle this problem, the MRC trial retests people returning a positive test prior to a full diagnostic investigation. This involves completing a second Haemoccult test, this time with dietary restrictions on intake of red meat and vegetables high in peroxidase; those testing positive are investigated. Those retesting negative are sent a third test three months later with positives being investigated; however, negative results are taken to imply that the initial positive test was false. The benefits of retesting are the maintenance of an acceptable participation rate without swamping the diagnostic facilities available; the cost is a small loss of sensitivity; the economic perspective is presented in Chapter Nine.

As part of the trial almost 35,000 people randomised to receive an offer of screening were further randomised to be tested over either three or six days. While six-day testing had a higher yield of neoplasia the difference was not significant, although the false positive rate and cost were higher and the participation rate was lower; an economic comparison of the two tests is contained in Chapter Five.

The trial has now completed recruitment and the initial offer of screening; most of the study group have also been offered rescreening at least once. No estimates of the impact on either colorectal mortality or on 'all cause' mortality will be made until the middle of the 1990s. In addition, mortality results from the trial cannot be discussed in advance of this report; this creates obvious problems for an economic evaluation; the solution is to present results in terms of interim measures like cases detected. Chapter Eleven makes estimates of the likely mortality reduction on the basis of the available published evidence.

Danish trial (Kronborg et al. (1987), (1989))

Many aspects of the protocol of the Danish RCT are similar to those of the MRC trial described above. A pilot study established the acceptability of screening in the Danish county of Funen prior to the recruitment of 61,735 people aged 45-74 from the general population in 1985. Comprehensive population and health registers were used to exclude those with pre-existing disease and also to follow-up the population. The study group were offered a three-day Haemoccult test every second year, with dietary restrictions during initial testing. Participation following the initial offer was 67% and the positive rate was 1%. Rescreening of previous acceptors alone resulted in participation of 93% with less than 1% testing positive.

Swedish trial (Kewenter et al. (1988))

Apart from comparing screening with no screening, this trial also sought to compare screening using the rehydrated development protocol for completed tests with screening without rehydration. 27,700 people aged 60-4 were

recruited in 1982-3 and randomly allocated to a study or control group. Those completing a screening test were asked to observe some dietary restrictions; those testing positive were asked to complete a second test. Completed tests were developed according to the following criterion: tests returned by those aged 60 or 61 when the trial commenced had their returned tests rehydrated, while those in the older age group had their tests developed as normal. Investigation was by BEXR combined with flexible sigmoidoscopy. Rescreening took place at an interval of between 16 and 22 months.

The main finding from this trial is that sensitivity was higher in the group whose tests were rehydrated; however, the figures for the group whose tests were not rehydrated are very low in comparison with other sources (86% for cancer versus 22% in the group whose tests were not rehydrated). Specificity fell from 99% to 96%, however.

There are great problems in calculating the sensitivity of a screening test in such a trial since there is no evidence on how many cases are missed, other than by observing the number of cases presenting with symptoms following a negative screen. These may be fast-growing cases which were not present at screening, however. Alternatively, very slow growing cases which are missed may not present in the interval.

The Memorial Sloan-Kettering Study, New York (Winawer (1991))

Commencing in 1974, 21,756 people aged 40 and above attending an out-patient clinic were recruited to the trial; recruitment ceased in 1979. The sample is thus untypical of the general population. Randomisation was by date of recruitment (i.e. clinic attendance); those attending between March and November were offered screening. This introduces a bias if those attending in summer are different to those attending in the winter.

This is not a trial of Haemoccult screening alone since all those recruited received a full history, physical examination and an offer of rigid sigmoidoscopy; the only difference between the two groups is that the study group were also offered Haemoccult testing. There is no group which was not investigated in some way. Testing was recommended over three days following a meat-free, high fibre diet. The test initially used, the Haemoccult I, was subsequently changed to the Haemoccult II (as used in all of the above trials) when this became available.

Participation with the initial offer of FOB testing was 70-80% with between 2 and 5% positive. It was soon appreciated that participation in rescreening was being affected by the means of recruitment since many people had originally attended the clinic on a 'one-off' basis: as a result participation was down to 35% of the original sample by the fourth round of screening. To cope with this, the trial population was sub-divided into 'one-off' and annual rescreening groups for purposes of analysis. Further problems have arisen in that the method of randomisation has resulted in an imbalance between the

annual screening and control groups, with the former being larger, including more women and having a higher median age. A further flaw in the trial was that initially no special funding was designated to follow-up the population.

Mortality results indicate that the 10-year survival rate is significantly higher in the group not previously screened before the trial. In the screened group there have been 3.6 deaths from colorectal cancer per 10,000 person-years of follow-up as compared to 6.3 in the control group; an observed reduction in colorectal cancer mortality of 43% was not statistically significant. Overall mortality is virtually identical at 77.2 deaths per 10,000 person-years of follow-up in the group offered screening compared to 77.9 in the control group.

Minnesota (Mandel et al. (1988))

Recruitment of 46,622 people aged 50 and over began in 1975 and finished in 1977. The trial population was drawn from a self-selected group: 30% were American Cancer Society volunteers while others were drawn from the membership lists of civic and fraternal organisations. Recruits were randomly allocated on the basis of age, sex and geographical area to one of three groups: a control group (who were not approached), a group offered screening every year, and a group offered screening every second year.

Initially, test development did not use rehydration, but early sensitivity results were lower than expected and, in order to achieve the intended mortality reduction, subsequent completed tests were rehydrated. As a result each screening round has a different ratio of rehydrated to unhydrated tests. Also comparisons of unhydrated and rehydrated testing are complicated because the former was used mainly in the earlier screening rounds when prevalence of disease is higher. While the sensitivity target has been achieved there is a high false positive rate. In addition, a review of mortality trends indicated that no significant difference was likely, hence screening recommenced in 1982 after a four year gap, with the intention of continuing until the end of 1995 (although nobody above the age of 80 will be screened).

A major problem for both American trials is the use of health-conscious groups as samples. In particular these people may seek screening if they are allocated to the 'no screening' arm of a trial; a survey in 1984 found 48% of physicians routinely follow recommended screening protocols (American Cancer Society (1984)). Another survey of randomly selected people aged between 40 and 75 found that 40% had completed a faecal occult blood test and 25% had done so within the last year; the figures for rigid sigmoidoscopy were 35% and 10% respectively (Polednak (1990)). Any advantages of screening will then affect the control group, making it harder to show a survival advantage for an offer of screening. It will also give a misleading impression of the participation rate.

Another problem is the follow-up of a population with fewer sources of centralised health data than in this country. In Minnesota, annual follow-up is by means of a postal questionnaire.

CONCLUSION

Colorectal cancer screening fulfils many of the criteria for a successful screening programme, although understanding of the disease is imperfect; a particular weakness is the lack of a clear risk factor for developing the disease. Screening for faecal occult blood as an indicator of malignancy is the most thoroughly evaluated means of screening. Clinical trials indicate that screening achieves many of the intermediate goals for an effective programme although these are subject to the usual biases. Mortality data from several American sources indicate that there is a gain to screening; conclusive evidence from the trials is awaited.

Chapter Three

METHODOLOGICAL ISSUES IN THE EVALUATION OF SCREENING PROGRAMMES

Economic evaluation compares the costs and benefits of two or more ways of allocating resources. This chapter discusses the application of standard economic evaluation techniques to screening. A literature review reveals a number of previous evaluations and the methodological strengths and weaknesses are discussed. The chapter concludes by describing the approach chosen for the present evaluation.

Types of economic evaluation

The various types of economic evaluation have been extensively discussed elsewhere (see e.g. Drummond et al. (1987)); where they are relevant the differences between the methods are discussed below. By way of introduction, there are commonly thought to be four types of evaluation:

i) cost-minimisation analysis, where the outcomes (or benefits) of each option are identical and hence the comparison is on the basis of costs alone;
ii) cost-effectiveness analysis, which compares options in terms of a ratio of costs to benefits where the latter are measured in terms of a 'natural' unit such as cases treated, lives saved or years of life ('life-years') saved;
iii) cost-utility analysis, which generates a unit of outcome called the Quality-Adjusted Life-Year (QALY) encompassing changes in both quality and quantity of life by expressing health status as a proportion of normal health and weighting life-year gains accordingly - differences between options can thus be compared to differences in cost;
iv) cost-benefit analysis, which seeks to value all effects in a common unit such as money, thus allowing a net benefit to be calculated.

Different types of study are appropriate in different circumstances, depending in part on the study question addressed.

Establishing a framework for the evaluation

The first stage of an economic evaluation is to determine the study question, which is not always as straightforward as it may sound. For example, "What is the value of screening for colorectal cancer?" and "Is it worth screening for colorectal cancer?" beg the (respective) questions "of value to whom?" (patients? the health service? society as a whole?) and "compared to what?" (other interventions in colorectal cancer? other screening initiatives? other health care programmes?). Thus, the question should determine the viewpoint to be adopted (i.e. determining the range of costs and consequences to be included) and alternatives to be evaluated.

Screening options

The parameters which go to make up a screening programme include who does what to whom, when, where, and how often. Each of these can be varied to produce a multitude of options based on the following:

i) which screening test(s), which diagnostic test(s), and which combination(s) of the two;
ii) how to define and identify the target population;
iii) what retesting protocol (if any) is appropriate prior to full diagnostic investigation, and should it use the same screening test;
iv) if they are endogenous, what levels to choose for the sensitivity and specificity of the screening and diagnostic tests;
v) the number and timing of rescreens (variable intervals are possible), plus the possibility of more frequent screening of high-risk groups;
vi) the appropriate administrative structure (a dedicated screening unit?), the method of invitation, use of means of increasing participation.

To illustrate the number of potential options, one evaluation of cervical screening using a mathematical model compared 100 different protocols (Koopmanschap et al. (1990)). Clearly, no clinical trial is capable of evaluating more than one or two of these options.

Identifying the costs and benefits of screening

The next step is to identify all the relevant costs and consequences of screening (direct, indirect, tangible, intangible) from the viewpoint specified in the study question. In the absence of any clear guidelines in this area, an evaluation can only hope to be explicit about what has been included and what has been left out. The figures can thus be reworked if some users of the information wish to adopt a different perspective. A comprehensive list of the costs and benefits of screening would include:

- NHS costs
  - screening (including administration, training, equipment and buildings)
  - diagnostic investigation of positive screening test results
  - treatment of those found to have disease
  - follow-up of those treated
  - treatment of recurrent and terminal disease
  - primary health care
  - costs of care as a result of longer survival

- other public sector costs
  - use of social services
  - convalescence

- costs borne by patient, family and friends
  - 'out-of-pocket' expenses
  - emotional costs (anxiety and depression)

- health effects
  - survival effects
  - complications and iatrogenic disease
  - quality of life effects

- psychological benefits
  - reassurance of a negative test (and of early treatment?)

- societal effects
  - effect on national income
  - valuation of non-working time

The next step is to measure and value these effects.

The costs of screening

Costing is an aspect of evaluation methodology which is frequently taken for granted; textbook comment (if any) is usually confined to the allocation of capital and overhead costs. For example, "It is often suggested that costs are easier to determine than benefits... [t]he process of quantifying - measuring - the inputs of a given programmatic alternative is generally straightforward..." (Warner and Luce (1982)). A traditional problem has been the lack of reliable data on which to base cost estimates; as one researcher found, "... [t]he attempt to compile costs for a small but identifiable group of patients in a large hospital proved less than straightforward." (Brooks (1981)).

Measuring resource use

Some procedures, particularly diagnostic tests, are relatively homogeneous in terms of resource use and a 'typical' procedure can be described. Other procedures, such as major surgery, are less predictable and there is potential for variation between patients. This latter group thus requires a more detailed approach to costing.

Homogeneous procedures can be modelled using data from surveys of the staff involved and by direct observation of a small sample of procedures to produce a representative cost for a 'typical' procedure. To cope with uncertainty a probability distribution can be used to calculate the expected cost of the procedure (where that term is used in the statistical sense).

If in-patient care is to be costed on an individual basis careful attention must be paid to the amount of research time involved. Data can be collected prospectively (by direct observation or using survey methods) but it is often difficult in this situation: it is not feasible for the researcher to be present all the time, while ward staff are reluctant to complete more paperwork. Inevitably, the main data source is from a retrospective review of the patient's case notes, but these are constructed as an aide memoir for medical and nursing staff. Not all relevant resource use is recorded, such as the time doctors and nurses spend with each patient.

Given that case notes are the main source, what exactly should be counted? This is a question of the level of detail required in the evaluation. As a rough guide, the amount of effort put into costing particular categories of resource use should be in proportion to the impact of that item on total cost. Thus, for most surgical procedures it is more important to pay attention to nursing and medical staff costs and general overheads rather than to diagnostic tests and drugs.

While the principles involved in costing care in this way are easy to describe the practice is infinitely more complex. A number of other problems arise as the case notes are analysed. A major problem proves to be obtaining a single set of figures which are easily generalisable across settings; factors which might cause bias include:

- medical practice variations between hospitals and between individual doctors e.g. due to different levels of expertise in technologies where a 'learning curve' effect exists;
- hospital cost structures, with teaching hospitals (which host many clinical trials) being especially expensive. A variant on this problem arises when patients in the sample are treated in more than one hospital, each with different costs. One option is to use a single set of unit costs to value individual patient resource use.

Other problems relate to exactly which costs incurred to include:

- activities carried out solely for research e.g. early admission pre-operatively to take part in other studies. These must be excluded wherever possible, but this is not always easy.
- treatment for unrelated conditions should be excluded if they would have occurred in any event. On the other hand, iatrogenic disease resulting, for example, from cross-infection while in hospital should be included. The best that can be hoped for is that the evaluation will clearly specify the inclusion and exclusion criteria used.
- while many patients are treated in NHS hospitals, a minority are either treated privately or are treated while overseas. This makes their notes difficult to obtain but also excludes their costs from the analysis if the perspective is that of the health care purchaser.
- unless all of the sample have already died, there must be a defined cut-off point in the costing to ensure consistency in the costs. This is most likely to be until death or until a given period of follow-up, whichever is shorter.
- a life-saving cure causes people to live longer and thus to increase the health service costs of treating other diseases of old age. One consideration is that including these costs makes comparison with the results of other evaluations (which commonly exclude these costs) difficult.
- screening changes the timing of health service resource use since, if successful, costs are incurred in the present to produce savings in the future. Economic theory holds that society prefers to defer costs

wherever possible, however, and hence a discount rate should be used to take account of the time profile.

By this stage it is easy to agree with the conclusions of one group of reviewers: "Costs are not immutable 'facts' lying ripe in the field waiting merely to be garnered, or even selectively winnowed, by diligent clerical officers." (Blades et al. (1987)). The costs derived are full of value judgements, although they are at least made explicit; none of the assumptions on the inclusion of particular items appears to seriously bias the results presented in the following chapters.

The benefits of screening

Identification

As noted above, purchasers are being driven by national targets on reductions in mortality rates. The significance of quality of life as another dimension of benefit is recognised, but inevitably the demands of the policy target dominate decisions. Other benefits are also neglected, such as the reassurance value of a negative test or the value people place on better information about their health. Purchasing these benefits may have an opportunity cost in terms of achieving mortality reduction targets.

Another issue relates to the inclusion of the indirect benefits of treatment, defined as the gain in economic productivity as a result of prolonged survival and good health. This whole area remains extremely controversial with little agreement among economists on the correct way to proceed (Ratcliffe (1993)). It could be argued that, other things being equal, a treatment which keeps a patient out of the workforce for a shorter period of time would be preferred to one requiring prolonged convalescence. This leads to ethical problems with giving priority to individuals with high earnings since their time is more 'valuable' on the basis of their wage; people who are not in paid employment receive a low priority on this basis. It is not clear that society is explicitly willing to prioritise treatments on these grounds. A second problem is that the principle aim of the health service is to increase health alone, with no judgement about what the individual should do with the good health once they have it; this would be implicit in any attempt to value extra life-years on this basis.

Measurement and valuation

Measures of the health benefits of screening found in the literature include:

i) intermediate e.g. cancers detected with the implicit assumption that early detection will be beneficial in health terms.
ii) pseudo final e.g. survival rate at five years following diagnosis;
iii) final e.g. measures of survival and quality of life gains.

Intermediate outcomes are acceptable when comparing ways of achieving a similar aim such as screening for colorectal cancer. 'Pseudo' final outcomes are liable to some of the biases described in Chapter Two, as well as having other deficiencies (McNeil et al. (1978)):

- it is not always evident why five years has been chosen as the cut-off point;
- it is implicit that survival of less than five years is of no value;
- it is implicit that the value of surviving five years is the same to everyone; and
- it takes no account of quality of life.

More sophisticated trials take account of many of these problems by providing information on the number of life-years saved as a a result of screening. However, this still takes no account of quality of life in the extra years.

A few clinical trials include measures of health status although the use of disease-specific measures creates problems for an economic evaluation, the aim of which is to compare the costs and benefits of screening with other uses of health service resources. Ideally, therefore, the health status measure would be as broad (and hence as generally applicable) as possible: this was the motive for the development of the QALY or Quality-Adjusted Life-Year.

To an economist, the correct measure of the benefits of a programme are the alternative benefits which must be given up to achieve them: if society is not prepared to give something up to achieve a goal then that goal cannot be worth very much. One way to assess this would be to invite society to choose between alternative 'packages' of benefits for the same costs. There are many ways to use the resources, however, hence it is argued that a common measuring rod of value, money, could be used instead to identify the best alternatives. While it is easy to see how this could work for investment decisions in industry and in some social projects, such as motorway construction, its application to health care has been problematic. Problems arise when this is in terms of gains in quantity or quality of life. Values are implicit in all everyday decisions but eliciting them is harder, although it has been attempted in screening for gonorrhoea (Goddeeris and Broncken (1985)). This has resulted in continuing reliance on the unsatisfactory QALY measure described above. There are many problems inherent in this measure and still more in its application; at present, however, few realistic alternatives are available.

Marginal analysis

A decision is clear cut when one option is cheaper than the other and has a higher yield. Where the more expensive option also detects more disease, however, average costs can conceal the true nature of the trade-off. In this case the appropriate figure to report is the extra cost required to detect the extra cases. The most famous example of a marginal cost comes from an earlier paper on colorectal cancer screening (Neuhauser and Lewicki (1975)). This is a neglected aspect of evaluations of screening, in spite of this example - as one review comments: "Despite the plethora of potentially

relevant margins, many studies take a very narrow perspective." (Cairns and Shackley (1993)).

Sensitivity analysis

Mass cancer screening is very complex in terms of the protocol options and to this must be added the intricacies of the disease process. As pointed out in Chapter Two, few details of the natural history of colorectal cancer are known for certain. One way to cope with this uncertainty is to assess the robustness of the results to changes in the underlying variables. Where the results depend crucially on particular variables more analysis of the assumption is required. In Chapter Four, for example, a simple model of screening identifies the most important underlying variables which are then the subject of more detailed work in subsequent chapters.

Decision criteria

The appropriate decision criteria to use depends on the type of comparison to be made. If the comparison is between two ways of screening for the same disease then the costs and disease yield of each option can be used, on the basis of marginal analysis. However, it is inappropriate to compare the cost per case detected by colorectal cancer screening with that of breast cancer screening and cervical cancer screening, for example. The case detected is only an intermediate measure of benefit; each type of cancer detected is likely to represent a different profile of health gain to the patient. Comparisons with other health services are more difficult unless a common measure of health benefit such as the QALY or £-value (through willingness-to-pay estimates) are used.

Issues of interest to potential users of the results

There are many practical issues which users of the results must address before implementing the findings of a programme. The evaluation presented in the following chapters concentrates on the initial round of screening and the two subsequent rounds, limited by the amount of data available from the MRC trial. The results of these later rounds are taken to be indicative of the costs and yields of the programme in its 'steady state'; the evaluation does not consider the costs of establishing the programme in terms of staff recruitment and training, purchase of computer equipment, etc.

A second set of practical issues surrounds the time profile of the costs incurred and the cost savings arising from screening. While these will have been taken into account in the cost data via discounting future costs to a net present value, this reduces a time profile to a point estimate. It may be more meaningful for purchasers to see what this will mean in terms of service use over a time horizon sufficient for the programme to reach its 'steady state'. This will reveal the likely 'bottlenecks' arising in the health service in terms of diagnostic and therapeutic capacity, etc. For monitoring purposes it will also be useful to have a prediction of when the health gains may be expected to materialise. These might take the form of the number of cases detected and

the change in five-year survival rate as an intermediate measure, with the reduction in colorectal cancer mortality for the longer term effects.

PREVIOUS ECONOMIC EVALUATIONS OF COLORECTAL CANCER SCREENING

This section reviews the literature on the evaluation of programmes of screening for colorectal cancer. The studies range from crude evaluations contained within a single paragraph of a review article up to complex mathematical models which attempt to predict the impact of the disease process not just on resource use but also on the disease process. This diversity makes the literature difficult to summarise adequately, but Table One outlines the type of evaluation used, the country of origin of the research, the screening test evaluated and the main data source:

Table One

1st author	EE type	Country	Test(s)	Data source
Agrez (1)	C	Australia	HM, HQ	Literature
Allison (2)	E	USA	HM	One-off study
Applegate (3)	E	USA	HM	Pilot of RCT
Atkin (4)	E	UK	FS	Literature
Bolt (5)	E	USA	RS	Literature
CRESGE (6)	E	France	HM, CS	Literature, expert opinion
Eddy	E	USA	HM, FS	Literature, expert opinion
England (7)	E	USA	HM, CS, FS	Literature
Farrands (8)	B	UK	HM	Pilot of RCT
GOICC (9)	B	Italy	HM	Literature, local data
Johnson (10)	E	USA	HM	One-off study
Joseph (11)	E	USA	HM, HQ	Literature
Kristein (12)	B	USA	HM	Literature, expert opinion
OTA	E	USA	HM	Literature, expert opinion
Neugut (13)	E	USA	CS	Literature

Notes to table

B - Cost-benefit analysis
E - Cost-effectiveness analysis
C - Cost analysis
HM - Haemoccult
FS, RS - Flexible and rigid sigmoidoscopy
CS - Colonoscopy
BXR - Barium enema X-ray
HQ - Hemoquant (a quantitative test for haemoglobin in faeces)

The numbers following the main author refer to references in the text below; the work of Eddy and of the Office of Technology Assessment (OTA) (Wagner et al. (1991)) is in the next section.

General comments

The type of evaluation performed mirrors that found elsewhere in the literature of economic assessment of health care programmes (Adams et al. (1992)).

None of the studies is based on conclusive evidence of health benefits from screening. Only one of the evaluations judged Haemoccult screening to be a failure but this had a particularly low take-up rate (10). Another study questioned the value of FOB screening (1), although the figures were challenged in an accompanying editorial (St John (1990)). Most studies are favourable to Haemoccult testing, including one claim that a screening programme would pay for itself in terms of treatment cost savings (12). Colonoscopy has a high yield as a screening tool but is relatively expensive (6,7,13).

What are the strengths and weaknesses of these evaluations? This issue is considered under six headings:

i) did the study identify, measure and value resource use in a suitable way?
ii) did the study make an appropriate estimate of the health benefits, taking account of all the biases inherent in the results of some clinical trials?
iii) did the study consider other aspects such as rescreening options, participation rate, choice of diagnostic investigation regime, age limits on target population?
iv) did the study perform a sensitivity analysis?
v) did the study compare options using a marginal analysis?
vi) what decision criteria were used and is it clear why the preferred option was chosen?

The studies can be divided into those that could be termed 'conventional' economic evaluations of one or two options for screening and those that are essentially mathematical models which are capable of evaluating limitless numbers of protocols; the latter are considered in a later section.

Cost coverage

Very few studies explicitly stated the viewpoint of their appraisal. All but one of the studies include the costs of screening and of diagnosis. Most consider treatment costs but only four studies also consider costs of symptomatic presentation (2,8,9,13). None of the studies include the costs of administering the programme, nor do they include non-health service costs. In terms of valuing resource use all of the American studies used hospital charges.

Benefit estimation

All of the evaluations consider cancer yield, at least as an interim step in the calculation of benefits. Four studies then apply five-year survival rate figures to the staging distribution (2,7,11,12). All of the cost-benefit analyses valued

the number of life-years saved using average earnings, with two allowing for participation in the workforce (8,9).

The benefits of adenoma excision are considered in only two studies (2,6). Both assumed that 5% of adenomas excised would eventually have become cancers and that they would have presented according to the symptomatic staging distribution observed.

Lead-time bias was mentioned in three studies; figures of 2, 5 and 15 years were used (3,10,12). Length bias was only taken into account by one study and this only let it affect workload, not benefits (9).

None of the studies estimates any benefits to screening beyond length of life (except where quality may impinge upon ability to undertake paid employment). One study mentions that benefits include extra productive time at work and 'other intangibles' (13).

Other aspects of screening

All of the studies assume that cancers presenting symptomatically are a homogeneous group, whether they occur in the 'no screening' option, among people refusing screening or in the interval following a negative test. Rescreening is considered in three studies, but in one this is only as a cost comparison with the initial screen (3). None of the studies considers the age limits most appropriate for screening, and very few are explicit about the economically optimal age range they would suggest.

Sensitivity analysis

Seven studies did not perform a sensitivity analysis (2,3,5,6,8,10,13), while the remainder vary factors such as test costs, prevalence, lead-time bias, test performance. Many key assumptions go unchallenged such as the remarkably high prevalence rate of 8 per 1,000 assumed in the Italian model.

Marginal analysis

Only four studies consider the marginal effects of different policies (1,4,6,7), with the French study alone adopting a true incremental approach (6).

Decision and conclusion

The cost-benefit analyses make a judgement on the basis of the net benefits of options (8,9,12). Three of the six of the cost-effectiveness analyses estimate the cost per life-year saved (2,7,11) while a further three report the cost per cancer detected or prevented (3,5,10). Another reports the cost per life saved (13). The French study presents the marginal analysis as its results (6).

Models of disease and screening

The mathematical models developed are as much about the disease process and how this is affected by intervention as they are about screening. In terms

of structure Eddy's model is the most sophisticated and has proved the inspiration for several simpler attempts. Originally the model was intended as an aid to the design of randomised controlled trials by identifying the option most likely to be economically efficient, given current knowledge. It has subsequently acquired a 'life of its own' with the results underpinning the recommendations for screening of various American medical societies and insurance companies.

It consists of a nine-state, time-varying Markov chain. The states considered include: alive and healthy; asymptomatic adenoma; asymptomatic cancer; cancer diagnosed (by Dukes' stage); died of colorectal cancer; died of other cause. This assumes that health states are discrete and there are specific probabilities of the likelihood of a given individual progressing from one state to another between two time periods. These probabilities are determined by a series of disease- and test-specific variables. For example:

i) The probability that an asymptomatic adenoma or a cancer will be present at a given point in time is estimated by age/sex-specific prevalence rates, other risk factors present, the length of the detectable pre-clinical phase of the disease and the individual's screening history.
ii) The probability that a screening test will detect an asymptomatic tumour is determined by the stage of the disease, its site within the bowel, the extent of the bowel covered by the test, and the test's random false-negative rate.
iii) The progression through adenoma to cancer and thence through the various Dukes' stages are determined by estimates of growth rates.
iv) The probability of dying of colorectal cancer in the coming time period is determined by stage-specific mortality rates for colorectal cancer.
v) The probability that the individual will die of other causes in the coming year is determined by general age/sex-specific mortality rates.

From a given population, the numbers in each health state at any point in time can be calculated; typically, the results are presented as an analysis of the optimal lifetime screening strategy for someone who is 50 years of age.

Several versions of Eddy's model have appeared and, while the structure is largely unaltered, the values of some of the key variables have changed to reflect developments in medical knowledge and opinion. For example, in 1980 it was assumed that 75% of cancers arise in adenomas, but this increased to 90% in the 1984 version and to 93% by 1987.

Other models have used a simplified form of this approach: for example the American Office of Technology Assessment (OTA) have evaluated the inclusion of colorectal cancer screening in the Medicare (publicly funded) programme for the elderly on a similar basis (Wagner et al. (1991)).

How do the models compare on the same criteria as those used above?

Cost coverage

- Eddy's model includes the costs of screening tests, diagnostic investigations (including adenoma excision), treatment (with separate figures for early stage and for advanced disease), and terminal care. The OTA model uses a similar approach including a full consideration of the five-year follow-up costs following diagnosis plus the implications of adenoma follow-up on the basis of one complete investigation post-convalescence followed by colonoscopy every three to five years until 85 years of age. Eddy's model does not include adenoma follow-up.
- Costs are based on hospital charge data in both cases.
- A discount rate of 5% is common to both models.
- The comparison in the OTA case is with "no screening" yet opportunistic screening is already so widely practiced in America that this may be inappropriate.

The annual net cost of the OTA programme is between $1.5 and $2.6 billion assuming full participation at all stages of the programme. The costs of adenoma follow-up are very important; screening itself constitutes only 4% of the total cost of annual FOB screening.

Benefit estimation

Both models calculate the years of life saved by screening net of losses due to diagnostic and operative mortality. Neither model considers quality of life during those years, although the OTA "assumed that such considerations would enter into individual clinical decisions about the value of colorectal cancer screening in a particular person." The source of Eddy's data is RCT evidence wherever this is available, supplemented by expert opinion; for example the latest version of his model is based on a questionnaire survey of 72 'experts' in the field of colorectal cancer.

To construct the disease model detailed estimates of the progress through the adenoma-cancer sequence and then through the various Dukes' stages are required. Stage-specific survival rates are then applied to the staging distribution.

In Eddy's model benefits are discounted at 5% per annum although the undiscounted results are also presented. It is unclear whether the OTA results are discounted or not.

Eddy has cross-checked the yield and benefits predicted by his model against the results of the on-going randomised trial in Minnesota and found them to be accurate to within 1% of the observed total.

Other aspects of screening

Eddy's model:

- does not include a variable for participation and hence has no non-responder group.

- assumes all symptomatic presenters, including those in the interval following a negative test, to be homogeneous with respect to cost and outcome.
- considers rescreening with varying intervals.
- evaluates a wide variety of screening protocols.

In the OTA model:

- cases occurring in the interval following a negative screen and among those declining testing are ignored, implicitly assuming cost and survival are unaffected by screening.
- the only protocol evaluated is screening by annual, unhydrated Haemoccult commencing at age 65 and finishing at 85.
- the implications of screening such an elderly group in terms of participation and operative mortality are not discussed.

Sensitivity analysis

In his 1984 paper, Eddy used expert opinion as his main data source: his results varied by 25-50% according to individual opinions. The most important assumptions are those relating to:

i) the proportion of cancers arising in adenomas;
ii) the time from an adenoma first becoming detectable by radiology or endoscopy until it becomes malignant;
iii) the time a pre-cancerous lesion is detectable on FOB testing.

The total cost is very dependent on the specificity of the screening test since diagnostic test charges are very high in America.

The key assumptions in the OTA model are those relating to the proportion of symptomatic presenters with early stage cancer; the prevalence, incidence and recurrence rate of adenomas; and the proportion of rectal adenomas which regress.

Marginal analysis

Eddy compares all options in terms of a trade-off of the net present value of total costs against life-years saved: this is presented explicitly as a value judgement which must be made. The OTA model does not include a marginal analysis.

Decision and conclusion

Both models predict that screening will not result in a net cost saving, implying a decision must be made on the trade-off of costs and benefits. The results arrived at can be compared in terms of the one common option evaluated, annual FOB screening. As noted above, the OTA assumptions are slanted against screening and this is reflected in the results which imply that the cost per life-year saved is $35,000 as compared to Eddy's figure of $2,420 (both compared to a situation of no screening).

Eddy chooses annual FOB screening with 3 to 5 yearly flexible sigmoidoscopy as the most technically efficient way of screening for

colorectal cancer; allocative efficiency is given less attention although some rudimentary comparisons with other programmes are made. While the information on each option is clearly presented, it is unclear how the final choice was made in the absence of further information on societal willingness-to-pay to save one more year of life. Additionally the recommended option in the most recent version of Eddy's paper has higher costs and lower benefits than several other options. He has sought to justify this on grounds of discomfort of diagnostic investigations, lower participation in some regimes, etc., being excluded from the figures. This is certainly the case but this information was not presented (or even mentioned) in the original conclusions. Subsequently, he has back-tracked still further by saying, "Colorectal cancer screening is an excellent example of a medical practice for which no single policy will be best for all patients, and the maxim 'talk to your doctor' applies." (1991). This seems at odds with the purpose of the analysis which was to identify the economic 'best practice' in this area.

A particular feature of the OTA analysis is the role of colonoscopic surveillance following excision of adenomas. Under the assumptions used any screening programme involving flexible sigmoidoscopy will eventually result in 55% of the population being involved in an adenoma follow-up programme.

Neither model considers how screening might be stopped if the results of the on-going clinical trials eventually show that their benefit calculations are too optimistic and screening has no benefit.

Summary

Mathematical models are of more value than 'one-off' evaluations since they can be applied to many different protocols. If the data used are good enough then clinical trials of each protocol become unnecessary, saving on the cost and time involved in evaluating a new test. They are commoner in America because of the problems of restricting the spread of new health care interventions in advance of proof of their effectiveness in that country. Nevertheless such models need RCTs in the first place to assist with their construction: for example, the predictions should be carefully validated against RCT evidence. Since they require data which are often not available, one possibility would be to design clinical trials to fill in gaps in the model; this will require a shift in current thinking, however. Until then, the assumptions made must be presented alongside the model results so that wary readers can judge the value of the analysis for themselves.

METHODOLOGY OF THE PRESENT WORK

The sections above have posed many methodological dilemmas which must be resolved. As noted at several points, there are often no right or wrong answers and the best that can be hoped for is a clear statement of the judgements made. These are listed below.

Type of evaluation

As set out in the Introduction, this evaluation addresses four related questions, assuming the viewpoint of the health service purchasing organisation. The first question asks simply for the costs and benefits of one of the options and hence is not an evaluation per se. The second question can be answered using cost-effectiveness analysis, comparing different ways of screening in terms of costs and cases detected. The third question can also be addressed by cost-effectiveness analysis if the outcome measure chosen is lives saved or some similar 'natural' unit. If both length and quality of life are thought to change then measuring benefits on a single dimension is inappropriate and cost-utility or cost-benefit analysis must be used. These comments are also appropriate to the fourth study question, since comparisons are now being made with a broad range of alternative uses of the resources; these may affect the health of those concerned in quite dissimilar ways.

Viewpoint of the evaluation

Identifying the appropriate viewpoint is not straightforward, however. At the start of the evaluation (October 1987) it was common practice to adopt a societal perspective (although the practice often fell some way short of this). With the onset of the 1990 NHS reforms, however, the role of evaluation altered to one of informing the decisions of those placing contracts i.e. health authorities and GP fundholders. This implies a greater emphasis on health service costs than was the case previously, where £1 of cost incurred counted equally whether it be incurred in the operating theatre or in terms of visiting by relatives. At the same time national policy documents such as 'Health of the Nation' (Department of Health (1991)) have established targets for purchasing in terms of health gains (or, to be precise, mortality reductions). This implies that the evaluation should focus upon health service costs and mortality reductions as the only benefit.

Costs

With the advent of the internal market, most acute hospitals can quote prices for most of the procedures they carry out. The divergence between prices and costs is well understood in North America. Some prices may be increased to reflect uncertainty over workload or to cross-subsidise between services. Guidelines state that prices should be equal to average costs, producing no net 'profit' (NHS Management Executive (1993)). However, this assumes that average costs are known in advance. Others have questioned the motives of hospitals: "Essentially, the first rule of charges is to make the institution solvent...Researchers attempting to assess economic efficiency must clearly defend their reasons for believing these gaps are small if they use charges as a proxy for cost; otherwise they should collect data on resource consumption and directly attempt to measure economic costs." (Finkler (1982)). The evidence implies that these gaps are very large in the NHS at present. For example, the prices of day case colonoscopy in five

acute hospitals in Glasgow lie between £550 and £650, but the cost of the resources used is little more than £100 (from Chapter Four). Given that American charges for colonoscopy are about $500 and a return ticket to America costs about £200 the Glasgow purchaser could afford to fly all patients requiring day-case procedures to New York for the same price!

There is a dilemma. On the one hand producing information for purchasers implies using the prices they face in the analysis. On the other hand the prices bear little relation to the resource costs and, once more detailed information becomes available, it is hard to believe that they will remain at this level. A possible resolution lies in the fact that the price tariffs relate to individual referrals by GP fundholders, whereas the majority of cases are dealt with under block contracts whereby the hospital agrees to deal with all the referrals made by non-fundholders for a given total reimbursement. Thus neither party knows exactly what is being charged. If these contracts are ever disaggregated it seems likely that costs will be set on the basis of work studies such as those carried out in the course of economic evaluations rather than by the allocation of blocks of costs as favoured by accountants. For this reason the present evaluation uses costs rather than charges.

It would not be appropriate to simply use the costs incurred in the course of the MRC trial, however. The main problem is to estimate the costs of the screening protocol as they would be if it were implemented nationally. The cost of the clinical trial provides some information but this is biased by the inclusion of research items and the exclusion of the costs of some facilities provided for free by the teaching hospital. Other items may be provided at a reduced cost such as items of screening equipment.

When costing hospital procedures the following rules for inclusion and exclusion of particular items are used:

- treatment of other conditions is included unless there are clear reasons for not doing so, assuming that most care related to the disease or its complications. Treatment of conditions noted in the patient's medical history at initial admission but unrelated to the onset of cancer were excluded
- treatments for colorectal cancer which are clearly experimental, such as resection of liver metastases, have been excluded since they would only be attempted in a few 'centres of excellence';
- patients treated privately or overseas are included wherever adequate information was available. In this context, 'adequate' was defined as including number of admissions, duration of stay and a description of treatment.
- various time cut-off points are used depending on the availability of data. Given the stage of the MRC trial, there is little information on what happens after the third round of screening, and even in that case the numbers are small; screening beyond this is not considered. Treatment costs incurred up to three years after a diagnosis of cancer were

recorded; while arbitrary, this is the resolution to the trade-off of sample size (to be maximised) against allowing follow-up costs such as those related to recurrent disease to be incurred.

- at present, the medium- to long-term effects of colorectal cancer screening are uncertain; median follow-up of the sample is only a couple of years. Over such a time horizon discounting at standard rates makes little difference to the results, although the potential implications are considered where appropriate.

Benefits

No data are available at present on the effects of screening on mortality from the disease; indeed, the MRC discourages discussion of this issue, in case the on-going trial is jeopardised in any way. This has an important bearing on the way in which each of the study questions can be addressed. The first question can only be answered in terms of interim measures of benefit such as cases of disease detected, although health gains can be included quite simply, once they are available. The second question can still be addressed in terms of cost per case detected, although cases must be defined either in terms of asymptomatic cancers or adenomas, but not both; this creates problems when different tests have different sensitivities to types of disease. The third issue can be addressed on the basis of literature values if it is accepted that lives saved is the main benefit dimension. The final issue is more difficult: QALYs are difficult to construct without survival figures, while cost-benefit analysis requires knowledge of the full benefits of screening for valuation purposes.

Indirect benefits have been excluded. Given the age of the target population it is unlikely that they would play a crucial part in the analysis. Additionally, review articles indicate that they are included in only about 10% of evaluations, so excluding them makes the results of this study more easily comparable with the majority of other analyses.

CONCLUSION

None of the studies found in the literature is entirely satisfactory, either because important aspects are not considered or because the analysis (particularly via mathematical models) relies so heavily on assumptions for the values of key variables. In the absence of a consensus among economists on evaluation methodology, the assumptions used in constructing the results from the following chapters have been set out. While the limitations imposed on the current study do not allow all of the study questions to be fully addressed, enough data is available to carry out much of the work; once survival data is available it will be relatively simple to complete the evaluation.

Chapter Four

BASIC MODEL OF SCREENING FOR COLORECTAL CANCER

Introduction

This chapter describes a simple model of screening for colorectal cancer in a defined population according to the MRC trial protocol for the study group. The costs of screening and of diagnostic investigation to the health service are calculated for the first three rounds of screening and compared to the yield of asymptomatic cases detected. An analysis of the robustness of the results with respect to changes in the variables then provides material for the subsequent chapters.

Modelling the screening process

Converting the experience of a research trial to a more general setting requires a number of assumptions. These have two purposes: firstly, to remove costs related solely to research (e.g. recruitment of GPs and randomisation of patients), and secondly to allow for differences between the MRC trial protocol and the likely form of a national programme if one were to be implemented. The MRC protocol of offering a three-day Haemoccult test to every person aged 50-74 every two years is evaluated in this chapter. To calculate the costs and cancer yield of screening the data are applied to a target population of 100,000 eligible people.

Administration

The MRC trial is administered by two whole-time-equivalent clerical officers. These officers have a variety of additional duties which would not be relevant outside the context of the trial; initially it is assumed that one clerical officer and two clerical assistants with the appropriate computer software would be sufficient. In addition, a supervisor of senior administrative grade would be appointed with responsibility for co-ordinating local efforts.

The programme would use Family Health Services Authority (FHSA) records to identify the target population. This data source is known to contain inaccuracies of up to 20% (Stilman (1984)), implying that a proportion of distributed tests would not reach their intended destinations, let alone be completed or returned. However, this can be expected to improve over time as more use is made of FHSA lists for other purposes and better linkage of information systems is achieved; thus, it is assumed initially that FHSA records are accurate.

It is assumed that the screening administration will be run from an office at the FHSA; an opportunity cost for rental value of 1,800 square feet of office space is included, valued at £10 per square foot (Ridley et al. (1991)). Access to the existing mainframe computer is assumed to be available at

minimal extra cost. There will also be set-up costs covering, for example, the purchase of furniture, computer equipment, and lines to the FHSA's mainframe computer. Administrative staff would need to be in post for some months prior to the initial invitations being sent out. Fixed costs such as power, heat and light are ignored on the assumption that they would be incurred irrespective of the existence of the screening programme.

Screening

The current cost of a 3-day Haemoccult test is £1.13, including reagent (based on the purchase price of one pack of 50 tests). Each test costs £0.41 to send (including postage, stationery and instruction leaflet).

Participation in screening is voluntary and the rate of 57.8% is adopted from the randomised evaluation of 3- and 6-day testing in the MRC trial (Thomas et al. (1990)). Publicity surrounding the introduction of a national programme may increase this but no effect is assumed. Completed tests are returned to the pathology laboratory of the local hospital; observation suggests that a sustained test development rate of 60 tests per hour is feasible and the time of a laboratory technician is costed accordingly.

Investigation

A positive rate of 1.29% of completed tests, as in the MRC trial, implies 746 positives tests out of 57,800 returned in the population of 100,000. Following the MRC trial protocol these people are invited to a screening clinic (run by a consultant surgeon with a nurse present), which assesses 6 patients per session; this unambitious throughput allows time for discussion of the results with the patient. Routine blood test are conducted and rigid sigmoidoscopy performed. Further investigation is by colonoscopy for preference: 88% of the patients are investigated by colonoscopy and 12% by flexible sigmoidoscopy plus double-contrast barium enema X-ray (BEXR). As no reliable costs are available for these procedures they have been valued separately: the details are provided in the appendix to this chapter. The cost of a colonoscopy is £105.10 while the cost of BEXR plus flexible sigmoidoscopy is £126.68.

Results

Cost and yield in the initial round of screening

The total cost of the initial round of screening for a population of 100,000 people appears in Table One (all cost figures are expressed in 1989/90 prices):

Table One

		Sub-totals	% of total
Programme overheads		**£103,818**	**28**
Administration	£37,818		
Supervisor	£30,000		
Office space	£36,000		
Screening		**£158,884**	**43**
Tests	£113,000		
Postage	£41,000		
Development	£4,884		
Investigations		**£109,775**	**29**
Out-patient clinic	£29,479		
Colonoscopies	£68,946		
BEXR / flexible sigmoidoscopy	£11,335		
Total		**£372,477**	**100**

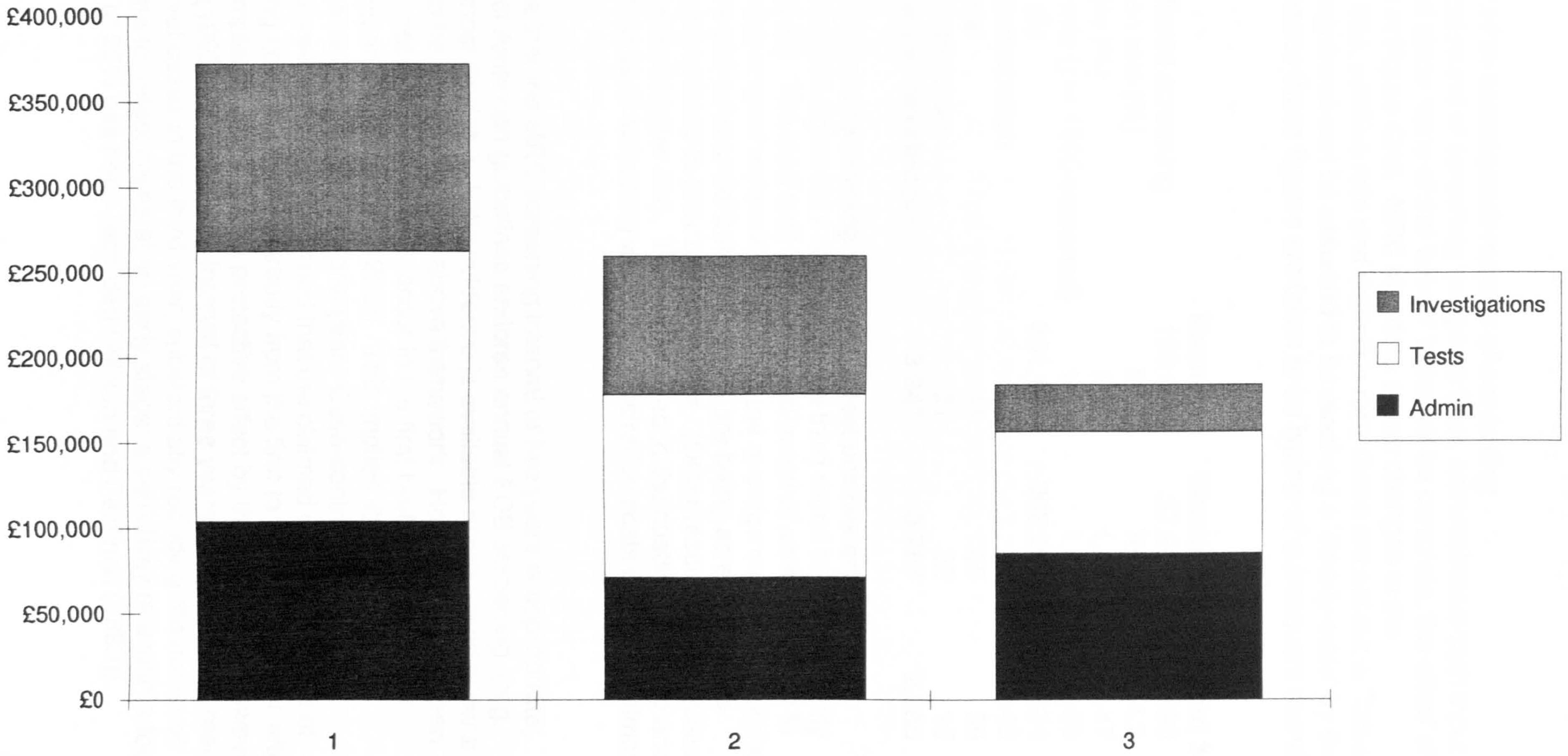
These results are presented graphically in the top half of Figure One.

In the MRC trial, three-day testing detects 1.97 asymptomatic cancers per 1,000 persons completing a screening test, i.e. 114 cases in the population modelled at an average cost of £3,271.

These are not the only effects, however, as noted in Chapters Two and Three above. In total, 7,456 people return a positive test. If they all agree to further testing, then seven will suffer a major complication as a result of investigation (on the basis of evidence presented in the Appendix to Chapter Four). Many of those with a positive test will have an adenoma detected will be entered into an endoscopic surveillance programme; depending on the protocol used, further complications (and deaths) will occur. There is also a danger of anxiety provoked by the positive test which is not fully dispelled by the investigation which shows a benign cause or finds no abnormality. It is possible that the 50,344 people who test negatively derive some reassurance value from the finding.

In addition, 61 people who had a negative Haemoccult test will present with symptomatic cancer within the next two years (interval cases), assuming three-day testing to be 65% sensitive (Thomas et al. (1992)). There will also be 78 cases in the group who refuse screening in the two year interval between screens, based on the incidence in that group of the MRC trial. This gives a total of 253 cases over the two years after the implementation of screening. Without screening there would have been 144 cases in the age cohort, based on the incidence rate reported in the MRC control group (Hardcastle et al. (1989)). Theory would predict such an excess as the prevalence of cases of disease is cleared, with future incidence reduced as a result. However, the number of symptomatic cases following screening is almost as great as the number without screening (139 versus 144).

Figure One - Programme costs for three screening rounds



Cost and yield in subsequent rounds of screening

For the second round of screening two years later, administrative staff input can be halved since many of the 'set-up' tasks will be complete; the effect of this is shown in Figure One. MRC trial data shows changes in the participation rate, positive rate and detection rate: these are set out in Table Two. The programme can be assumed to be reaching a 'steady-state' by the third round, hence these figures are taken to be typical of subsequent rounds.

Table Two

	Round 1	Round 2	Round 3
Number offered screening	100,000	57,800	44,390
Participation rate (%)	57.8	76.8	84.2
Positive rate (%)	1.29	1.20	0.47
Detection rate (per 1000 screened)	1.97	1.56	1.65
Total cost (£)	381,105	262,261	184,324
% on administration	27	32	46
% on tests	41	35	39
% on investigation	31	32	15
Cost per cancer detected (£)	3,347	3,787	2,989

In the second round of screening 533 people test positive and require investigation but this figure falls sharply on the third round to 176 (i.e. only seven per month). The total cost in the second round of screening is 31% lower than in the original screening round but the average cost is 13% higher; fixed costs are still substantial but less people are being screened since screening is only offered to previous acceptors. (Other rescreening policies are considered in Chapter Six). This is reflected in the costs of the third and subsequent rounds of screening where administration costs constitute almost half the total.

This assumes that the MRC screening interval of two years is appropriate; several sets of American guidelines endorse annual FOB screening, (e.g. American Cancer Society). Little evidence is available which would allow a comparison to be made within the above framework. However, it has been reported that most interval cancers occur in the first twelve months after screening (Allison and Feldman (1990)). This implies little benefit from reducing the screening interval to one year. Case-control studies of Haemoccult screening have confirmed that the claimed protective effect of FOB screening does not decline greatly from the first to the second year after the test is completed; there is little protective effect by the third year, however (Selby et al. (1993)). A screening interval of three years or more would result in many interval cases in the third year, substantially reducing the ability of the programme to detect cases at an early stage: a sensitivity rate with a four-year interval of 25% has been recorded (Allison and Feldman (1990)).

Who should be rescreened?

An alternative to only offering rescreening to previous acceptors is to concentrate efforts on persuading those who have previously declined screening to participate. In the MRC trial, two methods have been used. Firstly, non-acceptors have a card placed in the records of their general practitioner to prompt discussion at future consultations. Secondly, a postal invitation was sent to a sample of non-responders to the initial offer inviting them to request a test. Trials of these methods are still at an early stage, hence the numbers are rather small. Overall, 13% of previous refusers have accepted a subsequent invitation. The relative potential benefits of rescreening previous acceptors versus previous refusers is considered in more detail in Chapter Six.

Sensitivity analysis

Many assumptions have been used to derive the model; how robust are the results to relaxation of those assumptions? One way to look at the robustness of the model is to consider the effects of an equi-proportional change in each of the variables; this gives some indication of the relative importance of each variable:

Table Three

Variable increased by 10%	% change in average cost
Detection rate	-9.1
Participation rate	-6.3
Positive rate	+3.0
Haemoccult cost	+3.0
Target population	-2.5
Colonoscopy cost	+1.9
Postage cost	+1.1
Administration costs	+1.0
Office costs	+1.0
All other variables	<1.0

The detection rate emerges as the most important variable in determining the average cost per case detected, in line with the findings of other studies (Eddy (1991); Wagner et al. (1991)).

An increase in the participation rate reduces the average cost because more cancers are detected; however, more people now test positive so the total cost rises by 3%.

An increase in the positive rate with no offsetting increase in the detection rate is akin to a reduction in the specificity of the test; the effect is similar to abandoning the retesting protocol described in Chapter Two.

Increasing the target population spreads fixed costs more thinly but total costs rise by 7%; this assumes that the extra work involved can be handled within the existing administrative structure at little extra cost.

The cost of colonoscopy is also significant; while the cost calculations in the appendix are quite detailed they too encompass many assumptions.

The remaining variables have only a limited impact on the model, implying that these variables do not merit further detailed attention.

How can the cost per case detected by Haemoccult be reduced? Fixed administration costs represent between 14 and 28% of the total and the programme cannot influence the costs of the tests or of postage; differences in development costs are trivial. However, there is a potential for savings in asking those people who do not wish to complete a screening test to return it for re-use: if half of the non-responders did so the total cost would be reduced by £24,000 for three-day and by £53,000 for six-day testing. In the latter case this represents a 10% saving on total costs.

Another way to analyse robustness is to consider a number of different scenarios which may arise. This method also gives some indication of which of the variables is most likely to change from its baseline value. A selection of scenarios is presented below.

- The implementation of a national programme results in bulk purchasing of tests and a 10% discount on the purchase price per test is negotiated: the average cost falls by 3%.
- The local costs of investigations are 10% higher than those in Nottingham owing to practice variations: average cost increases by 2.2%.
- Local circumstances mean that the catchment population is 10% larger than that described above; total costs rise by 7.3% but average costs fall by 2.5% as fixed costs are spread more thinly.
- As a result of the publicity surrounding the implementation of a national screening programme participation achieves the breast cancer screening target of 70%; total costs rise by 6.5% while the average cost per case detected falls by 12.1%.
- FHSA registers are found to be very inaccurate, reducing participation by 20% and requiring 50% more administration and clerical time to correct. The average costs rises by 24% as fewer people complete a test and the yield falls, but the net effect on total costs is close to zero since the savings on investigating fewer people almost exactly offset the extra administration costs.
- Access to a fully-equipped endoscopy suite is not available, forcing surgeons to use the combination of BEXR and flexible sigmoidoscopy; average and total costs rise by 3.8% (assuming the cancer detection rate is unaffected).

- Long-term follow-up reveals the sensitivity (and the detection rate) of the Haemoccult test to be 10% lower than is currently thought, increasing the average cost by 9.1%. A 20% fall in sensitivity increases the average cost by 25%.
- As the programme relies on the postal system to distribute and return tests, a 10% rise in postage costs would appear crucial but the average cost only rises by 1.1%.

These relative effects are shown in Figure Two. Further factors can also be considered once survival data are available to convert cost per case figures into cost per health gain. A notable example would be the effects of inter-surgeon variability in survival rates; such variations have been previously demonstrated to exist, independent of case-mix and other measurable factors (McArdle and Hole (1991)).

Other limitations of the model

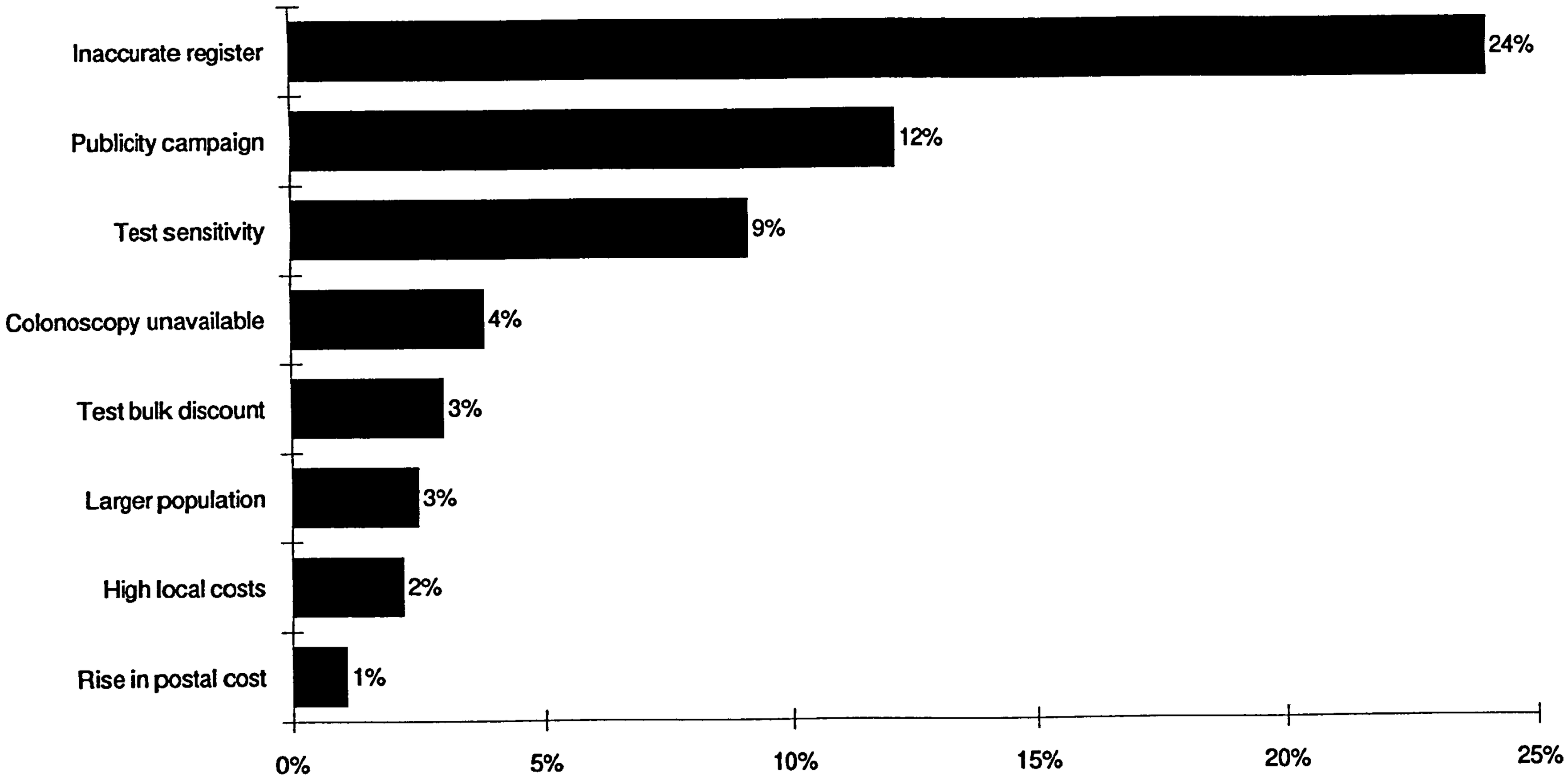
Diagnostic investigation costs departed from the health service perspective stated by including costs to patients: these represent 20% of the cost of colonoscopy and 22% of the cost of the combined radiology/sigmoidoscopy investigation. As the above figures show, excluding these figures would not have a significant effect on the results (i.e. less than 5% off the average cost). If the aim is ultimately to compare the costs and benefits of colorectal cancer screening with those of other uses of the resources, however, it may raise problems of comparability since so few other studies include these costs, despite the importance of the societal viewpoint in economic evaluation textbooks.

While the sensitivity analysis has helped with the variables that were included in the model it has not considered those factors which are excluded and therefore are implicitly assumed to have no effect. Two of the most significant exclusions are the net effect of screening on the costs of treatment and the health effect of adenoma excision. These are considered in more detail in later chapters. A further category of benefit excluded is the value of information to the individual: no satisfactory means of eliciting this value have yet been developed. It is, therefore, assumed that the benefits to the patient at least outweigh the costs of performing the test, awaiting the results, attending the clinic, and so on. On this basis, excluding both costs and benefits means that, from the patients' perspective, the results are the upper limits on the net societal cost per case detected.

Conclusion

The cost per case for a 'typical' target population was derived from a set of assumptions based on the MRC trial. It may be possible to rule out screening for some diseases on this basis alone if the cost per case is so high that it could not be justified by any plausible magnitude of health gain. FOB

Figure Two - Effect on cost of scenario changes



screening for colorectal cancer does not fall into this category, thus it is worthy of more detailed consideration. Ideally, data on health gain would be used but in advance of the analysis of the MRC trial results this is not possible at this stage.

Sensitivity analysis reveals that the key variables underlying the results are the performance of the screening test, the participation rate and the size of the target population; each influences the likely costs and yield. Other factors excluded to date also require consideration including treatment costs and the benefit of adenoma excision. These are considered in more detail in the following chapters.

APPENDIX

HOSPITAL COSTS OF DIAGNOSTIC PROCEDURES FOR COLORECTAL CANCER

Introduction

Diagnostic investigation of the colon and rectum can be undertaken either by radiology or by endoscopy. The barium enema X-ray was virtually the only technique of visualising the colon up until the 1970s. This involves coating the lining of the colon with barium as a contrast medium; the development of the double contrast method has enhanced accuracy, especially for small lesions. Endoscopy has been available for many years in the form of the short (25cm.) rigid sigmoidoscope, but developments in the 1970s led to the introduction of longer, flexible instruments using fibreoptic technology to allow direct visualisation of the colon. The colonoscope is capable of reaching the caecum if fully inserted, while the 60cm. flexible sigmoidoscope reaches beyond the sigmoid colon and towards the splenic flexure.

Resources used include: staff time, drugs, films and chemicals, pathology specimens, equipment depreciation, hospital overheads, the opportunity cost of the building, in-patient stay (for observation or to treat complications), and costs to patients of attending. Assumptions based on observations of all three techniques and discussed with the staff involved, are set out below.

Colonoscopy is performed by a consultant conducting seven investigations each day in a dedicated endoscopy unit with supporting nursing and clerical staff. The total staff costs of the unit are divided across capacity workload, calculated as the maximum number of colonoscopies that could be performed in one year. At seven patients per day the throughput of a single suite of the unit would be 1,820 cases; assuming there are two suites then half the nursing and support staff costs can be allocated to each. All patients receive a laxative and most are sedated. If a polyp is seen on examination, it is removed by snare diathermy via the colonoscope wherever possible and is sent to the pathology laboratory for histological analysis. The biopsy rate is dependant on the number of abnormalities; in the calculation the case-mix arising from the MRC trial is used.

Barium enema X-rays are performed by a consultant radiologist, with a radiographer and an auxiliary nurse present. Each day, twelve patients can be investigated; this requires a total of two hours of clerical time to administer. Patients receive a laxative but are not sedated. Each patient requires a barium contrast medium and ten films; chemical costs are calculated according to a 'rule-of-thumb' used in radiology departments.

Flexible sigmoidoscopy is performed in clinic by a consultant with an auxiliary nurse present. Twenty patients can be seen in one day, requiring two hours of clerical time to organise. Patients are given a laxative but no

sedation. The biopsy rate is assumed to be 72.9% of that of colonoscopy, based on MRC trial data for the distribution of cancer in the bowel.

Thus the staff, drug, film and chemical, pathology and equipment costs are as follows:

Staff costs (£)
The Personnel Department at University Hospital, Nottingham (UHN) supplied the following figures:

Table Four

	Gross annual cost (£)	Cost per day (£)
Consultant	38,379	147.60
Endoscopy staff (total)	47,885	191.50
Auxiliary nurse	6,676	25.70
Radiographer	12,543	48.20
Clerical officer	6,303	24.20

Gross annual cost represents the full NHS costs of employment (salary, national insurance and pension contributions). Note that the endoscopy staff cover two suites.

Drugs and disposables
Costs are based upon figures in the British National Formulary, March 1989.

Table Five

	Quantity used	Cost (£)	
Sodium picosulphate	2 sachets	0.60	All CS and BXR patients
Sodium phosphate enema	1 sachet	0.43	All patients
Diazepam	20mg.	0.00	95% of colonoscopies
Hyoscine butylbromide	20mg.	0.19	25% of colonoscopies
Naloxone hydroxide	0.4mg.	3.57	5% of colonoscopies

A preliminary assessment of the costs of disposable items, such as pads and syringes, indicated that these were negligible and they are excluded.

Films and chemicals
The barium X-ray medium costs £4 and one film costs £0.88 with 10 used per patient (Radiology Department, UHN). The total cost per procedure is £14.34 (= £4 + (10 x £0.88) + £1.54 for chemicals valued at 17.5% of film costs).

Biopsy rate and cost
The following is based on discussions with staff of the Histopathology Department, UHN). The cost of a biopsy is £5.76, comprising £4.77 for labour, £0.56 for equipment depreciation and £0.43 for materials. The biopsy rate in the screening study is 1.51 per colonoscopy and 1.1 per flexible sigmoidoscopy based on the following assumptions:

12.3% of investigations find a carcinoma, requiring 4 biopsies;
37.3% find adenomas, requiring 2 biopsies;
27.2% find other abnormalities, requiring 1 biopsy;
22.8% are normal, requiring no biopsy.

Equipment depreciation and maintenance

All equipment is valued at replacement cost, excluding Value-Added Tax. A simple division of the cost by working life is made and the annual cost, including maintenance, is then divided by the annual workload.

Colonoscopy

Colonoscope and forceps cost £9,998 and last for one year if used full-time under the workload assumption; annual maintenance costs £1,000 (figures from Keymed Supplies). The electro-surgical unit, suction apparatus, video display unit, trolley and monitor cost £19,144 in total but last for ten years.

Barium enema X-ray

The X-ray machine itself costs £240,000 and lasts for ten years with annual maintenance costs of £4,630 (figures from Radiology Department, UHN). A processor costs £10,000 with a ten-year life. Cassettes cost £3,000 for 15 and last for five years.

Flexible sigmoidoscopy

Sigmoidoscope and forceps cost £3,740 and are assumed to last for one year if used intensively; annual maintenance costs £250.

Costs to patients of attending

There are costs to patients in attending for investigation including the cost of travel and the value of the person's time. A survey of patients attending a follow-up clinic for colorectal cancer at UHN measured the time and travel costs involved. Time was valued using the Department of Transport's figures for the value of travel time savings updated to £2.23 per hour:

Table Six

Mode	Cost (£)	% using mode
Public transport fare	1.90	35
Car mileage	0.55	60
Walking time	1.55	5
Vehicle time	1.19	95

The expected cost of attending the hospital is £2.90. In addition, the person will be in the hospital for about two hours, valued at £4.46. About 10% of people are accompanied by a friend or relative: this increases the expected cost by £0.29. The average total cost of attending the hospital is £7.65. None of those interviewed reported any lost income from paid employment as a result of attending.

This figure is taken as a proxy for the costs of attending for a flexible sigmoidoscopy or BEXR, with an adjustment to allow for the person being in hospital slightly longer. For a colonoscopy, however, sedation is necessary and hence patients are advised not to drive themselves home. Assuming that these people therefore use public transport, the cost of the journey is £3.72. While the procedure itself is relatively quick, the patient will be required to stay in hospital for about five hours, including time to recover from sedation. If a similar proportion are accompanied the total cost is £15.27. This also assumes no lost income as a result of the procedure. (Note that in the Australian breast cancer screening study 52% were accompanied (Hurley and Livingstone (1991)); this would raise costs to £16.49).

In-patient costs and general service overheads

In-patient admission can occur if the patient requires overnight observation before or after investigation. The experience of the MRC trial is that 6% of colonoscopy patients are kept in for overnight observation while about 0.6% are admitted for two nights. An expected in-patient cost based on these figures is included. Admission following the other investigations is rarer; in the absence of similar series admission rates of 0.5% following BEXR and of 0.1% following flexible sigmoidoscopy are assumed.

A diagnostic investigation can also be the cause of morbidity and mortality. Treatment of complications such as haemorrhaging and perforation of the colon require in-patient care and even surgery. If care for minor complications is included in the above admission rates then the proportion of patients experiencing major complications from each type of procedure is taken to be as follows: 0.09% of colonoscopies (Hall et al. (1991)), 0.04% of BEXR (Ott et al. (1985)) and 0.05% of flexible sigmoidoscopies (Selby and Friedman (1989)). A major complication requires surgery and the total in-patient stay is valued at £2,000; this figure is tested in the sensitivity analysis.

Total general service expenditure from UHN cost returns is divided by a factor reflecting the time each type of patient (inpatient, outpatient, day cases) spends in the hospital, to give a cost per hour. In 1989-90 total spending on general overheads at UHN was £14,266,694. Assumptions are required on the number of hours patients of each type spend in hospital, as follows:

Table Seven

Type	Number	Duration (hours)	Total
In-patient days	372,188	24	8,932,512
Out-patient attendences	309,730	2	619,460
Day cases	7,888	8	63,104
A&E attendences	166,913	3	500,739

In total there were 10,115,815 patient-hours using the assumed durations. These require adjustment to reflect the fact that the day of admission plus the

day of discharge of an in-patient admission are only part days. Assuming that they make up one full day between them then allowing for the number of episodes (69,179) gives a revised total of 7,272,216 patient-hours. The average overhead cost per hour is £1.70 and the overhead costs of each type of contact are £40.80 per in-patient day, £3.40 per out-patient attendance, £13.60 per day case attendance, and £5.12 per A&E attendance.

Colonoscopy patients are assumed to occupy five hours of hospital time (at £1.70 per hour), while the other procedures require three hours since there is no need to recover from sedation. From cost returns, the in-patient cost for general surgery comprises £40.80 for service overheads (i.e. 24 hours at £1.70 per hour) and £76.10 for patient care excluding theatre, diagnostic and support service costs.

Opportunity cost of the building

An allowance must also be made for the opportunity cost of the building. Ridley et al.'s estimate an opportunity cost in terms of vacant rental value of £10 per square-foot per year (in 1989/1990) and an area of 2,702 square-feet for an intensive therapy unit. This gives an annual total of £27,020 or £103.92 per working day. Assuming that endoscopy and radiology suites are of a similar size, the same costs can be allocated across total workload.

Results

The average costs per investigation borne by the hospital for the three diagnostic techniques are:

Table Eight

Category	**Colonoscopy**	**BEXR**	**Flexible sig**
Medical staff	21.09	12.30	7.38
Nursing and support staff	13.18	12.62	7.75
Drugs and disposables	1.26	1.03	0.43
Films and chemicals	0	14.34	0
Biopsy	8.70	0	6.34
Equipment	7.09	9.69	0.77
General services	8.50	5.10	5.10
Building	14.85	8.66	5.20
Inpatient stay	8.04	0.70	0.14
Complications	1.80	0.80	1.00
Patient costs	20.63	13.74	13.74
NHS variable	57.35	40.30	21.90
NHS direct	87.79	63.74	32.96
NHS total	97.63	65.13	34.08
NHS plus patient total	105.10	78.87	47.81

Definitions:

NHS variable cost = medical & nursing staff, drugs & disposables, biopsy

NHS direct cost = NHS variable cost + equipment, overheads and building
NHS total cost = NHS direct costs + admissions and treatment of complications
NHS plus patient total = NHS total + patient cost

Rather than considering BEXR and flexible sigmoidoscopy individually it would be more realistic to consider them as a combined strategy. The components of the costs of each option are compared in Figure Three. The choice of diagnostic strategy, comparing the test cost against accuracy, is considered in more detail in Chapter Nine.

Sensitivity analysis

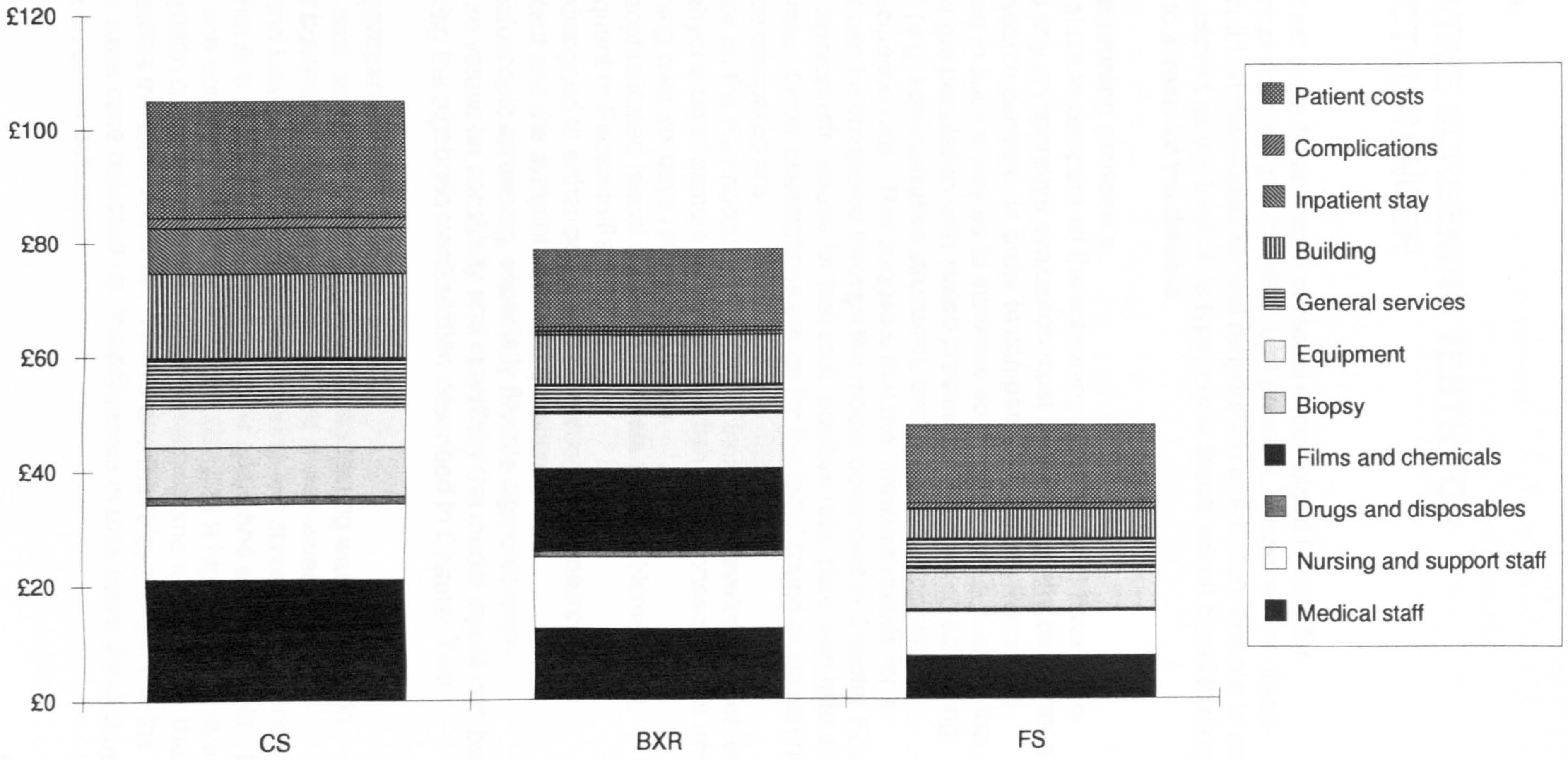
Clearly, the above calculations rest upon a number of assumptions and thus an analysis of robustness is required. Each of the variables was increased by 10% and the effect on the cost of each diagnostic procedure noted (the NHS plus patient definition was used). The most important variables are listed in Table Nine below, showing changes relative to the baseline values in Table Eight above:

Table Nine

Variable changed by 10%:	**Colonoscopy**	**BEXR**	**Flexi sig**
Throughput	-4.9	-4.2	-2.8
Working days per year	-3.6	-4.0	-3.0
X-ray film cost	0	+1.8	0
Consultant salary	+2.0	+1.6	+1.6
Time patient is in hospital	+2.0	+1.6	+2.6
Opportunity cost of space	+1.4	+1.1	+1.1
Endoscopy staff costs	+1.3	0	0
Value of patient time	+1.2	+0.9	+1.5
Overhead costs	+1.1	+0.7	+1.1
Clerical officer salary	0	+0.8	+1.4
Pathology cost per specimen	+0.8	0	+1.3
Equipment replacement cost	+0.6	+1.0	+0.2

None of the other variables has an impact greater than 1% for any of the test costs. In general labour costs are more important in endoscopic procedures and capital costs are more important in radiological procedures, despite the extreme assumption that endoscopes are replaced every year. The analysis reveals that the results are quite robust although account should be taken of local workload factors when they are being generalised for use elsewhere. Increasing throughput and reducing the numbers of support staff would reduce costs, although this could affect the quality of care and ultimately increase total costs if more repeat procedures were required.

Figure Three - Investigation costs



Chapter Five

ALTERNATIVE SCREENING TESTS FOR COLORECTAL CANCER

Introduction

Chapter Four set out a basic model of faecal occult blood testing for colorectal cancer according to the MRC trial protocol. However, the three-day Haemoccult test developed without rehydration of the stool sample is not universally accepted as the best of its type, nor is faecal occult blood testing the only way to screen for the disease.

Alternative screening protocols

No clinical trial could compare all the screening protocols that have been advocated so any comparative evaluation must be based on the data from a number of research centres. In order to compare the options the results should be used in such a way as to minimise sources of bias e.g. apply the figures to a single population with stated prevalence, a common screening 'environment' (e.g. administrative structure), and (where appropriate) a common participation rate. This suggests that the literature results for a variety of test can be compared through the model described in Chapter Four but using the appropriate values for test cost, positive rate, detection rate and participation rate. Other assumptions are as for the initial round of screening. The tests to be compared are:

1. Variants on the Haemoccult test protocol including development of test with rehydration of sample as in the Swedish and Minnesota trials and screening over six days rather than three.
2. More sophisticated faecal occult blood tests such as Hemeselect, Hemoquant or Fecatwin/Feca EIA.
3. Tests designed to enhance participation such as Coloscreen, Ez-Detect and the symptom (or risk) questionnaire.
4. Sigmoidoscopic screening, especially flexible sigmoidoscopy.

Given literature values on sensitivity and specificity the model inputs can be calculated using the algebraic manipulation described in Chapter Two.

Method for comparing the options

The average cost per case detected for three-day testing was reported in Chapter Four but the potential for this to mislead is well-established (Neuhauser and Lewicki (1975)). When comparing two screening protocols it is assumed that A is preferred to B if A has lower costs and a higher yield. In cases where one option is more expensive but also has a higher yield there is no simple decision criterion; the choice rests on a judgement of whether the extra yield justifies the extra expense. To this end the relevant figure is the extra cost per extra case detected i.e. the difference in total costs divided by the difference in cases detected.

Haemoccult protocols (without rehydration)

The greater sensitivity of six-day FOB testing to known cancers was commented upon in Chapter Two. As part of the MRC trial the two test periods have been compared and the relative performance is set out in Table One:

Table One

	3-day	6-day
Cost of test (£)	1.13	2.26
Cost of sending test (£)	0.41	0.60
Rate of developing test (number per hour)	60	30
Participation rate (%)	57.8	53.8
Positive rate (% of those completing)	1.29	1.69
Cancer detection rate (per 1000 acceptors)	1.97	2.56

Six-day testing has a lower participation rate resulting from the extended testing period and higher screening costs (two three-day tests, greater postage and slower test development). Extended testing detects more cases of disease which bleed sporadically but at the expense of a greater number of people giving false positive reactions, whether through 'natural' blood loss or dietary factors. One result of this is a higher total cost of diagnostic investigation. The detection rate difference in the clinical trial was not significant given the sample size, but assuming that a larger study would confirm the results then six-day testing on the same basis would have a sensitivity of 84% compared to 65% for three-day testing (Thomas et al. (1992)).

Applying the data to the model from Chapter Four gives the following results:

Table Two

	3-day	6-day
Number testing positive	746	902
Total cost (£)	372,477	531,708
% on admin	28	20
% on tests	43	55
% on investigation	29	25
Cases detected	114	136
Average cost (£)	3,271	3,920
Extra cost per extra case (£)		7,238

In comparison with the three-day testing, the incremental yield and cost of six-day testing is 22 cancers and £159,321, giving a cost per extra cancer detected of £7,238.

Note that the difference between the specificity rate of the two tests is 0.3% (98.6% and 98.9% for 6- and 3-day testing respectively) which could be lost as a rounding error but the cost of this difference is 134 extra false positives

(156 extra positives minus 22 extra true positives). The health service costs are over £14,000 or 9% of the difference in costs of the two options. Care should therefore be taken to be precise when estimating the specificity of a screening test.

The above data are based on a comparison of 3- and 6-day testing in randomised groups; an alternative, which has also been used in the MRC trial, is to use a single population but to number the order in which the squares on the test were completed. A sample of 9,000 asymptomatic people were offered a six-day test with the squares numbered and the respondents asked to provide samples in the order specified. Of the 140 people who completed a test which was found to be positive, 70% had at least one positive test out of the first six squares numbered. Positive results were retested and persistent positives investigated, according to the usual protocol. Of the 41 who would have been negative over three days of testing, 27 were still negative. Two of the 14 investigated had cancer while five others had large adenomas, compared to nine and seven cases respectively in the 105 people who had one or more positives out of the first six squares.

The cost of this extra yield consists of the extra tests sent to the whole group, retesting those positive on the second six squares plus the subsequent diagnostic investigations. The extra cost of six-day testing is £1.32 (test plus postage plus £0.08 for the extra time to develop the tests returned). To this must be added 41 retests at £3.03 each (six-day test, postage, development) and 14 colonoscopies. The total cost is £12,959 (of which 85% is on the extra initial tests). The cost per extra cancer detected is thus £6,480, within 12% of the figure calculated as the incremental cost of six-day testing as compared to three-day testing above.

Rehydrated Haemoccult protocol

This protocol is modelled from the Swedish trial which reports a sensitivity of 85% and specificity of 94.5% (Kewenter et al. (1988)). Note, however, that this trial is aimed at people aged between 60 and 65; the positive rate is known to be age-related and hence may be different in another population. The model inputs are as follows:

Table Three

	Rehydrated case
Cost of test (£)	1.13
Cost of sending test (£)	0.41
Rate of developing test (number per hour)	45
Participation rate (%)	57.8
Positive rate (%)	5.8
Cancer detection rate (per 1000 acceptors)	2.54

Rehydration of the sample prior to development has no effect on participation, on the cost of the test or on the cost of postage. The sensitivity of the test is increased but at the cost of a greatly increased number of false positives.

Table Four

	Rehydrated case
Number testing positive	3,352
Total cost (£)	757,891
% on admin	14
% on tests	21
% on investigation	65
Cases detected	150
Average cost (£)	5,043
Extra cost per extra case (£)	10,706

The rehydration method detects more cancers than unhydrated testing but at greatly increased cost - total and average screening costs are almost doubled and cost per cancer detected is 54% higher. The extra test development costs of rehydration is a minor factor; the principal source of the difference is the reduced specificity resulting in more diagnostic investigations of healthy people. Rehydration increases the yield of cancers by 36 relative to unhydrated three-day testing at an additional cost of £385,414 i.e. at an incremental cost of £10,706 on average (approximately three times the average cost without hydration). In comparison with six-day testing the incremental cost is £16,156 per extra case.

Other FOB protocols

Hemeselect is an immunochemical test specific to human haemoglobin i.e. dietary factors cannot cause false positive results. While there are no published data on the performance of the test in an asymptomatic population of comparable age to that considered here, one study has reported results for a study of people attending a surgical out-patient clinic with lower gastrointestinal symptoms (Thomas et al. BJC 1992). The results indicate a sensitivity of 94% and a specificity of 84.1%. Note, however, that the data source may underestimate the specificity as the sample used is more likely to have rectal bleeding than the general population: in the same group the three-day Haemoccult test (unhydrated) had a specificity of 96%, compared to the 98.9% found in the general population. Assuming that Hemeselect would experience the same proportionate increase in specificity the following model inputs are used:

Table Five

	Hemeselect
Cost of test (£)	2
Cost of sending test (£)	0.41
Participation rate (%)	57.8

Positive rate (%)	13.6
Cancer detection rate (per 1000 acceptors)	2.69

The cost of developing the test is approximately £4, including technician time, chemicals and overheads. Once again an improvement in the sensitivity of the test is 'purchased' at a greatly increased false positive rate.

Table Six

	Hemeselect
Number testing positive	7,861
Total cost (£)	1,733,334
% on admin	6
% on tests	27
% on investigation	67
Cases detected	155
Average cost (£)	11,148

Compared to three-day (unhydrated) Haemoccult screening the extra cost per extra case detected is £33,192; the figure compared to six-day testing is £63,243 and compared to three-day (rehydrated) Haemoccult is £205,089.

Hemoquant provides a quantitative measure of the blood in the faecal sample and thus allows an explicit choice as to the criterion for a positive test result, allowing a trade-off of sensitivity against specificity. The feasible trade-offs are:

83% sensitivity, 92.7% specificity if levels greater than 1.5mg of haemoglobin per gramme of stool constitutes a positive test;
74% and 94.7% if 2.0mg. is the cut-off; and
62.9% and 97.3% if 3.0mg. is the criterion (St John et al. (1992)).

None of these options is preferred to three-day Haemoccult (rehydrated) and the test is far more expensive to develop. Thus, assuming that Hemoquant has no advantage in terms of improved participation, it would not be chosen as the most efficient screening test.

Fecatwin/Feca-EIA is a combination test intended to follow a highly sensitive test (sacrificing specificity) with a more specific test and only investigating those positive on both. This test was used in the early stages of the MRC trial and has also been tested in a separate group (Pye et al. (1989)). The results indicate that it has a sensitivity of 67% and a specificity of 91%; it is therefore outperformed by the more sensitive Haemoccult options since it has no cost or participation advantage.

Tests designed to maximise participation

Coloscreen and ***Ez-Detect*** are self-completion tests designed to maximise participation. The results are interpreted and reported by the patient with the help of a reference card showing what a positive and negative test should look like. Tests commonly involve dropping prepared cards or sheets of

paper into the toilet bowl after passing a motion with testing over three days; haemoglobin produces a change in colour.

Both tests have been evaluated using samples of surgical out-patients with symptoms of lower gastrointestinal disease. The results of the two studies are remarkably similar: Coloscreen was completed by 86% and was found to have a sensitivity of 33% and a specificity of 94.1% (Pye et al. (1990)); the comparative figures for Ez-Detect were 88%, 36.4% and 89.3% respectively (Tate et al. (1989)). These convert into positive rates of 6.08% and 11.08% and detection rates of 1 and 1.09 per 1,000 acceptors for Coloscreen and Ez-Detect respectively. Assuming each test to cost £1 plus £0.41 postage, the model provides the following results:

Table Seven

	Coloscreen	**Ez-Detect**
Number testing positive	5,229	9,750
Total cost (£)	1,014,635	1,680,333
% on admin	10	4
% on tests	14	5
% on investigation	76	91
Cases detected	86	96
Average cost (£)	11,798	17,518

Thus, both tests are easily dominated by the more sensitive Haemoccult protocols. A later study (Tate et al. (1990)) confirmed that in a further sample of out-patients with large bowel symptoms Ez-Detect was significantly less sensitive than Haemoccult with no gain in participation. These tests are not considered further.

The ***symptom questionnaire*** is a postal survey of bowel symptoms, allowing an assessment of risk without faecal sampling. The experience in Nottingham is that the response rate is poor (34%) with 24% of responders reporting one or more symptoms (Farrands et al. (1984)). Further investigation revealed that none of these people had cancer; however, two asymptomatic cancers were later detected by Haemoccult in the same group. Even the most optimistic data indicate that a questionnaire has too high a false positive rate to be preferred to the more sensitive Haemoccult tests.

Endoscopy screening

Some researchers have supported colonoscopy as a screening test (Reasbeck (1987)). The cost of colonoscopy is £105.10; pooled trial data give a sensitivity of 94% (England et al. (1989)). Assuming the same participation rate as for FOB testing and a specificity of 100%, the model predicts the total cost will be £6,178,598 (57,800 colonoscopies plus admin costs) and 165 cases will be detected at an average cost of £37,446. In comparison with the next most sensitive option, Hemeselect, the extra cost per extra case is £439,526 each.

Given the availability of endoscopic facilities support for a flexible sigmoidoscopy screening programme would appear to be more realistic (Atkin et al. (1993); Jass (1989)). A postal invitation to 100,000 people would cost £0.33. Claims of 70% participation seem ambitious (Atkin et al. (1993)): of three supporting references, two are of doubtful comparability (based on a Norwegian population and a single general practice in Hertfordshire). The third reference is incorrectly quoted: the true figure is 44%. In the MRC trial 46% of the population accept three-day testing with no prompting implying that they are well motivated in health matters. As this is in line with the reference mentioned above, participation of 45% is assumed.

The prevalence for the target population is 3.03 per 100,000 of which 61.6% are in the range of the flexible sigmoidoscope (Atkin et al. (1993)); within this range a sensitivity to cancer of 98% and a specificity of 100% is assumed. Given that the cost per examination is £47.81 (from Chapter Four), the total cost is £2,288,268 (of which 94% is the cost of the investigations). The yield is 134 cases, at an average cost of £17,076 each. In addition the protocol proposed would colonoscope everybody with a 'high-risk' adenoma (diameter of 1cm. or more, villous element) representing 5% of the those screened; this adds £236,475 to the total and a few cancers may be found beyond 60cm. As a whole, the option is inferior to 6-day Haemoccult testing and rehydrated Haemoccult testing. If the claims of 70% participation were justified the cost would rise to £3,483,518 (plus £367,850 for colonoscopy) and the yield to 208 cases, at an average of £16,753 each; this would be preferred to colonoscopy as the 'Rolls Royce' screening option.

Such a screening programme would have the detection and excision of adenomas as one of its main aims, yet no benefit is assumed here. The reduction in future cancer incidence likely to result from excision is discussed in Chapter Seven.

Comparison of options

The options evaluated can now be compared in terms of their incremental cost (i.e. the extra cost divided by the extra number of cases detected):

Table Eight

Test	Total cost (£)	Cases	Incremental cost (£)
3-day (unhydrated)	372,477	114	
6-day	531,708	136	7,238
3-day (rehydrated)	757,891	150	16,156
Hemeselect	1,783,334	155	205,089
Colonoscopy	6,178,598	165	439,526

The incremental costs are plotted in Figure One and the scatter of plots of cost and yield in Figure Two. Ideally a screening test should be as close to the origin of the graph as possible. This indicates the trade-off to be

Figure One - Extra cost per extra case

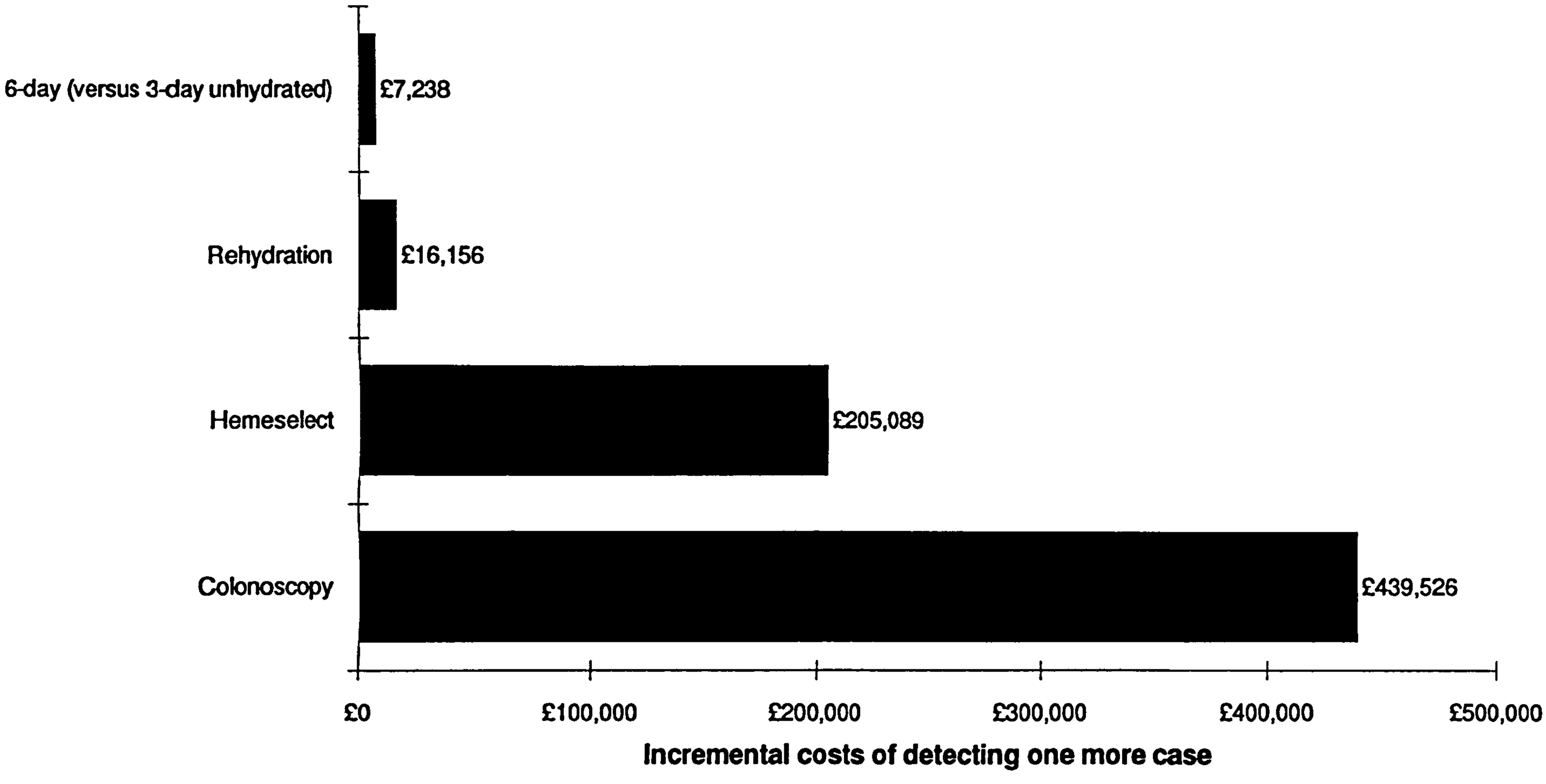
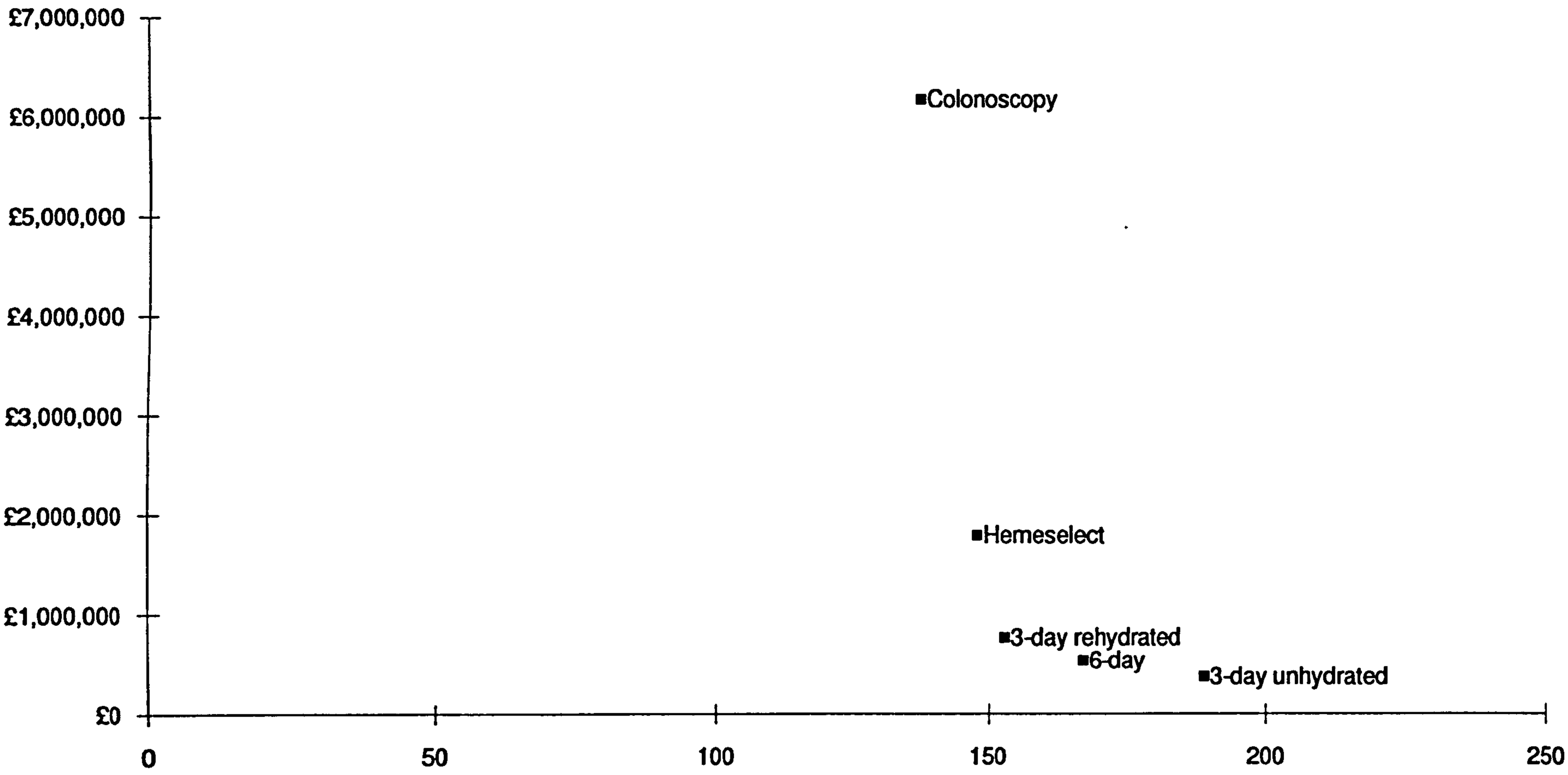


Figure Two - Cost versus cases missed



addressed as a policy issue. To resolve it, information on the health benefits of early detection will be necessary, allowing the cases to be converted into health gain.

Discussion

The same deficiencies apply to these results as to those quoted in Chapter Four, since the same model is used. However, several biases work against the more sensitive tests. As they detect more adenomas any benefit from preventing future cancers bears more on them. Similarly if there is a cost in terms of extra treatment costs to missing a case there should be an extra penalty to a missed case. A particular factor working against endoscopic screening is the one-off screening considered above: in fact, FOB testing will have to be regular, whereas the interval between endoscopic investigations can be considerably longer.

Given their positive rates it is unlikely screening by Hemeselect or colonoscopy would be considered with the resources currently available to the health service. Even rehydrated Haemoccult testing requires substantial health gains to justify its extra cost.

THE DEFINITION OF A POSITIVE SCREENING TEST

So far it has been assumed that any one positive square on a test means that the whole test is judged to be positive; it is possible that different diagnostic protocols should be adopted for each number of positive squares on a test. The issue can be explored using unpublished MRC trial data based on screening the general population.

Table Nine

	Six-day		Three-day	
Squares positive	**Number**	**Cancers**	**Number**	**Cancers**
1	45	2	114	8
2	69	7	117	12
3	30	3	48	5
4	19	4	28	4
5	29	3	23	4
6	19	2	31	9
7	17	4		
8	14	4		
9	3	0		
10	8	4		
11	2	1		
12	8	3		
Total positive	263	37	361	42

Since occasional bleeding among 'healthy' people is adequate to trigger a false positive, there are grounds for regarding one positive square as a

chance finding. On three-day testing this would have averted 114 colonoscopies (saving of almost £12,000) but the sensitivity of screening, already criticised for being too low, would fall by 20% of its original value. The implication is that so many cancers bleed only intermittently that all positive results (following retesting) should be subject to the same investigations.

The sixth stool guaiac re-revisited

The questions of the most efficient criterion for a positive Haemoccult test and of the appropriate number of squares to be used on such a test was first addressed by Neuhauser and Lewicki almost twenty years ago. Based on the results of an early clinical trial (Greegor (1969)) they considered the extra cost and yield of increasing the number of squares on a Haemoccult test. The rapidly diminishing gain in yield, approaching zero by the sixth square, when compared to the constant average cost per square, resulted in extremely high marginal costs.

This example of the contrast between average and marginal costs has become a celebrated piece of health economics but there are a number of reasons for doubting its relevance to colorectal cancer screening policy. First, the calculations were performed using results from a very small sample of symptomatic patients (n=278), amongst whom only two cancers were detected; as both of these cases had eleven out of twelve positive squares on their test this implied that there was very little benefit to having more than one or two squares on the test. As has been shown above, this is not the case in larger samples of asymptomatic people. Second, they assumed that each additional square would identify an identical proportion (11/12 or 91.7%) of the cancers remaining undetected by the previous number of squares. Third, the effect on the participation rate of increasing the number of squares was not considered. Fourthly, the specificity of each square was miscalculated, to the extent that the false positive rate for a single square, calculated as 37% in the original article, has been re-estimated at 0.47% by subsequent researchers (Brown and Burrows (1990)). Finally, the study used charges to value diagnostic investigations rather than resource costs.

While Neuhauser has defended his original work against the criticism directed at it (in particular that the re-estimated figures do not predict Greegor's results), he has also accepted that it is open to improvement (Neuhauser (1990)). Given the established manufacturing processes, the number of squares placed on a Haemoccult test is unlikely to change in the foreseeable future, although it is feasible that those offered screening could be requested to only complete (say) two of the six squares. What are the costs and benefits of implementing such a policy?

As noted above, the only experience available on the use of the Haemoccult test in an asymptomatic population is either over three or six days i.e. six or twelve squares. For each of the key variables in a screening programme

(participation, positive and cancer detection rates) there will be a function relating them to days of testing. Participation, for instance, is 53.4% for a six-day test but rises to 57.8% for a three-day test. The straight line joining these points has the equation:

$$\text{Participation rate (\%)} = 62.2 - 1.467.D$$

where D denotes the number of days of testing (at two squares per day). This implies that the maximum participation rate that can be expected in the MRC trial population is 60.7% with one day of testing. Similar equations for the other variables are as follows:

$$\text{Positive rate (\%)} = 0.89 + 0.133.D$$

$$\text{Cancer detection rate (per 1,000 acceptors)} = 1.38 + 0.197.D$$

An asymptotic functional form may be more intuitively appealing but the straight line appears to be an adequate approximation other than at the extremes: for example, the equations predict that 1.38 asymptomatic cancers per 1,000 people will be detected with no screening at all! However, for values of D of one or more they seem to offer a reasonable approximation up to six-days of testing; at this point non-linearities may set in as limit values are approached. Studies in symptomatic patients suggest little gain to testing beyond eight days (St. John and Macrae (1982)).

The model derived in Chapter Four can be used once the equations have been solved for each day of testing. Some examples are provided below:

Table Ten

Days of testing	Participation (%)	Positive (%)	Detection rate
1	60.7	1.02	1.58
3	58.7	1.29	1.97
5	54.9	1.56	2.37
8	50.5	1.95	2.96

Given the linear nature of the relationships, the incremental cost per extra case detected of adding an extra day to testing is virtually identical throughout the range. The extra cost per extra case as a result of extending the test period is approximately £7,000.

CONCLUSION

The simple model can be used to compare the costs and disease yields of screening by various means, although the exclusion of any benefits from the excision of adenomas biases the results against some options. A sub-set of options clearly dominates the remainder in terms of the cost and cancer yield; these protocols use three- and six-day Haemoccult testing (with and without

rehydration), an immunological test and colonoscopy. On the grounds of economic efficiency, none of the other options for screening should be considered given their current performance. The choice between the options within the sub-set depends on the exact health gains and the willingness of purchasers to pay for these.

Chapter Six

PARTICIPATION IN SCREENING

"Occult blood screening is unique in that it requires a combination of considerable preparation and action by the patient outside the medical office." (Knight et al. (1989)).

Introduction

Participation in faecal occult blood screening trials is voluntary, hence the proportion of people completing and returning a test is less than 100%, implying that the full health gain of screening (if any) is not realised. If participation is independent of prevalence then 50% uptake means half the cancers in the population are neither detected nor treated at an early stage. This could jeopardise the success of the whole programme, given the limited sensitivity of the Haemoccult test: 58% participation and 65% sensitivity as in the MRC trial of three-day testing means that at most 38% of cases benefit from early detection.

Participation in the American screening trials is 80% but these use volunteer groups; in an unmotivated population a figure of 15% seems more appropriate (Winawer et al. (1985)). Participation in Scandinavian countries has been closer to 70%.

Colorectal cancer is not a well-known disease: a Cancer Research Campaign poster described bowel cancer as two taboos in one subject. Initially, participation was low in the MRC trial but increasing sophistication in techniques of invitation and persuading non-responders to reconsider has raised the figure from 38% to over 60% in the later stages of the trial.

The economic significance of participation

This improvement has only been achieved through considerable effort devoted to persuading more people to take up the offer of screening: "It is essential that we approach the topic of FOB screening aggressively and with rigor to develop a means to increase participation" (Blalock et al. (1987)). Little attention has been paid to the economic costs and benefits involved: maximising participation must take some account of the costs involved. For efficiency, the expected benefits of increasing participation would be compared to the extra costs involved.

A rational individual's decision to participate in a voluntary screening programme depends upon a subjective cost-benefit calculus of expected value. However, the results of a number of studies (Jansen (1984); Klaaborg et al. (1986); Nichols et al. (1986); Pye et al. (1988); Silman and Mitchell (1984)) suggest that participation is also a function of the method of invitation to participate (e.g. whether the invitation is made by the hospital or the

general practitioner), whether the subjects receive their tests automatically or whether they have to obtain them for themselves, and so on. Participation can be increased by giving more appropriate information regarding the test and the disease, the specific targeting of individuals at risk (e.g. those with a family history of cancer) and reminders to complete the test. Each measure involves extra cost, but there is no general rule as to which is the most efficient.

The decision to participate in the screening programme

Subject participation in the major European colorectal cancer screening trials lies within the range 55-70%, implying that a significant minority of individuals in each case are not taking up the offer of screening. On the other hand, the vast majority of individuals who participate in a mass screening programme derive no health benefit. Research into cervical screening programmes (Elkind et al. (1989)) has identified technical reasons for non-participation including:

i) inaccessibility (e.g. unrecorded change of address or death);
ii) ineligibility (already screened or does not fulfil screening criteria);
iii) unsuitability (medical problem prevents completion of test);
iv) communication failure (e.g. offer of screening not received, target illiterate).

These seem equally appropriate to faecal occult blood testing: the potential size of the 'wrong address' problem in existing age-sex registers has been noted already. The most significant set of variables relate to attitudes to and perceptions of screening, however.

The Health Belief Model (Rosenstock (1975)) offers a convenient way of thinking about the behaviour of the rational individual (taking all factors into account) when faced with this problem. The model attempts to predict individual behaviour on the basis of the individual's general health motivation, self-perceived risk of disease, perceived consequences of the disease, and any 'cues to action' in the individual's environment (Gillam (1991)). Variations between the behaviour of individuals can therefore be explained in terms of different perceptions, motivations and spurs. This suggests that the rational individual is more likely to participate if they perceive (i) that they are personally susceptible to the disease, (ii) that the consequences of the disease would be severe for them, (iii) that screening and subsequent treatment would result in an improved prognosis if the disease were present, and (iv) that the costs involved are small. It has been found that self-perceived risk is a far better predictor of screening participation than objectively calculated risk (Vernon et al. (1990)).

This framework can be used as a model of the expected net value of participation in screening. Participants obtain new information of one of two types - test results may be positive or negative for cancer. This information may prove either true or false. The individual will assign both values (or utilities) and subjective probabilities to each of the outcomes, yielding the

following expression for the expected value of information (I) received from the test result:

$$EV[I] = v_a p_a + v_b p_b + v_c p_c + v_d p_d \quad (p_a + p_b + p_c + p_d = 1)$$

where subscript 'a' refers to a true positive result, 'b' to a false positive result, 'c' to a true negative result and 'd' to a false negative result. This information is not free, and the costs facing the individual are the expected disutility and inconvenience incurred in undertaking the test (v_e), and the opportunity cost of time foregone (v_f). It follows that an individual will participate in screening if:

$$EV[I] > E[v_e + v_f]$$

(Note that in private health care systems there can also be a fee or charge for the test). An earlier formulation of this model used the expected net values of participation and of non-participation in cervical screening although the differences otherwise appear small (Haycox and Haran (1988)).

For a given individual, the subjective probabilities depend upon the individual's knowledge of cancer prevalence, prognosis and of test performance, as well as their perceptions of personal risk. True results take a positive sign while false results take a negative sign; one exception would be those who do not believe early detection increases the chances of cure, and they place a zero or negative value on v_a. A lack of confidence in the predictive power of the screening test would imply high values of p_b and p_d, thus inaccuracies in screening techniques are not widely publicised. Individuals 'convinced' they have cancer value p_a as one, and those equally convinced they do not value p_c as one.

The framework can be used to interpret typical justifications given in surveys of non-participants (Dent et al. (1983); Farrands et al. (1984)):

i) "Cancer worries me, I would rather not know and I'm afraid of what the test might show" implies high values for p_a and for v_a but the latter is negative.
ii) "I have no symptoms" / "I feel all right" / "I have not had this problem before so am not likely to have it now" implies p_c is one, p_a is zero and v_c is negligible.
iii) "The whole idea put me off" / "I couldn't face it" / "The test seemed messy, unpleasant, distasteful" indicates a high value for v_e.

Research findings are consistent with the predictions of the model:

- Participation is higher amongst those who perceive the effects of colorectal cancer as serious and who view screening as improving the prognosis i.e. v_a and v_c valued high and positive, p_b and p_d presumed low (Jansen (1984)).

- Past experience with FOB testing is a positive influence upon participation (Blalock et al. (1987)): ex post v_e is lower than ex ante expectations. In the MRC trial, re-testing of initial positives produces very high participation, suggesting consistency on the part of initial acceptors; those refusing the retest may be scared by the implied increased chance they have the disease.
- Screening that increases costs (v_f) to the subject reduces participation e.g. six-day testing, dietary restriction.
- The presence of gastrointestinal symptoms increases participation (Thompson et al. (1986)) presumably because of greater self-perceived risk of the disease (p_a high) and a high positive value to confirmation (v_a and v_c high).
- Embarrassment and distaste are significant causes of non-participation (v_e high) (Macrae et al. (1986)).
- Those who regard themselves as being especially susceptible to cancer (p_a high) are either particularly likely to participate (v_a high and positive) or particularly likely to refuse (v_a high and negative) (Macrae et al. (1984)).
- Tests which do not require taking stool samples (reducing v_e) increase participation (Pye et al. (1990), Tate et al. (1989)); but these sacrifice sensitivity and specificity, ultimately undermining p_a and p_c.
- Mant et al. (1990) found 96% of false positives would not be deterred from participating in screening programmes in the future; this implies that ex post v_b is small and negative.

Supporting evidence from studies of take-up of breast cancer screening includes the increased participation when a mobile mammography unit is used, reducing access costs (Haiart et al. (1990)). It was shown that a 10% increase in the 'crow-fly' distance from the unit to the individual's home reduced the uptake by 2.4%. Other studies have shown that women with positive health concerns and a knowledge of screening methods are more likely to take part (high positive values on v_a and v_c and realistically low values on p_b and p_d) (Vernon et al. (1990)). Many British women appear to believe in the effectiveness of mammography screening but do not perceive themselves as being at personal risk of having the disease (Fallowfield et al. (1990)). Other research indicates that it is better to educate people about screening rather than individual risk since this can be an important constraint on take-up (Vernon et al. (1990)). Non-participants are a heterogeneous group with either very good or very bad self-perceived health (Hunt et al. (1988)), implying feelings either of invincibility or of fatalism.

The model assumes rationality on the part of the individual, but is this always the case? There is some evidence that non-participants are coping with their fear of ill-health by not thinking about it and specifically by not considering their personal risk of getting ill or of dying (Fallowfield et al. (1990)). Making the calculation may itself be costly to the individual in terms of the anxiety it induces: from this perspective it may be an entirely rational decision.

Health education based upon the health belief framework has been used as an enclosure to the Haemoccult test at the initial screening invitation in the MRC trial (Pye et al. (1988)). The leaflet was designed to persuade the target population: (i) bowel cancer was more common in their age-group than they probably believed (p_a is high); (ii) "unlike most cancers it's the one most treatable if caught early", i.e. v_a was strongly positive; and (iii) a negative result is reassuring (v_c is high). No mention was made of the possibility of a false test result (implying p_b and p_d were zero). However, there was no significant difference between the group randomised to receive health education and the group who were not, implying that the invitation letter alone "recruits all those susceptible to postal recruitment and there may be little scope for further improvement". Another study reported similar results (Nichols et al. (1982)). The implication appears to be that health education is either ignored or amounts simply to preaching to the converted.

Methods of inviting people to complete screening

The findings from the last section suggest that the method of screening invitation might influence the individual's decision through the costs to the individual. Health education offered with the invitation could be more personal (e.g. from the individual's GP). This section draws on several British studies of participation in an asymptomatic population of similar age and with no special *a priori* motivation for undertaking screening. Note that while the remainder of the evaluation uses the viewpoint of the NHS purchaser, it is appropriate in this example to also consider the viewpoint of the individual and the costs involved in participating.

Five different methods of invitation were compared varying the source of the invitation between the individual's GP, a health visitor or a postal invitation from a screening unit. The alternatives are as follows:

1) A test is posted to each member of the target population from the screening unit (Thomas et al. (1990));
2) Individuals are invited by letter to consult their GP at a specified time to discuss screening (Nichols et al. (1986));
3) Individuals are invited by letter to make an appointment to discuss screening with their GP (Box et al. (1984), Nichols et al. (1986));
4) Individuals are invited by letter to discuss screening with a health visitor at a specified time (Nichols et al. (1986));
5) Individuals are invited by letter to collect a screening test from reception at their local GP surgery (Nichols et al. (1986)).

In modelling the invitation regimes it is assumed that:

- a register of the population to be screened is already available;
- completed tests can be developed at any health care facility;
- non-participants do not return unused tests for re-use; and
- general administration costs, development costs and overheads are constant for all options and can thus be excluded.

A accepting a test and K completing and returning a test. All data are from the studies cited with the exception of:

(i) the number of invitations for methods 2 to 5 (assumed to equal T), and
(ii) the number of acceptors for methods 2, 4 and 5 (assumed that 10% of those taking a test away with them do not complete and return it in line with the results of method 3).

The costs of each method of invitation can be expressed algebraically:

1. $T(c + 3s + 2p_2 + t)/K$
2. $\{T(c + 3s + 2p_1) + A(l + 10g + 2s + p_2 + t)\}/K$
3. $\{T(c + 3s + 2p_1) + A(l + c + 10g + 2s + p_2 + t)\}/K$
4. $\{T(c + 3s + 2p_1) + A(l + 15h + 2s + p_2 + t)\}/K$
5. $\{T(c + 3s + p_1) + A(l + 2c + 2s + p_2 + t)\}/K$

where the letters denote the following, with associated costs (£):

b	telephone call, per minute	0.014
c	clerical assistant time, per minute	0.05
g	general practitioner time, per minute	0.34 (Hughes (1991))
h	health visitor time, per minute	0.07
p_1	postage, letter only	0.15
p_2	postage, test	0.20
s	stationery, per item	0.01
t	3-day Haemoccult test	1.13

'l' denotes the cost to the individual of attending the GP and is calculated as follows. There are two elements to this cost: travel and time. A national survey of the accessibility of surgeries (Ritchie et al. (1981)) found that:

i) 49% of the population live less than one mile from their local surgery, most of this group walk to the consultation;
ii) 26% live 1-2 miles from their GP, half travel by car with the remainder equally split between walking and bus;
iii) 25% live more than 2 miles from the surgery, two-thirds travel by car with the reminder using the bus; and
iv) average waiting time is 15 minutes.

The average time taken to consult a GP is about one hour altogether. Time is valued at £1.75 per hour according to Department of Transport (1987) in 1985 prices, updated at 5% p.a. to £2.23 in 1990 prices. About half the population only incur 'shoe leather' costs but the remainder incur travel costs assumed to be £1 (at £0.3/mile). The average cost to the patient of a GP visit is thus £3. This is consistent with the finding that 90% of the population regard visiting the GP as 'easy' (Ritchie et al. (1981)).

The cost of each alternative, based on the given assumptions, is presented Table One. All calculations are normalised for a target population (T) of 1,000 persons:

Table One

Method	% accepting	% completing	TC (£)	AC (£)
1	100	46	1,610	3.50
2	59	53	4,953	9.34
3	30	27	2,720	10.07
4	50	45	3,080	6.84
5	19	17	846	4.98

Notes
TC - Total cost
AC - Average cost per acceptor

The results are also presented in Figure One, demonstrating the trade-off between cost and uptake; ideally an option would be located in the bottom left hand corner of the graph. The MRC trial method of postal invitation offers the lowest cost per person completing a test. The health visitor option has the lowest total cost but this places the greatest onus upon the individual (v_f is highest). While GPs can achieve higher participation (an extra 70 cases in option 2) the extra cost involved is £3,343, an incremental cost of £48 per extra participant.

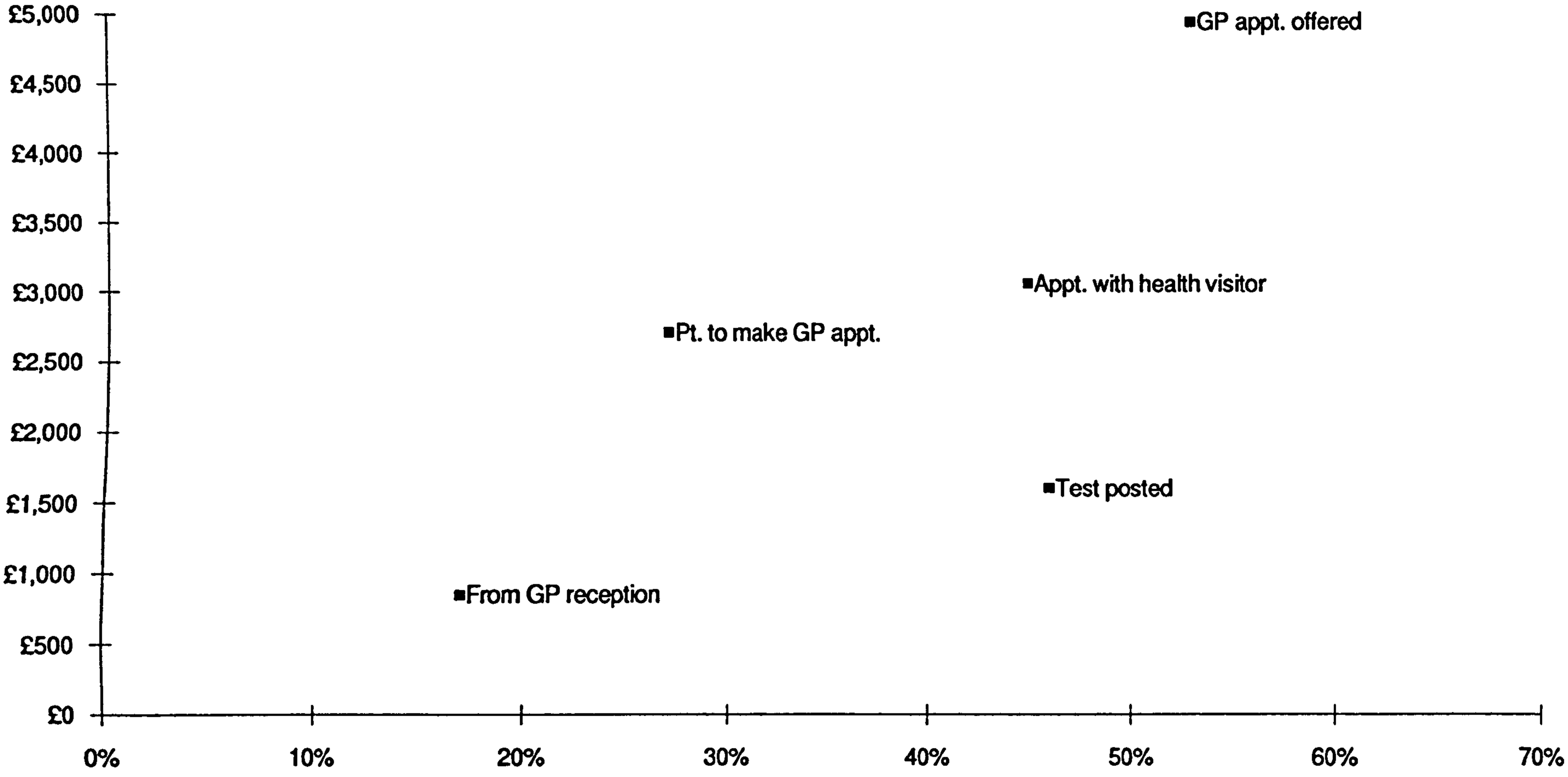
Enhancing participation in postal-based mass screening programmes
The European trials of colorectal cancer screening all use the postal invitation method (number 1 above), although each trial also uses a variety of means of persuading people who have not completed a test to do so.

- The MRC trial follows the initial invitation with a reminder letter to non-responders, pointing out the potential benefits of early detection and the ease of completion of the test. People who still do not respond have a 'flag' placed in their notes at their GP surgery, which will prompt the GP to discuss screening with them when they next attend.
- In the Swedish trial, a reminder letter is sent, with persistent refusers receiving a second reminder letter with another test.
- In the pilot study of the Danish trial a reminder letter was sent, with persistent refusers contacted by telephone or a second letter (Adamsen and Kronborg (1984)).
- The main Danish study uses a more complex protocol based on two reminder letters and a telephone call. At each mailing (including the initial mailing), a questionnaire is included, to be returned by those not wishing to complete a test. Those completing a questionnaire were telephoned to correct misperceptions and to attempt further persuasion (Klaaborg et al. (1986)).

These methods can be costed using a similar method to that in the previous section. The appropriate cost equations, based on similar lettering, are:

i) For the initial invitation: $(c + 3s + 2p_2 + t)$
ii) For each reminder letter: $(c + 3s + p_1)$

Figure One - Cost and response for different screening invites



iii) For the GP prompt card: mark notes (2c + s), assume GP discusses with 50% (2g). Each 'convert' sent test (c + 3s +$2p_2$ + t).
iv) For telephone contact: 2c per target; (5c + 5b) each contact, each acceptor sent test (c + 3s +$2p_2$ + t).

The results of costing the four alternatives are presented in the table, indicating for each trial, the average cost and participation rate and the subsequent incremental cost per incremental responder and the additional participation rate at each stage (Danish (1) and Danish (2) are the Danish pilot and main studies respectively):

Table Two

	MRC		Sweden		Danish (1)		Danish (2)	
	£	%	£	%	£	%	£	%
Invitation letter	3.53	45.7	3.15	51.1	2.64	61.0	2.77	58.5
Reminder letter 1	1.13	10.0	1.10	9.8	1.83	4.7	1.05	5.0
Reminder letter 2							1.20	1.9
Reminder 2 with test			13.69	4.6				
Phone refusers					19.29	1.3	6.12	3.3
GP prompt card	5.97	2.5						

Using the MRC protocol, 45.7% of those contacted respond to the initial invitation at an average cost of £3.53 per responder. A reminder letter is sent to the 54.3% who do not reply: another 10% of the *whole* sample reply at an average cost per extra responder of £1.13. A total of 44.3% of the sample have not yet replied and a prompt card is placed in their GP notes: this adds 2.5%age points to the total response at a cost of £5.97 per extra responder.

A high response to the initial invitation results in a low average cost but low initial participation is often followed by a good response to the first reminder. Thereafter, persuasion becomes progressively more expensive, because of the smaller number of further successful 'conversions'; this mirrors earlier findings in screening for asymptomatic bacteriuria (Rich et al. (1976)).

Optimal participation rates

Mass population screening programmes often aim for 70% participation (e.g. Forrest (1986)) as an 'optimal' participation rate. This section sets out to determine the economic optimum on the basis of the balance of the extra costs and benefits involved in increasing response.

In a given population of N people, n complete a screening test and the prevalence of disease is p. All screen detected cases are early stage and all symptomatic cancers are advanced stage. The benefit of treating early stage disease is B_1 and of late stage disease is B_2 with treatment costs of C_1 and C_2 respectively. The net benefit of the screening programme is $np(B_1 - B_2)$.

The cost of screening is nC_t (where C_t represents the cost of the test plus postage and development) plus the net costs of treatment:

$$nC_t + np(C_1 - C_2)$$

Screening is efficient if marginal benefits exceed marginal costs:

$$np(B_1 - B_2) > nC_t + np(C_1 - C_2)$$

However, the probability of participation might be enhanced by additional expenditure i.e. participation can be a direct function of cost, $n = f(C_t)$ or $C_t = g(n)$. Optimal participation from the economic point of view, where marginal benefits equal marginal costs, can be derived from the equation:

$$np(B_1 - B_2) = n[g(n)] + np(C_1 - C_2)$$

i.e. $g(n) = p[(B_1 - B_2) - (C_1 - C_2)]$

To make the model operational requires some proxy measure of the net value of screening in the absence of trial results. Based upon UK cancer registration figures (OPCS (1988)), the average age of new cases of colorectal cancer is 70 years for females and 67 years for males; life expectancy at these ages is 84 and 79 years respectively (OPCS (1987)). If early stage treatment has a 100% cure rate, patients have an average life expectancy of 13 years after treatment. By contrast, the average life expectancy of those treated for late-stage cancer is approximately 3 years, i.e. there is a net gain from early treatment of 10 life-years. Based on the most recent evidence from the MRC trial, $p = 0.0035$. The cost of cancer treatment is estimated to be £2,000 (Tuck et al. (1989)). All benefits and costs are discounted at 6% per annum. It is assumed that treatment for screen-detected cancer takes place immediately after screening and that for late-stage cancer occurs three years later.

The function $n = f(C_t)$ can be estimated for each of the four trials from the data presented in Table Two, assuming that the relationship is continuous. A logarithmic form

$$n = a + b.\ln(C_t)$$

was chosen assuming that (i) a minimum level of participation will occur without any encouragement, and (ii) there will be diminishing returns to persuasion. The equation coefficients and test statistics (in brackets) are presented in Table Three:

Table Three

	a	b	R^2 (%)	F
MRC	0.3695	0.0944	67	2.026
	(3.123)	(1.423)		
Sweden	0.4698	0.0667	71	2.462
	(5.587)	(1.569)		
Danish pilot	0.6032	0.0226	67	2.063
	(18.278)	(1.436)		
Danish main	0.5328	0.0674	88	15.165
	(18.369)	(3.894)		

Although these equations fail classical statistical significance criteria, they are satisfactory under the criteria for very small sample sizes estimated by Leamer (1978).

Reworking the optimal participation equation gives:

$$n = b.\ln\{p[(B_1 - B_2) - (C_1 - C_2)]\} + a$$

and this permits the calculation of the implicit value of a life-year gained for any given level of participation. These are approximately £448 (MRC), £464 (Denmark main), £734 (Sweden) and £871 (Denmark pilot) for the present levels of persuasion used. If the value of a life-year exceeds these amounts then it is worth incurring additional persuasion costs. A participation of 100% would be optimal only if the value of a life-year exceeded £34,780, £123,700 and £45,000 for MRC, Swedish, and Danish main trials respectively; the Denmark pilot study would require a valuation in excess of £1 billion for optimal participation to be 100%. Note that the technical reasons for non-participation noted earlier restrict the realistic maximum participation to less than 100%.

These values are proportional to cancer prevalence (p) i.e. life-year values will halve if prevalence doubles. Changing the assumption of life-years gained as a result of early detection, from 10 to 9 years raises the implicit life-year valuation by approximately 8% for all four protocols. Halving the number of life-years gained to five increases the values by 71-74%. The results are not sensitive to treatment costs: doubling the costs to £4,000 only raises the implicit value of a life-year by between 2 and 7%.

Discussion

The results assume that participation is independent of prevalence, thus the gain in yield for a given number of extra participants is constant. However, in the Danish trial the proportion of colorectal cancer deaths is significantly higher amongst refusers than among acceptors (Kronberg et al. (1987)). The

MRC trial accepts "...some degree of selection bias is apparent in the high frequency and poor prognosis of cancers diagnosed among non-responders" (ardcastleHardcastle et al. (1989)). Prevalence of colorectal cancer and participation in FOB screening are both age-related, with screening uptake being lowest where the incidence (and prevalence?) is highest, i.e. in the older age groups of the target population; this is considered in more detail in Chapter Eight. A similar effect has been noted for breast cancer screening: in the Edinburgh trial the proportion of breast cancer deaths was far higher among refusers than acceptors (34% versus 12%), while participation also fell with age (64% of ages 45-49, 57% of ages 60-64) (Roberts et al. (1990)).

This cannot be taken to mean that it would be those refusers most at risk who were amenable to persuasion of the types described above. This question can only be conclusively answered by long-term follow-up of the on-going randomised trials. In the case of breast cancer screening, the New York HIP trial found breast cancer yield and mortality rates were lower among 'reluctant' acceptors than among those who needed no persuasion (Fink and Shapiro (1990)). Table Four shows the number of breast cancers per 1,000 life-years in each group and the rate of breast cancer deaths per 10,000 life-years:

Table Four

	Case rate	**Death rate**
'Keen' acceptors	2.35	5.6
'Reluctant' acceptors	1.95	3.8
Refusers	1.89	4.9
Controls	1.95	5.0

(Note - the original article does not record whether these differences are statistically significant).

This would suggest that the women who were initially reluctant to accept but succumbed to persuasion were less at risk than the other groups; in terms of the research findings outlined in an earlier section they may have been those with a very high regard of their own health (and with some reason). The poor prognosis in the refuser group (less cases, more deaths) implies that this group may consist of women with a pessimistic view of their own health (again, with reason). However, a higher mortality rate from other causes was noted in the refuser group and (to a lesser extent) in the 'reluctant' acceptor group, suggesting that co-morbidity may reduce the incentive to take part in screening.

The list of technical reasons for failing to participate earlier can be corrected without any recourse to changing individual behaviour, although this is not to say that this means of increasing participation is free. Inaccurate address registers can be very significant as 3 million individuals in Britain change addresses annually. A study of FHSA-based breast screening in Inner

London found 16% of invitations were returned by the Post Office (McEwan et al. (1989)). In inner cities FHSA registers have incorrect address for up to 50% of cervical and breast screening targets (Bowling and Jacobsen (1989)). GP records, used as the source of recruitment for the MRC trial, are generally more accurate with 2-3% test returns via the postal system. Birth dates are also inaccurate in 5% of entries (Ross (1989)); this is of some relevance when the target population is defined on the basis of age.

Protocols for other forms of screening have used other means of persuasion involving alternative data sources and direct contact with screening targets. One study of cervical screening used the electoral register to cross-check the population coverage of cervical screening (Cook and Wald (1985)). Of 500 names drawn from the register, 309 were excluded on age or gender grounds, with the remaining 191 being cross-checked against cytology records; 86 were found to have undertaken a recent test. The remaining 105 were contacted by post (42 requiring further reminders), enclosing a clinic appointment. The 22 individuals who did not respond to this offer were visited at home by a district nurse - 11 people attended following the posted appointment and a further 2 following a home visit. Applying the cost data used earlier in this chapter and assuming (i) 6 electoral register names and 2 cytology names can be checked per minute, and (ii) home visits take 30 minutes each, the average cost for the final 13 responders equals approximately £8 (for an additional participation increase of 6.8%). When compared to the other methods of pursuing non-responders this result seems reasonable.

Conclusion

A model based on expected utility theory is consistent with the observed behaviour of populations offered screening and can help in the design of health education to accompany the initial contact. In terms of motivating people to complete a test a postal-based system of Haemoccult screening outperforms screening organised through general practitioners. This difference is likely to be accentuated if the trial results used related to well-motivated GPs. In terms of prompting people to complete a screening test, some persuasion appears to be efficient. Going to the extremes used by some trials, however, means that the marginal costs are high relative to any extra benefits. The economic optimum for participation is less than 100% as the costs of persuasion rise once a 'hard core' of people, apparently determined not to complete a test. is reached.

Chapter Seven

THE BENEFITS OF ADENOMA EXCISION

Introduction

Although it is commonly accepted that the initial malignant stage of many, if not most, colorectal carcinomas is within an adenoma the benefit of adenoma excision is difficult to quantify. While the evaluation to date has excluded any effects of adenoma excision, the sensitivity analysis of the simple model of FOB testing in Chapter Four showed that the cancer detection rate was the most significant variable. This implies that quantifying the effect of breaking the adenoma-cancer sequence is potentially crucial for the evaluation: cases *prevented* may prove as important as cases *detected*. No systematic analysis of the data on adenoma detection in the MRC trial is yet available, hence the first section of this chapter gathers together the relevant figures for the first time. The next section compares various ways of estimating the impact of adenoma excision on future cancer incidence. The final section considers the costs and benefits of adenoma excision.

Adenomas in FOB screening trials

At the time this work was carried out (mid-1991) 139 people in the control group of the MRC trial had had adenomas excised, as had 552 members of the group offered screening. The ratio between the sexes is identical in each group with 61% of the patients being male. The mean age in the study group was 65 with their control counterparts being one year older. More study group cases had multiple adenomas (29% versus 19% of cases in study and control groups respectively).

The commonest way to classify adenomas is according to the diameter of the largest recorded specimen for each person:

Table One

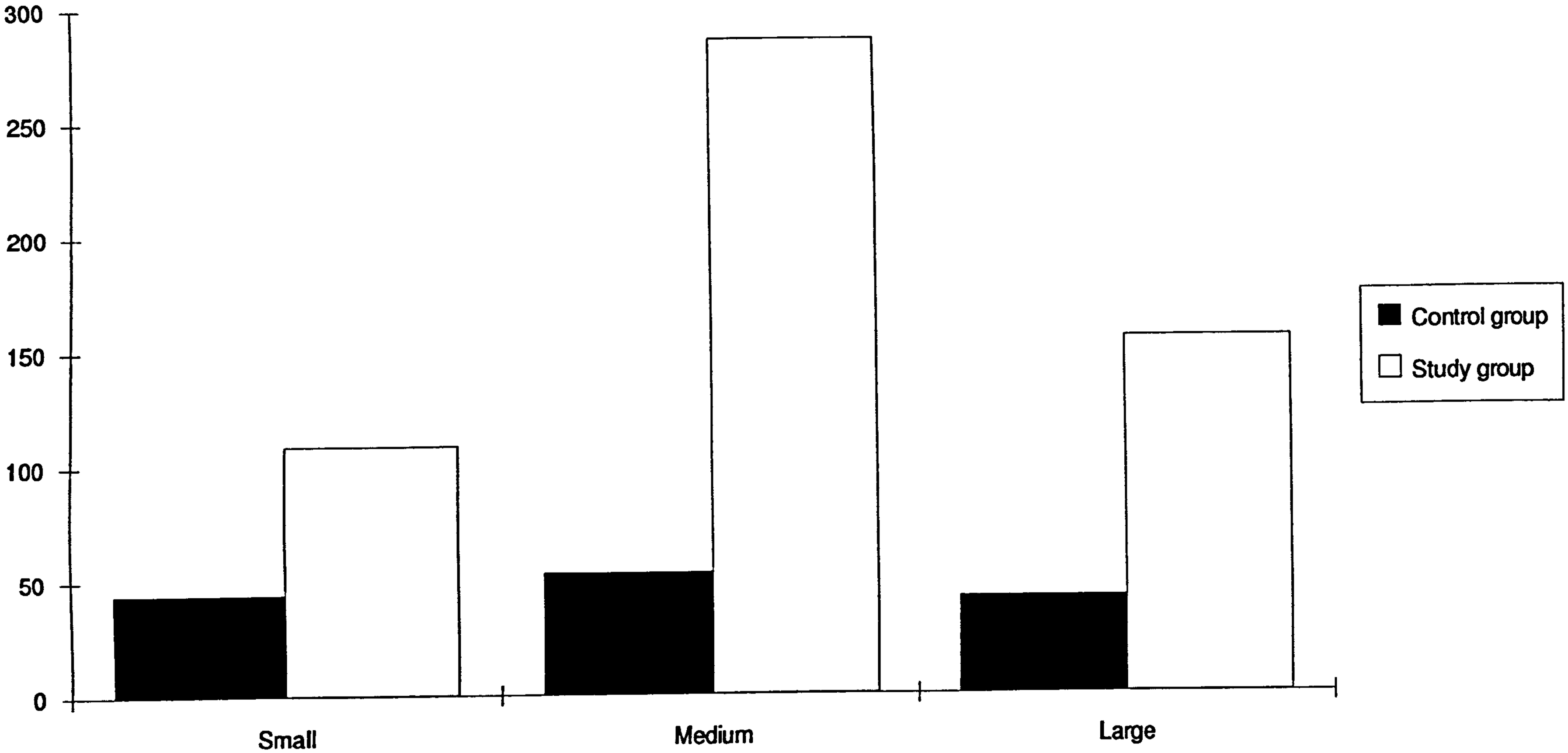
Group	Small	Medium	Large
Control	44	53	42
Study	109	287	156

'Large' - 2 cm. or more in diameter; 'medium' - 1 cm. to 2 cm. diameter; 'small' - less than 1 cm. diameter.

These data are presented in Figure One. The mean age of each of the six sub-groups was identical at 65 years old.

Five times as many adenomas were found in the study group as in the control group; the excess was particularly notable for medium and large adenomas. Screening detects 9.6 adenomas per 1,000 people completing a test at the initial screen and 3.8 per 1,000 at the second screen. Theory predicts a

Figure One - Adenomas by size and trial group



large yield from the early rounds of screening as the prevalent adenomas are cleared from the study group, with differences in rates of detection equalising over time to reflect incidence alone. Note, however, that post mortem series do not find a similar excess of medium and large cases (Williams (1982)). In the Danish trial the figures for adenomas detected at the initial screen are substantially lower although figures for the second screen are similar (Bech et al. (1991)).

Table Two shows the rate of adenoma detection (per 1,000 persons) in the control group (C) and for the study group (S) following recruitment to the trial.

Table Two

	All		Large		Medium		Small	
Year	**C**	**S**	**C**	**S**	**C**	**S**	**C**	**S**
1	0.3	2.1	0.1	1.1	0.1	1.6	0.1	0.4
2	0.3	1.2	0.1	0.4	0.1	0.6	0.1	0.1
3	0.4	1.4	0.1	0.2	0.1	0.8	0.2	0.4
4	0.5	1.0	0.1	0.1	0.2	0.5	0.2	0.3
5	0.5	0.8	0.1	0.1	0.2	0.6	0.2	0.1
6	0.4	0.8	0.1	0.2	0.2	0.3	0.1	0.3
7	0.4	0.5	0.1	0.2	0.2	0.3	0.1	0.3

The incidence in the control group rises slightly over time as the population ages, although part of the effect may be attributed to improved methods of diagnosing and excising polyps as the trial progresses. The yield in the study group falls steadily as prevalent cases are removed and by the seventh year of the programme the rate is almost equal to that of the control group: after three rounds of screening the yield is only slightly above the incidence rate. The fall in the rates of detection of medium and large adenomas is much more marked than the figures for small adenomas.

In terms of the site of the largest adenoma, screening seems to detect a large number of cases in the left side of the colon:

Table Three

	Control			Study		
	Small	**Medium**	**Large**	**Small**	**Medium**	**Large**
Right side	1	0	2	2	1	3
Transverse	2	1	1	9	4	7
Left side	16	26	18	50	202	110
Rectum	25	26	21	48	80	36

Right side - caecum and ascending colon; transverse includes both flexures; left side - descending and sigmoid colon; rectum includes the recto-sigmoid junction.

There is very little difference between the groups beyond the descending colon but there is a very heavy predominance of adenomas in the sigmoid colon. This is not borne out either by a British post mortem series (Williams et al. (1982)) or by the Danish screening trial where the distribution is more uniform.

In summary, the MRC screening trial has been very successful at detecting and excising adenomas. However, the data show some inconsistencies with other reports in terms of the number of adenomas detected at the initial round of screening, their site within the colon and their size.

Screening and the reduction of future cancer incidence

A number of methods have been used to estimate the benefits of adenoma detection and excision. As has been noted in Chapter Three the results of several economic evaluations of colorectal cancer screening depend crucially on these estimates (Eddy (1990), Wagner et al. (1991)). In this section six means of calculating the significance of the adenoma-cancer sequence are explored in more detail. Each method is described below and then applied to a standardised sample of adenomas.

Method One: Eddy and expert opinion

In Eddy's model 93% of cancers begin as adenomas and hence are potentially preventable by adenoma excision (Eddy (1990)). 5% of the adenomas which reach a diameter of 5mm are destined to become cancers after an average duration of seven years. This is based on reviews of 'expert opinion', although the lack of scientific data must raise questions as to what this 'opinion' is based upon.

Method Two: Hoff's growth rates

A Norwegian study achieved ethical approval to leave 35 newly diagnosed adenomas of less than 5mm diameter unexcised for two years (Hoff et al. (1986)). At the follow-up investigation 17 had grown, 13 were the same size and 5 had shrunk. Growth was most common in the sigmoid colon. On average the adenomas that grew were 2.8mm. in diameter and originally and 4.1mm .subsequently, while those that shrunk went from 3.6 to 2.4mm. on average. From these figures it was estimated that a small adenoma takes 10 years to double in size. Assuming that the sequence takes between 10 and 15 years from the small adenoma to cancer the annual risk of cancer is 1% in small adenomas, 20% in medium-sized adenomas and 50% in large adenomas (Hoff (1987)).

Method Three: Eide's prevalence calculations

On the basis of a post mortem series Eide, a Norwegian pathologist, estimated age-specific prevalence rates and hence the total number of prevalent adenomas in the Norwegian population (Eide (1986)). These figures assumed similar distributions of characteristics (size, histology, dysplasia) in the general population and in the post mortem series. Based on

cancer incidence data he then calculated the annual probability that an adenoma with one of the above characteristics would become malignant and be detected. Overall, the annual conversion probability was 0.25%, but for adenomas larger than 1cm. in diameter the probability was 3% per annum, for a villous element to the histology the annual probability was 17%, and for severely dysplastic adenomas the annual risk was 37%.

Method Four: Stryker's observational study

American researchers reviewed the records of 226 patients with medium or large colonic adenomas where diagnosis was not followed by excision either because of the patient's general medical condition or their preferences (Stryker et al. (1987)). Regular BEXR surveillance was carried out and the adenoma excised only when 'significant' growth was observed. In total, 47% of cases eventually underwent resection, with 21 cancers found at an average follow-up of 9 years (range 2-19 years). Cumulative malignancy probabilities over time were also calculated: the chance of cancer developing at the adenoma site was 2.5% after 5 years, rising to 8% after 10 years and 24% after 20 years.

As the authors admit the study has a number of weaknesses: for example, it is impossible to be sure that malignancy was not present in the first place. However, if a curve is fitted to the data, the equation gives the cumulative probability of carcinoma arising in an adenoma (P) on the basis of the number of years since diagnosis (Y). The equation passes through the origin (assuming malignancy was not present at initial diagnosis) and takes the following quadratic form:

$$P = 0.003505\,Y + 0.000426\,Y^2$$

Differentiating gives the probability that malignancy develops during any given year following detection.

Model Five: Stryker adjusted for small adenomas

As the above data apply only to medium or large adenomas there is no possibility of small adenomas growing to pose a threat of malignancy. To allow for this the growth rates described by Hoff above can be used: these would imply that small adenomas become medium-sized after ten years; after this the equation above applies.

Method Six: Stryker adjusted for histology and dysplasia

Method Four does not take any account of the malignant potential of adenomas in terms of histology (tubular, tubulovillous, villous) or degree of dysplasia (mild, moderate, severe). Method Three, in line with conventional wisdom, implies that villous and severely dysplastic adenomas have the greatest malignant potential. How can the Stryker equation be adjusted to reflect this?

There is no longitudinal study upon which to base these weights, yet leaving the equation unadjusted implies no extra weight for these factors. One possible data source is the report of a cross-sectional endoscopic study of over 1,000 excised adenomas including the proportion showing malignancy according to each characteristic (Muto et al. (1975)). Taking the Stryker equation as representing the malignant potential of the most common type of adenoma (medium-sized, tubular) weights can be derived from the ratio of the malignancy rate for any pair of characteristics compared to the rate for the medium tubular. For example, malignancy was found in 4.5 times as many large, tubulovillous adenomas as in the base case (medium tubular): thus, the Stryker equation for large tubulovillous equations has coefficients 4.5 times those of those of the base case. The weights used are presented in Table Four:

Table Four

	Size			**Dysplasia**		
	Small	Medium	Large	Mild	Moderate	Severe
Histology						
Tubular	0.1	1.0	3.4	0.2	0.9	2.7
Tubulovillous	0.4	0.7	4.5	1.4	3.1	3.3
Villous	0.9	1.0	5.3	3.5	4.0	4.9
Dysplasia						
Mild	0.03	0.3	4.1			
Moderate	0.2	1.4	4.9			
Severe	2.6	2.4	4.7			

Unfortunately, the cross-section study only reported malignancy rates for pairs of characteristics. A given adenoma, therefore, falls into three different categories; for instance, a small, severely dysplastic, villous adenoma has a weight of either 4.9 (villous, severe dysplasia), 2.6 (villous, small) or 0.9 (small, severe dysplasia). It is not clear which characteristics dominate or how they interact, hence the analysis uses the middle estimate for each case.

Data sample

In order to compare these methods the characteristics of the adenomas detected in each group of the MRC trial were compared. The aim was to specify the characteristics of the adenomas which were in excess in the study group i.e. those which were detected by screening but which would not have presented symptomatically. The results are presented in Table Five

Table Five

Size	**Histology**	**Dysplasia**	**Control**	**Study**
S	T	Mild	4	24
S	T	Moderate	6	25
S	T	Severe	3	3
S	TV	Mild	1	2
S	TV	Moderate	1	3

S	TV	Severe	0	0
S	V	Mild	0	0
S	V	Moderate	0	1
S	V	Severe	0	0
M	T	Mild	8	32
M	T	Moderate	5	77
M	T	Severe	2	15
M	TV	Mild	2	11
M	TV	Moderate	6	48
M	TV	Severe	4	14
M	V	Mild	0	0
M	V	Moderate	0	2
M	V	Severe	0	2
L	T	Mild	1	10
L	T	Moderate	3	15
L	T	Severe	4	4
L	TV	Mild	3	6
L	TV	Moderate	4	38
L	TV	Severe	5	17
L	V	Mild	1	4
L	V	Moderate	2	6
L	V	Severe	1	5

Note the table only records cases with complete data for all three characteristics, excluding 34% of study group cases and 53% of control group cases.

For example, Table Five shows an 'excess' of 34 large moderately dysplastic tubulovillous adenomas in the study group over the control group. Rather than working with the exact number of excess cases for each category, the proportion of the excess falling into each category was used to define the characteristics of a group of 100 'typical' excess cases. The composition of this group was as follows:

- 24 medium tubular moderate dysplasia
- 14 medium tubulovillous moderate dysplasia
- 12 large tubulovillous moderate dysplasia
- 8 medium tubular mild dysplasia
- 7 small tubular mild dysplasia
- 6 small tubular moderate dysplasia
- 5 medium tubular severe dysplasia
- 4 large tubular moderate dysplasia
- 4 large tubular severe dysplasia
- 3 medium tubulovillous mild dysplasia
- 3 medium tubulovillous severe dysplasia
- 3 large tubular mild dysplasia
- 1 small tubulovillous moderate dysplasia
- 1 medium villous moderate dysplasia

- 1 medium villous severe dysplasia
- 1 large tubulovillous mild dysplasia
- 1 large villous mild dysplasia
- 1 large villous moderate dysplasia
- 1 large villous severe dysplasia

As the average age of the sample was 65 years, their remaining life-expectancy was 15 years (OPCS (1987)).

Each of the six methods of estimating the adenoma-cancer sequence was applied to this population in turn, taking account of the appropriate restrictions e.g. small adenomas were excluded from the original Stryker model (method 4).

Results

Table Six shows the predictions of the six methods of the number of cancers that would have developed in the 100 adenomas if they had not been excised in Year 0 (Stryker (2) refers to the modification to take account of small adenomas and Stryker (3) refers to the weighted Stryker equation):

Table Six

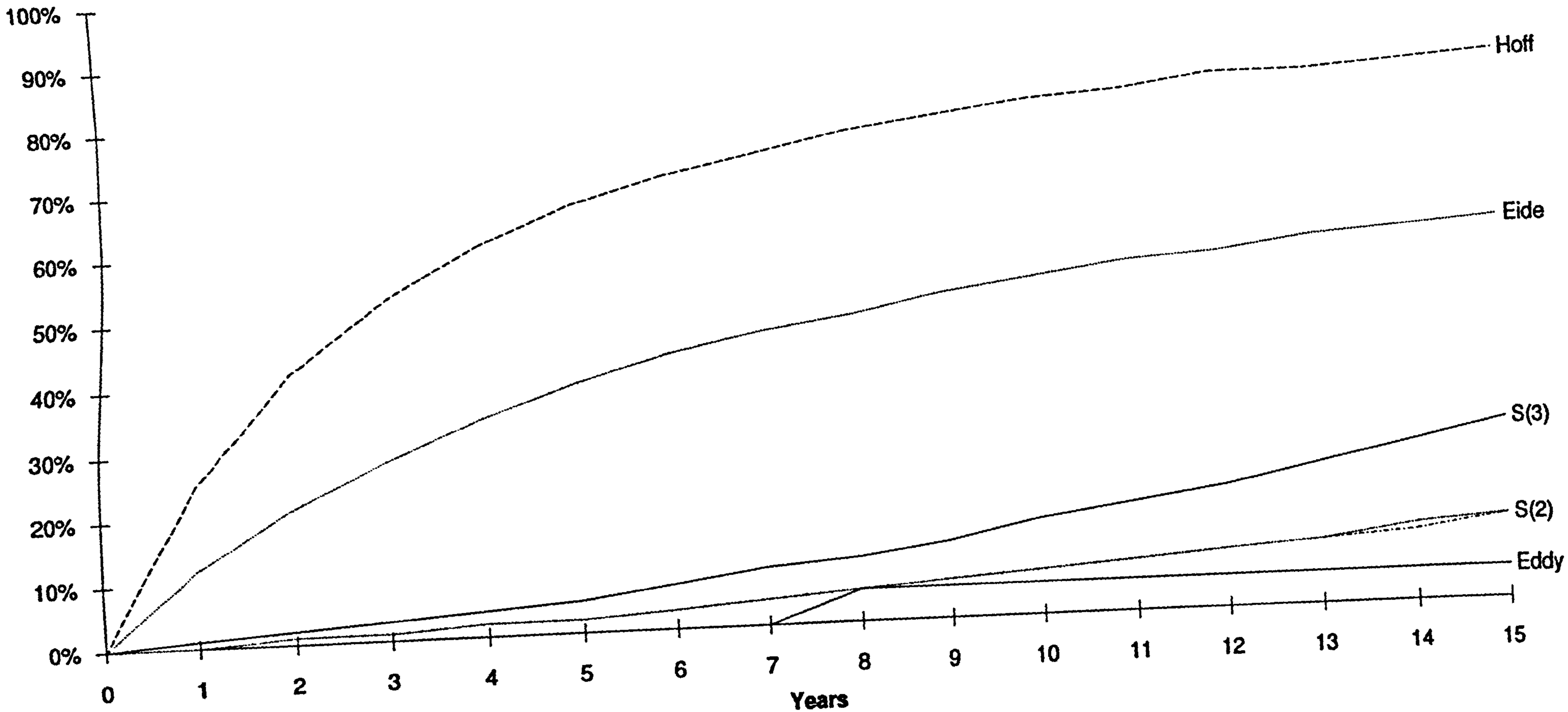
Years	Eddy	Hoff	Eide	Stryker	Stryker (2)	Stryker (3)
1	0	25	12	0	0	1
2	0	42	21	1	1	2
3	0	53	28	1	1	3
4	0	61	34	2	2	4
5	0	67	39	2	2	5
6	0	71	43	3	3	7
7	0	74	46	4	4	9
8	5	77	48	5	5	10
9	5	79	51	6	6	12
10	5	81	53	7	7	15
11	5	82	55	8	8	17
12	5	84	56	9	9	19
13	5	84	58	10	10	22
14	5	85	59	11	12	25
15	5	86	60	13	13	28

The results are also shown in Figure Two. The range of results is extremely wide: unless there are grounds for believing that one method is far superior to the others (implying more weight should be attached to one set of results), the magnitude of the benefit will be very uncertain.

Discussion

Ultimately, these predictions can be compared with observed declines in the incidence rate following screening: for example, Eddy claims that he constantly revises his assumptions in the light of new trial data. However, this requires many years of follow-up and in the interim models will be judged

Figure Two - Adenoma-cancer predictions of different methods



on the perceived validity of the assumptions used. A randomised trial of adenoma removal has been proposed (Pollock and Quirke (1991)), leaving a group of diagnosed adenomas unexcised, but the ethical problems are severe.

How can the differences in results be explained? One possibility is that colorectal cancer is a more aggressive disease in Norway than in America, hence a higher proportion of Norwegian adenomas eventually become malignant. A second explanation is that there is a difference in definitions between the research studies; the Norwegians calculated the proportion in which cancer was clinically present, while Stryker was calculating the proportion in which symptoms occurred. If this were the case then the Norwegian surplus cases would be 'lifetime latent' cancers.

Another explanation is that Eide assumes all cancers begin as adenomas: to the extent that some arise de novo in the bowel, his figures are over-estimates. If 40% of cancers arise de novo, then Eide's method would predict 51 cancers rather than 60 as in Table Six.

While each explanation holds some truth it is hard to escape the conclusion that all of the researchers are making (more or less) intelligent guesses. Adenoma excision almost certainly has a health benefit but we simply do not know how large this is: as few as 5% of adenomas may have become malignant or as many as 85%. But is this all that can be said? Leaving this factor out of the evaluation altogether means that the benefits are underestimated and the programme may not be implemented as a result. In particular, screening the younger age groups may be hard to justify if the only source of health benefit from screening is via cancer detection. Which estimate is more realistic? The Norwegian figures seem to bear less resemblance to the available facts since they imply a large reduction in cancer incidence in the year following screening, an effect that has not been observed in any of the clinical trials. Eddy's method is also discarded since, despite its simplicity, it is the average of many subjective opinions; while this is not inherently bad, the guesses are not based on any evidence. The Stryker equation is unsatisfactory but is at least based on an observational study.

Which of the Stryker variants is most appropriate? The adjustment for small adenomas makes virtually no difference to the results: this is not surprising since there are only 14 cases in the sample and they can have no effect whatsoever until five years before the end of the patients' lives. The weighted Stryker equation has an intuitive appeal, but is not based on observed experience. This leaves the simple Stryker model alone, although as noted above, the figures are subject to several sources of bias. For example, they are an overestimate to the extent that any of the cancers eventually detected were present at the first diagnosis of an adenoma. They are also calculated on the basis of the time taken for them to double in size

rather than to cause symptoms and thus affect health: the high proportion of excised cancers at Dukes' stage A in their study illustrates this. The figures are upper limits on the true benefits, therefore.

Possible extensions to the analysis include applying the Stryker equation to each patient according to their individual life-expectancy: a preliminary estimate implies that this does not make much difference to the results. Another possibility is to take account of the number of adenomas found in each patient: those with multiple adenomas are thought to be at higher risk of developing cancer. While there are good grounds for thinking that this is a relevant factor, the problem is once again of a lack of data to use. Resorting to cross-sectional studies would imply weights of between 1.3 (O'Brien et al. (1990)) and 2.1 (Muto et al. (1975)). There is the danger, of course, that these would produce yet more estimates which could not be verified.

All of the discussion to date has assumed that preventing cases of cancer in the future will carry with it health gain. Some have suggested, however, that adenomas are merely the benign manifestation of the factors which cause colorectal cancer. This school of thought would argue that an adenoma detected on screening has remained asymptomatic for so long that there is no reason to believe it would ever have affected the patient's health. Even an enthusiastic advocate of endoscopic screening for adenomas has written, "One must certainly consider the possibility that prophylactic polypectomy might lead only to a reduction in the incidence of the more benign-behaving tumours." (Jass (1989)).

While the health benefits of adenoma excision are uncertain there are also cost implications in terms of the cost of conducting an endoscopic polypectomy and the benefits in terms of treatment cost savings. Can adenoma excision be justified on these grounds alone, leaving any health benefit as a 'bonus'?

Economic evaluation of adenoma excision

The distribution of treatment cost savings over time is calculated by multiplying Stryker's prediction of cancer incidence by unit treatment costs, appropriately discounted into present values. An earlier study (Tuck et al. (1989)) calculated the mean length of inpatient stay for the treatment of colorectal cancer patients to be 17 days. Using an average cost per inpatient day of £140, the treatment cost per cancer amounts to £2,380. Future costs are discounted at 6% per annum, in line with Treasury recommendations for public sector projects.

The costs of excision are the additional resources of the procedure itself, the need for any follow-up investigations and the associated risks to the patient's health status. In the MRC trial, adenomas are excised endoscopically during the course of a diagnostic colonoscopy with a small marginal cost: excision requires few additional resources. The proportion of cases where a follow-up

investigation is undertaken either to complete the excision or to review the completeness of the procedure is high, however: in the MRC trial 30% of patients require BEXR and 6% need a second colonoscopy to complete the investigation.

With respect to health risks, a study of complications in 3,153 colonoscopic excisions (Macrae et al. (1983)) found further treatment was required in a total of eight cases (0.25%); six patients required minor operations to control haemorrhages and two to remedy perforations. No deaths from excision *per se* (as opposed to investigation) were reported.

Thus for 100 patients diagnosed as having adenomas at colonoscopy the cost of polypectomy is:

30 patients requiring BEXR to complete @ £78.87 each
6 patients requiring a second colonoscopy @ £105.10 each
0.44 patients requiring an operation for complications @ £2,380 each

The total cost to the NHS is thus £4,044.

The discounted cost saving is £17,851, implying a net saving on each adenoma of £138. Assuming that the Stryker equation is an approximation to the number of cancers prevented, excision can be justified on cost-saving grounds alone with any health benefit being an additional gain. People who have an adenoma excised may subsequently develop cancer, which would serve to reduce any savings.

CONCLUSION

A review of the literature has found several different ways of calculating the effect of adenoma excision on future cancer incidence. Applying these to a standard population derived from previously unpublished MRC trial results shows that each gives a different answer. The range involved means that there is no consensus. Some methods can be ruled out because they do not correspond to observed behaviour, although it is harder to choose between those in the plausible range. It is concluded that about 13% of excised adenomas would have become cancers within the lifetime of the sample, but only further observation of the MRC trial population will confirm this. Polypectomy is justified in terms of the cost savings it produces alone, irrespective of any health benefit from cancer prevention.

Chapter Eight

AGE AS A RISK FACTOR FOR TARGETING SCREENING

Introduction

Most types of screening are targeted at a particular group: the MRC trial protocol uses age as its discriminator on the basis of the incidence figures presented in Chapter Two. The age range was not arrived at on the basis of a rigorous comparison of the costs and benefits of alternatives, however. This chapter assesses the effects of screening *either* only a sub-set of the 50-74-year-old group *or* of extending screening to younger and/or to older groups.

Epidemiological versus economic considerations

The most common age limits for colorectal cancer screening trials are from 45 or 50 up to 75. These are based on considerations of incidence to determine the lower age limit and the ability to undergo treatment to determine the upper age limit. (See for example American Cancer Society recommendations, European research trial protocols).

Economic evaluation has been used before to compare different target groups for screening. Opportunistic cholesterol testing, for example, costs £200 per quality-adjusted life-year (QALY) for men aged 40 to 69 (with diet therapy) but for men aged 25 to 39, the cost is £14,150 per QALY (Standing Medical Advisory Committee (1990)). Similar analyses of breast cancer screening for women aged less than 50 have provided an economic input into that controversy (Eddy et al. (1988)). For colorectal cancer screening strategies the only study to date indicates that commencing screening at 40 rather than 50 has little benefit but doubles the cost of a lifetime screening programme (Eddy (1990)).

Data for the evaluation

The MRC trial results were divided according to the patients' age and grouped into five-year age bands. (Rescreening results use the age of the patient at the rescreen). Projected screening costs, diagnostic costs and the yield of cancers and adenomas were made for each age group using the model derived in Chapter Four. This allows a comparison of results to be made on the basis of a common size of target population offered screening.

Table One - Initial offer of screening

			Detection rate (per 1,000)	
	Participation (%)	Positive (%)	Cancer	Adenoma
TOTAL	52.8	1.51	2.217	3.364
45-49	34.9	1.58	0	2.764
50-54	54.2	0.81	1.673	0.907

55-59	56.1	1.17	0.627	2.932
60-64	55.3	1.33	1.768	2.321
65-69	53.4	1.74	2.741	5.187
70-74	48.2	2.7	6.309	6.419

Table Two - Second round of screening

			Detection rate (per 1,000)	
	Participation (%)	Positive (%)	Cancer	Adenoma
TOTAL	76.3	1.23	1.735	2.167
45-49	78.4	3.09	1.931	1.531
50-54	73.7	0.76	0	2.029
55-59	75.4	0.87	0.713	0.94
60-64	77.9	1.27	1.563	2.842
65-69	78.2	1.17	2.423	2.625
70-74	75.5	1.4	2.8	2.75
75-79	73.1	2.35	4.334	1.812

Table Three - Third round of screening

			Detection rate (per 1,000)	
	Participation (%)	Positive (%)	Cancer	Adenoma
TOTAL	87	0.79	0.621	1.755
50-54	83.5	0.38	0	0
55-59	88.2	0.63	0	1.548
60-64	89.6	0.47	0.334	1.797
65-69	87.8	0.81	0.7	3.074
70-74	86	0.98	0.978	1.26
75-79	81.6	1.77	2.214	1.204

To calculate the significance of adenoma excision the Stryker model described in Chapter Seven was used, based on the life-expectancy for the midpoint of each age group (OPCS (1987)). The cumulative probability that an adenoma of 1 cm. diameter or more would eventually have progressed to cancer is as follows (LE denotes life expectancy):

Table Four

Age at detection	Life Expectancy	Cumulative chance (%)
45-49	30	47
50-54	25	36
55-59	21	26
60-64	17	19
65-69	14	13
70-74	11	9
75-79	8	6

The combined total of cancers detected and prevented was then used as the output of the programme to calculate a cost-effectiveness ratio for each group. Cases detected in the future were not discounted to a present value,

in line with Department of Health recommendations* (Parsonage and Neuburger (1992)).

Results

The results for each round of screening broken down by age group are as follows (AC1 denotes the cost per case detected alone and AC2 the cost per case either detected or prevented):

Table Five - Initial screen

	Cases Detected	Cases Prevented	AC1 (£)	AC2 (£)
TOTAL	117	23	3,271	2,677
45-49	0	45	-	7,542
50-54	91	18	3,607	3,018
55-59	35	43	10,212	4,609
60-64	98	24	3,792	3,035
65-69	146	25	2,727	2,330
70-74	304	19	1,491	1,405

Table Six - Second round of screening

	Cases Detected	Cases Prevented	AC1 (£)	AC2 (£)
TOTAL	132	21	2,190	1,924
45-49	151	56	3,978	2,898
50-54	0	54	-	6,085
55-59	54	18	6,359	4,736
60-64	122	42	3,213	2,388
65-69	189	27	2,007	1,759
70-74	211	19	1,896	1,742
75-79	317	8	1,572	1,533

Table Seven - Third round of screening

	Cases Detected	Cases Prevented	AC1 (£)	AC2 (£)
TOTAL	139	19	2,190	1,924
50-54	0	0	-	-
55-59	0	35	-	9,245
60-64	30	31	10,308	5,097
65-69	61	35	5,712	3,636
70-74	84	10	4,402	3,945

* This point is contentious. The article referenced points to illogicalities in applying the Treasury discount rate for public sector projects to health benefits: the arguments in favour of a 0% discont rate are weaker and have not been widely accepted by health economists. Indeed, using a 0% discount rate limits the comparability of results with other well-designed evaluations which use rates of 5 or 6%. The use of the 0% figure is on the assumption that the Department of Health will seek to enforce its views in the future, both in evaluations carried out by its own economists and also via the award of research funding.

75-79	181	6	2,538	2,457

Note that the rescreening figures for AC1 do not correspond exactly with those in Chapter Four owing to the different sample size bases used in each calculation.

Discussion

Despite the large sample size of the MRC trial, the numbers in each sub-group can be quite small and hence the figures are subject to the wide confidence intervals. For instance, 45-49 year olds were only recruited in the early stages of the trial, with initial screening offered to 1,447 but only to 661 at first rescreen and only 13 at second rescreen. In most other groups the sample size is around 8,000 but by the second rescreen the cancer yield is extremely low (0.6 per 1,000 acceptors) and the problem re-emerges. Thus, the figures should be taken to be indicative rather than a precise calculation.

Substantial differences between age groups exist with regard to the rate of participation, likelihood of a positive test and yield of neoplasia. Participation tends to be highest in the middle of the population age range i.e. around 60 years of age. The positive rate rises with age, as does the yield of cancer (whether detected or prevented). The exception is the cancer yield in the 55-9 age group on the initial screen, but this may be a chance finding. In other respects the trends in the results are in agreement with those reported by the Danish trial (Kronborg et al. (1987) and (1989)).

The cost per case either detected or prevented ranges from a low of £1,405 to a high of £9,245; in one sub-group no neoplasia at all have been detected to date. This suggests that a screening protocol which ignores age will have higher costs and lower benefits than one that adopts different policies for different sub-groups.

By virtue of including differential probabilities for the chances that an excised adenoma would have become malignant during the lifetime of the age group, the analysis has moved one step beyond that of Chapter Four. However, cost per case is not a basis for prioritising screening intensity because it takes no account of the remaining life expectancy (and hence of the ability to benefit) of each group. A full evaluation of must await the results of the on-going trial, but what are the chances of each sub-group achieving the necessary survival gain for screening to be justified?

Examination of existing cost-effectiveness league tables implies that treatments with a net cost per life-year saved of up to £2,000 are usually judged efficient. Thus, if it costs £6,000 to detect (or prevent) a cancer then it follows that it is necessary for the average benefit per case to exceed 3 life-years. On this basis, screening of all groups can be justified if early detection results in a gain of 4 life-years on average. For comparison, Eddy has

estimated that people with colorectal cancer die, on average, 6-7 years ahead of full life-expectancy.

Several studies have been carried out to test whether age is an independent prognostic factor for colorectal cancer. Those under 60 years of age have the best prognosis after controlling for all other differences between age groups (Svendsen et al. (1989)).

In terms of identifying the most efficient age group to screen, the effects of cancer prevention by adenoma excision are extremely important. On the basis of cancers detected alone it is hard to justify screening those younger than 55 on economic grounds; in particular, no cancers have been detected in those aged under 50 in either the MRC or the Danish trials. Given the life expectancy of this group, however, there is a high probability that excised adenomas of 1 cm. or more in diameter would have become malignant within the person's life-time. The results above imply that the under 50s should be considered for screening before anybody aged 50-64. The importance of the prevention effect lessens with the age at which the adenoma is excised, and is countered by the increasing yield of asymptomatic cancers detected with age. As a result, screening the two extremes of the age distribution gives the lowest average costs. Unfortunately, these groups are also the least likely to participate.

Possible extensions

The MRC trial protocol could be varied in three different ways to take account of the results above: (i) using a different test for different age groups; (ii) screening age groups at different intervals; and (iii) screening a different age range. These points are considered in turn.

A different test for different age groups?

The sensitivity of three-day Haemoccult testing for asymptomatic cancer is 67% (Thomas et al. (1992)). This is an average across all patients: sensitivity may be age-related which would open the possibility of using a different test for different groups. Unfortunately, no data are available to explore this further, but there may be justification for using a very sensitive test where ability to benefit is greatest i.e. in the youngest age groups. Flexible sigmoidoscopy has been proposed in this role, although some of the assumptions of advocates have already been challenged in Chapter Five (Atkin et al. (1993)). The option is considered in more detail in Chapter Eleven.

Different intervals for different age groups?

The second option is to rescreen different age groups at different frequencies. In terms of cost and yield, it is easy to justify offering a second screen to those aged 60 and over and even offering a third screen to the oldest age group; it is harder to justify offering repeated screening to those in the middle of the age range. (Numbers are insufficient to comment on the

youngest age group). A potentially efficient protocol would offer screening once or twice to those aged 45-50 with more frequent screening beginning at 60 or 65. The rising average costs for all ages in the third screening round indicates that it may be necessary to extend the interval between screening rounds at some point after the second round.

Are age limits of 50 and 74 appropriate?

The third possibility is to extend the target age to include those aged less than 50 and those aged over 75. In the former group, medium or large adenomas stand a very high chance of eventually becoming malignant (62% according to the Stryker equation). However, data from post mortem studies contain too few cases in this age group to allow further analysis. Account must also be taken of the likely participation rate in this group.

Over 75s have too short a life expectancy to benefit much from adenoma excision but cancer incidence increases dramatically in these age groups; on the other hand, the chance of a patient of this age surviving a colonic resection diminishes. In one surgical series 29% of over 70s died in hospital (Umpleby et al. (1984)); a second study found a mortality rate of 6% for the 70-79 year olds but 20% for those aged 80 or over (Lewis and Khoury (1988), Ozoux et al. (1990)). Patients detected as a result of screening would be generally healthier and would not be emergency admissions, however, implying that a lower figure is more appropriate. Screening this older age group may also encounter problems of poor participation; this effect has been shown to be smaller than expected in breast cancer screening, however (Hobbs et al. (1990)). If this can be taken to be indicative of a constant level of health consciousness in the older age groups then the results may apply to colorectal cancer screening as well.

If choices must be made between age groups, surveys of public opinion favour treating younger people ahead of older people (Wright (1986); Charny et al. (1989)). This would suggest wider approval of a policy which sought to extend screening to younger rather than to older age groups.

CONCLUSION

Including the effects of adenoma excision implies that screening younger age groups offers a low cost per case prevented or detected. Screening the older age groups has a similarly low cost per case owing to the high prevalence of cases in this group. The figures are consistent with a policy of one-off screening of the population in their 50s using a very sensitive test with more frequent screening in later life. Further evidence on the adenoma yield in people aged 45-55 would be helpful in determining the precise age at which this early screen would be most efficient.

Chapter Nine

THE INVESTIGATION OF THE POSITIVE SCREENING TEST

Introduction

An accurate diagnostic test is an important element of any screening programme: inaccuracies result in either cases being missed or unnecessary treatment. The problem in colorectal cancer screening is that there are several diagnostic options, but no definitive comparison by clinical trial to compare them. The most widely used tests are colonoscopy, flexible sigmoidoscopy and the double-contrast barium enema X-ray (BEXR). Endoscopic techniques are capable of biopsying suspicious areas of the bowel and hence are more specific than radiology. However, colonoscopy is more likely to result in complications such as haemorrhaging or perforation of the bowel wall, while sigmoidoscopy covers only a limited area of the bowel. Patients undergoing colonoscopy often require sedation, so some patients will be unsuitable for this type of investigation except when the operator is an expert endoscopist. Radiological investigations are safer but require a subsequent endoscopy to biopsy any findings.

A review of the clinical literature shows little consensus: some clinicians advocate colonoscopy (Aldridge and Sim (1986); Lindsay et al. (1988); Maxfield (1984); Thorson et al. (1986)), while others favour radiology (Feczko and Halpert (1986); Ott et al. (1985)) or even combinations of methods (Abrams (1982); Irvine et al. (1988); Kalra and Hamlyn (1988); Reiertsen et al. (1988)). The costs of the test is one factor but clinical recommendations are not based on the results of a full economic evaluation.

A model for evaluating diagnostic strategies

Algebraically the expected cost of detecting a cancer can be expressed as C/NS , where C is the total cost of using the diagnostic procedure, S is the sensitivity of the investigation, and N is the number of cases of disease present. Taking the cost per case detected to be the decision criterion, then for two tests, A and B, A will be cost-effective relative to B if

$$C_A/NS_A < C_B/NS_B$$

or, by rearranging, if:

$$C_A/C_B < S_A/S_B$$

for an identical case-mix. The yield depends on the prevalence of disease in the group with a positive screening result and the sensitivity of the test. The cost calculation is slightly more complicated and includes the following elements:

i) the cost of the preferred diagnostic procedure multiplied by the number to be investigated;
ii) the number referred for the alternative investigation, either because the initial investigation is incomplete, equivocal, or (in the case of a referral from radiology to endoscopy) requiring biopsy/excision of a lesion detected - this is referred to below as the 'cross-referral' rate;
iii) the proportion who experience a complication serious enough to require active in-patient care (as opposed to observation which is already included in the costs calculated in Chapter Four) multiplied by the cost of such care;
iv) the cost of a false negative result in terms of the additional treatment cost multiplied by the number of cases missed.

Initially, the model is applied to the investigation of symptomatic patients since most of the available data apply to this group; diagnostic investigation of positive screening results is considered subsequently.

Investigating symptomatic patients

Prevalence

Investigation of people with lower gastrointestinal symptoms reveals cancer in 8% of cases (Irvine et al. (1988), Lindsay et al. (1988)), although in the absence of a definitive diagnostic test it is difficult to say how many cases are missed; these comments also apply to sensitivity estimates.

Sensitivity

Estimates of the sensitivity of colonoscopy and BEXR to cancer vary. Combining the results of several studies (Aldridge and Sim (1986); Durdey et al. (1987); Irvine et al. (1988), Lindsay et al. (1988)) implies that the yield of cancers on colonoscopy is 40% higher than the yield on BEXR in similar populations. As the sensitivity of colonoscopy is commonly thought to be in the region of 95% this implies that the sensitivity of BEXR is 68%.

Cross-referrals

One study reports cross-referral rates of 27% from BEXR to colonoscopy and 8% in the reverse direction (Lindsay et al. (1988)).

Procedure costs

Assumptions on the frequency and cost of treating complications are as in Chapter Four. The cost of colonoscopy is £105.10 but complications are now endogenous to the model, hence a net figure of £103.30 is used; for BEXR the cost of £78.87 is adjusted to £78.07 on a similar basis.

Treatment costs of missed cases

The cost of a missed case is based on the difference between an elective and an emergency admission. This assumes that all cases diagnosed at investigation are of the former type while all cases missed are in the latter category. Patients admitted a s emergencies stay in hospital for four days longer on average (see literature review in Chapter Ten), valued at £116.90 per day (as in Chapter Four) implying a total of £468 per case missed.

Results

In a population of 100 people, 8 of whom have cancer, the total cost of the regime which prefers colonoscopy is £11,322 while the cost of the regime which prefers BEXR is £11,874, a difference of £552 or £5 per person investigated. In addition colonoscopy would detect 7.6 cases of disease while BEXR would detect 5.44 cases. There is some evidence that the radiological misses are most likely to be small stage A cancers in the sigmoid colon (Aldridge and Sim (1986); Farrands et al. (1983); Kalra and Hamlyn (1988)). As the cure rate for these cases is particularly high the health 'costs' of a missed diagnosis may be severe.

Analysis of robustness

Clearly, the model is constructed from many different sources with variations in sample sizes and study design making bias possible Under what circumstances would the conclusion be reversed? While it is easy to say what values of the variables give the BEXR strategy a lower cost, this still does not take account of the higher sensitivity of colonoscopy. A rise in the relative cost of colonoscopy by 7% or a fall in the cost of BEXR by the same amount would produce this effect, as would a cross-referral rate from colonoscopy exceeding 15% or of less than 22% in the opposite direction. Large changes in other variables would be necessary, although it is interesting to note that BEXR has a cost advantage if the prevalence of cancer in the group investigated falls by half.

One extension of this model is to compare colonoscopy against a regime using both BEXR and flexible sigmoidoscopy (referred to below as the combined strategy). The same studies used above suggest that there is very little difference between the sensitivity of these two strategies (Aldridge and Sim (1986); Irvine et al. (1988); Durdey et al. (1987); Farrands et al. (1983); Kalra and Hamlyn (1988)). The cross-referral rate is also similar. On the basis of the investigation costs calculated in Chapter Four, colonoscopy has a slight advantage as the initial investigation (£103 versus £125). Cross-referrals from colonoscopy will not incur the full cost of the combined strategy since only BEXR will be required in most cases. The referral rate from the combined strategy will still be high if lesions are detected beyond 60cm. Finally, the need for two investigations is likely to delay the diagnosis slightly, increasing the anxiety to the patient and the risk of the cancer obstructing and resulting in an emergency admission. Counting against colonoscopy, however, is the comparatively high risk of fatal complications. If a factor to reflect the risk of mortality is included in the costs in Chapter Four, valuing a life at £500,000 in accordance with Department of Transport assumptions (Department of Transport (1987)) then radiology is preferred under almost all circumstances.

This model has been constructed with the detection of cancer in mind, although it has included the cost of cross-referrals to treat adenomas by polypectomy. Most of the studies above have also found colonoscopy to be

more sensitive to adenomas, hence the assumption of no benefit to excision is likely to count against colonoscopy, making this scenario the 'worst case' for that investigation. In general, BEXR tends to over-diagnose very small lesions as it is unable to distinguish faeces from polyps in some cases: this reduces its specificity.

Ultimately, the choice of investigation must depend upon the local availability of resources and the skill of the practitioner: colonoscopy is a technically difficult procedure, and in untrained hands the rate of incomplete investigations and even of complications could be high. This model has identified the best way of detecting symptomatic cancer but this may not be the aim of the doctor taking a decision in an out-patient clinic. More significant may be to find a source of the patient's symptoms, and in this respect BEXR may redeem itself by its higher sensitivity for other bowel diseases.

INVESTIGATION OF POSITIVE SCREENING RESULTS

Two issues are considered under this heading. Firstly, the need to minimise the number of false positive investigations is discussed and the option of retesting positive results prior to diagnostic investigation considered. Secondly, the choice of diagnostic test following a positive screening test is evaluated using data for the relevant population to replace that for a symptomatic group wherever possible.

The role of retesting positive results

On the basis of Chapter Four, 57,800 complete a test out of a population of 100,000 and 3.4% have an initial positive test, implying 1,965 would require colonoscopy. Even allowing for the fact that this would be spread over two years, the workload would swamp the facilities currently available; a report from the British Society of Gastroenterology suggested that at least 200 investigations per 100,000 population would be required but 500 should be allowed for (Working Party (1991)). The screening programme alone would double this figure, with extra investigations in the group not accepting screening, those outside the age range and those being followed-up after adenoma excision to be added on. At present, therefore, any such policy would not be feasible. One way to cope with this is to reduce the false positive rate prior to diagnostic investigation; the problem is to achieve this without any loss in sensitivity.

As outlined in Chapter Two, the MRC trial protocol retests all positive results using a six-day FOB test to be completed while observing dietary restrictions: retest positives are investigated (Thomas et al. (1989)). Retest negatives receive a third test after three months and once again positives are investigated while negatives are assumed to have had a false initial positive and are not contacted again until the next round of screening. Participation is high for the first retest but falls to 79% for the second retest. Investigation is

by colonoscopy which, for the purposes of this section alone, is assumed to be definitive and fully accurate at the first attempt

Data from the MRC trial show that, of 647 people who carried out the retest, 251 (39%) were positive and hence were investigated without further delay; 35 cancers were found in this group (prevalence 14%). Of the 396 retesting negative, 79 refused the three-month retest (20%); there have been no symptomatic cases of disease in this group to date after a minimum follow-up of 18 months. Of the remaining 317 people, 31 (10%) were positive at the second retest; investigation found four cases of cancer in this group. Follow-up of the 286 negatives at this retest has found one case of advanced disease after a median follow-up of 24 months.

How many cases of disease were present initially? Retesting finds 39 cases (35 retest positive, 4 second retest positive) and misses one case. It is possible that other cases of disease which retested negative were still asymptomatic at 24 months. Also 79 people refused the second retest: if the decision to participate with this retest was independent of whether or not the person had asymptomatic disease then there were also 1.2 cases of disease in this group which had not presented within the period of follow-up allowed. Thus the total yield of initial colonoscopy would have been 41.2 cases. In addition 5% of those offered retesting do not respond i.e. the above figures excludes 32 people who were initially positive but refused the retest. Based on the disease prevalence among acceptors, this implies that two cases of disease in this group had also not presented within the follow-up period.

How accurate are the imputed prevalence rates for refusers? Since participation is negatively correlated with age, one would expect a higher prevalence rate in this group than in those who accepted. However, if the cases have not presented in the follow-up period then they behave quite benignly. Nevertheless, in line with the remainder of this analysis, the intermediate outcome measure of screening is taken to be the number of cases of malignant disease detected, and this assumption is continued here.

The cost of the MRC trial protocol is 679 initial six-day retests plus 396 further six-day retests at three months, plus colonoscopy for 251 positives at the first retest and 31 positives at the second retest. As calculated in Chapter Five, six-day testing costs £3.03 including postage and development while the cost of colonoscopy is £105.10 (from Chapter Four). In addition, the cost of case presenting in the interval following a negative screen has been shown to exceed that of a screening detected case by £612, after allowing for discounting to a present value (Tuck et al. (1989)): this is the cost 'penalty' of a missed case, the health 'costs' of which are still unknown.

Retesting detects 39 cases but one is missed and a further 3.2 are among those refusing one of the retests. The total cost of retesting is £35,465 (of which colonoscopy represents 84%) and the average cost is £909.

Colonoscopy of all those testing positive initially (no retesting at all) would have cost £71,363, assuming 100% participation. The yield would have been 43.2 cases at an average cost of £1,732 each.

The difference in average costs seems quite small, but the extra 4.2 cases are detected at an extra cost of £35,898, or £8,547 each. A more realistic assumption on the participation rate for colonoscopy is the MRC trial figure of 92%: this would reduce the cost of the 'no retesting; strategy to £65,654 with a yield of 39.7 cases (assuming participation to be independent of disease prevalence). The extra cost per extra case is now £43,127.

A third possibility is to stop offering a second retest three months later: compared to the full retesting protocol this would save £4,458 but miss four cases.

Discussion

A protocol of colonoscopic investigation following a single positive test is expensive; 94% of colonoscopies performed will be negative for cancer. Over 90% of the yield can be obtained at 50% of the cost, even if full participation is assumed for colonoscopy. Under more realistic assumptions, 98% of the yield is obtained for 54% of the cost, in line with other clinical research findings (Elliot et al. (1984)).

The Haemoccult test may lack sensitivity in the asymptomatic population but it is accurate as a means of identifying a high-risk group who require investigation from among initial positives. Those who have two negative retest results are returned to the general population; they are a low-risk group rather than a 'no-risk' group but the cost of detecting additional cases of disease appears high. By reducing the false positive rate one of the objections to mass screening by FOB testing is removed (Simon (1985)).

Of particular interest in this respect are the Swedish trial data on the effects of retesting following a positive test result using the rehydration technique for developing completed tests. Chapter Five showed that, while the sensitivity rate was high, almost 6% of the population required diagnostic investigation, counting heavily against the regime overall. This suggests a role for retesting.

Only limited data are available (Kewenter et al. (1990)). Of 579 initial positives, 544 completed a retest (94%) and 186 were positive (34%). As part of the research trial, all 544 people were offered investigation, although 27 refused (5%). Cancer was found in 14 people in the positive retest group (7.7%) and three in the negative retest group (0.9%). Using the same method as that described above 0.5 cases are present among those completing a retest but refusing investigation and one case among those refusing the retest. Investigation of all positives would have cost £57,810

based on the observed participation rate in investigations of 95%; the yield would have been 17.6 cases out of the total 18.5 cases present in the whole group. The average cost would have been £3,285. Retesting costs £1.65 per person (including postage and development) giving a total cost of £20,504 with a yield of 14 cases at an average of £1,465. The extra cost of detecting the 3.6 extra cases is £37,306 or £10,363 each.

Retesting in the Swedish trial could substantially reduce the number of people requiring investigation (to 32% of those testing positive initially), but this is at a loss of sensitivity: if the test is 85% sensitive on the basis of the initial test alone then the retesting protocol reduces this to 70% and much of its advantage over unhydrated testing is lost.

In summary, retesting initial positive results can reduce the health service costs of screening as well as the number of people investigated unnecessarily; this is crucial in making screening a practical possibility given currently available diagnostic facilities. The trade-off is the small number of cases initially positive but retesting negative: the acceptability of this protocol depends on the policy maker's willingness-to-pay to detect one more case.

Choice of diagnostic test

The results above are based on the study of patients presenting with symptoms of lower gastrointestinal disease. Differences between this case and that of the asymptomatic screened population include the greater proportion of neoplasia in the sigmoid colon and the added significance of adenoma detection and excision.

In terms of cancer detection, several reports from the randomised FOB screening trials have been critical of BEXR in comparison with colonoscopy. The proportion of cancers screening positive but missed on BEXR alone has been reported as 11% (Kewenter et al. (1987)), 25% (Winawer et al. (1985)) and 36% (Mandel et al. (1988)). An uncontrolled study of asymptomatic screening has reported that nine out of the 12 cancers detected were negative on BEXR, including all of the stage A cases, with sites in the sigmoid and transverse colon and in the caecum (Elliott et al. (1984)); this confirms the evidence presented above for symptomatic cases. The accuracy with regard to adenomas is harder to calculate because of the lack of prevalence data. As noted above, the health 'cost' of missing a stage A cancer are potentially large, but a further complication in the screening case is that if such lesions are asymptomatic then they are also most likely to be 'lifetime latent' cases.

Swedish trial reports acknowledge the weakness of BEXR alone but believe that in combination with flexible sigmoidoscopy an accuracy at least comparable to that of colonoscopy can be achieved (Jensen et al. (1990)). All the other trials use colonoscopy wherever the patient is fit enough, with recourse to BEXR only in cases where a full examination proves impossible.

To address this issue a similar approach to that outlined for symptomatic people is used.

Prevalence

The results of three-day testing in the MRC trial show that the prevalence of cancer in the group undergoing investigation is 15.3 per 100 (Thomas et al. (1990)). Of these 13% were in the rectum, 69% were in the sigmoid colon, 8% were in the transverse colon (including flexures) and 9% were in the caecum. The Danish trial shows a slightly more even spread, albeit with smaller numbers; the combined series gives figures of 20%, 60%, 11% and 10% respectively. To estimate what proportion of these cases would be identified by each strategy site-dependent sensitivity data are required together with an estimate of the probability the segment would be satisfactorily visualised.

Cross-referral rates following investigation

To assess the number of patients requiring a second (or subsequent) investigation when colonoscopy was the investigation of choice, the MRC trial records of 163 patients investigated following a positive test result were examined. 12% were unsuitable for the first choice investigation, colonoscopy, and were investigated by BEXR; half this group also underwent flexible sigmoidoscopy. The remaining 88% of patients underwent colonoscopy; 30% subsequently required a BEXR to complete the investigation and a further 6% needed a second colonoscopy. In all, only 56% of the sample were investigated by a single colonoscopy alone, with 9% of the sample requiring three investigations. Of those initially deemed unsuitable for colonoscopy, over one-third (7/19) were eventually investigated by that method.

In total, the 163 patients were offered 165 colonoscopies (1 refusal), 65 BEXRs (3 refusals) and 10 flexible sigmoidoscopies. On the basis of the costs in Chapter Four, the total cost was £22,604 (76% on colonoscopy), or £139 per person.

The combined investigation (BEXR plus flexible sigmoidoscopy) has a cross-referral rate of at least 19%, based on data from the Swedish screening trial (Jensen et al. (1990)). This is a low estimate owing to the rehydration screening method having such low specificity. In a study of symptomatic patients the view on the combined investigation was unsatisfactory in 18% (Irvine et al. (1988)) to which must be added the need for polypectomy of any findings beyond 60cm.

In the early stages of the MRC trial, the combined investigation was preferred owing to the unavailability of colonoscopy, producing a cross-referral rate of 21%. Unfortunately the screening protocol at the time offered tests to people aged less than 50, used the Fecatwin/Feca-EIA test which has a high false positive rate, and did not use the retesting technique described above. The

results are also biased estimates of the figures obtained following three-day testing without rehydration but with retesting.

Completeness of investigations and sensitivity

In the MRC trial, 30% of colonoscopies are incomplete and require a BEXR; rates as low as 8% have been reported, however (Rex et al. (1990)). It is assumed that this means that 30% of colonoscopies fail to reach the caecum and 15% do not visualise the transverse colon or the flexures. The scope will always pass through the rectum since rigid sigmoidoscopy will previously have proven possible; however 1% of investigations do not get beyond the rectosigmoid junction (e.g. due to strictures or to obstructing carcinoma). Site-dependent sensitivity for colonoscopy is taken to be 98% for the rectum and sigmoid: the former can be missed by the endoscopist when the scope is initially inserted and in addition there is some evidence from the Swedish trial that endoscopy can miss cancers in the rectosigmoid junction. When they can be visualised sensitivity in other sites is 99%: studies which have undertaken repeat examinations either immediately or at a six week interval have failed to demonstrate further cancer (Hixson et al. (1990)).

The combination investigation is assumed inadequate in 20% of cases, based on results for symptomatic patients (Irvine et al. (1988), Lindsay et al. (1988)). As an initial assumption the completeness of investigations of the sigmoid is taken to be 97%, based on reports of a 'blind spot' between the sigmoid and descending colon. The radiological investigation may also have some problems visualising the caecum (one case was missed in Rex et al. (1990), with several missed cases on follow-up of negative BEXR in the MRC trial). Arbitrary assumptions are made that the sensitivity is low (95%) but that sigmoidoscopy makes up for many radiological failings in the rectum and sigmoid colon: nevertheless, the Swedish trial has shown that misses occur and a rate of 98% is assumed for both sites.

The initial version of the model uses the following assumptions, although these can be varied:

Table One

	Colonoscopy		Combined Investigation	
	Complete (%)	**Sensitivity (%)**	**Complete (%)**	**Sensitivity (%)**
Rectum	100	98	99	98
Sigmoid	99	98	97	98
Transverse	85	99	99	99
Caecum	70	99	95	95

Costs for the three investigations are as in Chapter Four but exclude the cost of complications, as above The figures used are £103.30 for colonoscopy, £78.07 for BEXR and £46.81 for flexible sigmoidoscopy (implying a cost for the combined investigation of £124.88).

Results

The results for 100 people with positive results on unhydrated three-day Haemoccult testing with retesting (prevalence of 15.3 cancers) are as follows:

Table Two

	Total cost (£)	Cases
Colonoscopy preferred	13,782	14.83
Combination preferred	15,055	14.69

The two strategies are very evenly balanced with the difference only amounting to £13 per person investigated and 0.14 cases (9% and 1% of the lower total respectively). Nevertheless, colonoscopy is both cheaper and more accurate under the assumptions used.

Sensitivity analysis and discussion

This result is robust in the face of changes in many assumptions because sizeable minorities of the people investigated under each protocol are cross-referred, thus muting the effects of (say) a change in test cost. A 10% reduction in the cost of the combined investigation still leaves the 'colonoscopy preferred' option cheaper. A shift in the site distribution towards the right-hand side of the colon would favour the combined investigation given the low colonoscopy completion rate but this would have to be large to reverse the conclusion. Given the number of assumptions on the sensitivity and completeness of investigations it is difficult to present the results of a systematic variation; however, a scenario can be constructed in which the combined investigation would be more accurate than colonoscopy. Suppose that the combined investigation achieved a sensitivity of 99% for rectal cancers and 98% for sigmoid and caecal cancers, with the sigmoid fully visualised in 99% of cases then it would detect only 0.01 extra cases at an additional cost of £1,204. The extra cost per extra case detected is £114,331. Note that this scenario exceeds the claims of 'combination' advocates: the Swedish trial reports sensitivity in the sigmoid area of only 90% (Jensen et al. (1990)).

Studies of patient preferences have shown little significant difference between each type of investigation, although a weak trend is favour of colonoscopy can be detected (van Ness et al. (1987), Aldridge and Sim (1986), Lindsay et al. (1988)). The unpleasant side-effects of out-patient flexible sigmoidoscopy, such as incontinence while travelling home, has also been demonstrated (Thomas et al. (1990)).

The choice of a diagnostic investigation in a screening programme brings to the fore the question of what the programme is trying to detect. The calculations above have assumed that the objective is to detect asymptomatic cancer, but many claim advantages to detecting and excising adenomas (Atkin et al. (1993)). Adenomas are included above as a cost factor, increasing the number of patients with radiological findings beyond 60cm.

requiring colonoscopy as well as causing a proportion of the repeat colonoscopies for excision in expert hands. There is evidence, however, that any benefit would also accrue to colonoscopy; even the Swedish study claims no decisive advantage for either form of investigation (Jensen et al. (1986)).

CONCLUSION

One researcher has written, "We may have stepped unwittingly into a diagnostic quagmire when we embarked down the road of colon cancer screening for the general population." (Lieberman (1990)). Certainly, the choice of diagnostic test is finely balanced, but the evidence suggests that colonoscopy should be preferred in centres where the caecum is reached on at least 70% of examinations.

Chapter Ten

THE IMPACT OF SCREENING ON TREATMENT COSTS

Introduction

By changing the staging distributions of cancers at detection, screening can affect the costs of treating the disease. Treatment may be simpler, for example by allowing the endoscopic removal of small cancers (polypectomy). Patients should generally be fitter than symptomatic presenters, implying better post-operative survival, and a shorter hospital stay. This chapter analyses data from the MRC trial to consider the size of this effect. Preceding this is a discussion of the methodology of costing hospital care.

The need for better cost data - case proven?

Routinely available cost data suffer from a number of faults, particularly that it rests so heavily on average costs and thus does not relate directly to the costs incurred in treating any individual patient. Attempts to improve on these figures have centred on increasing the relevance of costs to each individual case. These gains in accuracy for the individual are expensive in terms of research time, however, hence all such refinements must be considered against the criterion of whether the increased accuracy affects the results of the evaluation. As one textbook puts it: "It is important to question the purpose of refinements since they are likely to involve considerable time and effort. One should not necessarily aim to produce the best possible information every time." (Drummond (1980)). A number of studies reflect this concern (e.g. Brooks (1981); Piachaud and Weddell (1972)).

An example of a study where detailed work made no difference to the ranking of the options and hence to the results of the evaluation is that of Culyer and Maynard (1981). They compared treatment options for duodenal ulcers and showed that a 'refined' approach to costing gave a cost per operation one-third lower than that derived from multiplying the national average length-of-stay for the procedure by the national average cost per day for hospitals of that size. However, re-working their results shows that the use of the national average figure would not have affected their conclusion.

In general technical efficiency questions (identifying the cheapest way of providing a given service) generally require less refined costings since the options often have common elements which cancel each other out. For example, Culyer et al. (1983) found no significant difference between groups using different types of wound dressing in terms of length-of-stay, and hence were able to ignore it.

The chance of a refinement affecting an evaluation is related to the significance of the costs involved in the overall total cost; thus, a small error in a major item is potentially more important than a major error in a small

item. The major items in acute hospital budgets are nursing and medical staff time and service overheads. It is easy to get bogged down in counting the number of pathology tests or aspirins but unless these items are particularly important to the treatment process the impact on the total cost will be minimal.

One commentator has summarised the position as follows "It would be incautious to conclude that a major increase on expenditure on costing is necessary. Costing is itself a costly activity, and since it is likely to be subject to diminishing returns in improving decisions, it will not be worthwhile seeking to abolish our ignorance entirely. It is approximate information that should be sought rather than perfection." (Hurst (1978)).

Many of the issues arising can be illustrated by a detailed consideration of the handling of nursing costs which constitute over one third of acute sector spending (Department of Health (1992)).

NURSING COSTS: Alternative allocation methods

The accountant's approach to allocating nursing costs is to determine the total nursing bill for each specialty and then to divide by some measure of workload such as the occupied bed-day. Thus, within specialties, patient costs per day are assumed to be homogeneous. (Some evaluations using these figures have also attempted to verify that the relevant patients are indeed typical (e.g. Gravelle et al. (1982))).

To reflect the care received by individuals some evaluations have used measures of nursing dependency as a proxy for the allocation of nursing time. In an evaluation of treatment for patients with bleeding peptic ulcers (Rees (1987)), ward sisters estimated the care required by a patient in the four categories of a standard dependency measure, the Telford Scale. Scores for each category were obtained by the ratio of the care in that category to the care given to a patient in the least dependent category. Summing these scores across the patients on the ward gave an estimate of the total workload and hence a cost per workload unit was derived from the daily total nursing cost. The costs used were as follows:

Table One

Category	Workload score	Daily cost (£)
Ambulant (pre-op., convalescing)	1	8.80
Semi-ambulant (recovering)	2.5	22.01
Partially helpless (a few days post-op.)	4.5	39.61
Totally helpless (immediately post-op.)	8	70.42

Nursing records were used to determine the patient's dependency category day-by-day, and hence the cost of nursing them in the course of their stay.

An even more thorough analysis using a similar method was conducted as part of the heart transplant study; the researchers carried out extensive samplings of the time grades of staff spent with various patients (Buxton et al. (1985)). However, this is unlikely to be a practical possibility for any but the best funded study: others have chosen the simpler method developed by Rees (e.g. Ridley et al. (1991)).

An application to the care of patients treated for colorectal cancer

To explore the increased accuracy of the dependency-based method, data were gathered for a sample of 65 patients undergoing a resection for colorectal cancer over a six-month period at University Hospital, Nottingham (UHN). In two cases, patients were readmitted within one calendar month of discharge and the two care episodes are treated as one continuous episode for each patient. As UHN is a research centre, patients are often admitted several days earlier than necessary in order to participate in medical trials: only the three days preceding the operation are considered.

The staffing structure of the relevant surgical ward was costed using total salaries plus national insurance and pension contributions, as in Table Two:

Table Two

Grade	Number	Unit cost (£)	Total cost (£)
Sister	1	13,489	13,489
Staff Nurse	4	10,278	41,112
State Enrolled Nurse	2	8,896	17,972
Auxiliary Nurse	2	6,676	13,352
Student Nurse			
(Third year)	2	6,634	13,26
(First year)	4	5,742	22,968
TOTAL			121,161

Average occupancy of the 28 beds was estimated to be 91.7%, implying 9,375 patient bed-days per annum. The average nursing cost per day per occupied bed was £13.03.

The division of time between shared care and care for individual patients was determined by the ward sisters. The proportion of total time spent on 'shared' care as opposed to that spent with individual patients was estimated to be:

Sister	60%
Staff nurse	40%
State Enrolled nurse	25%
Auxiliary nurse	50%
Student nurse	40%

The employment cost for each grade was then divided according to these proportions, to yield total costs of shared care (£50,202) and patient-specific care (£71,959).

Dependency data are routinely collected at UHN using a system developed in North Lincolnshire District Health Authority and based on a scale similar to that described in the peptic ulcer study. Ward sisters estimated the nursing time required for each dependency category, as in Rees's study. These data were then treated as units of work, one unit being the daily care received by a category 1 patient. Table Three displays workload ratios, the observed average dependency mix for the wards and the derived total number of daily workload units.

Table Three

Category	Workload ratio	Patients	Total work per day
1	1	5	5
2	2.5	13	32.5
3	4.5	5	22.5
4	7.5	3	22.5
Total		26	82.5

Thus the average daily workload was 82.5 units, or 30,116 per annum, giving a patient-specific cost per workload unit of £2.39.

Results

The mean length of stay of the sample was 15.1 days within a range of 10 to 38. Figure One shows the pattern of changes in dependency category for the sample by day of treatment up to the tenth day post-operatively. For any given day of treatment the top point of the line represents the highest dependency category recorded in the sample, while the bottom point represents the lowest category (0 implies the patient was discharged). The mean dependency category is denoted by the mark within the range. Most patients were in categories 1 or 2 prior to operation and in state 4 on the day of the operation, with a gradual reduction in dependency to state 1 or 2 before discharge, although some patients are very dependent for several days after the operation.

The 'average' patient requires less care in the earlier and later stages of the episode than implied by the average method, although this is compensated by above-average needs immediately after the operation. This is illustrated by the cost per day under each method in Figure Two.

Total nursing costs across the length of stay were calculated for each patient using each method. The mean total cost for each method were identical at £195.90. For individual patients total costs calculated by each method were highly correlated (Spearman rank correlation coefficient, $r = 0.93$, $p<0.01$),

Figure One - Dependency state by day of stay

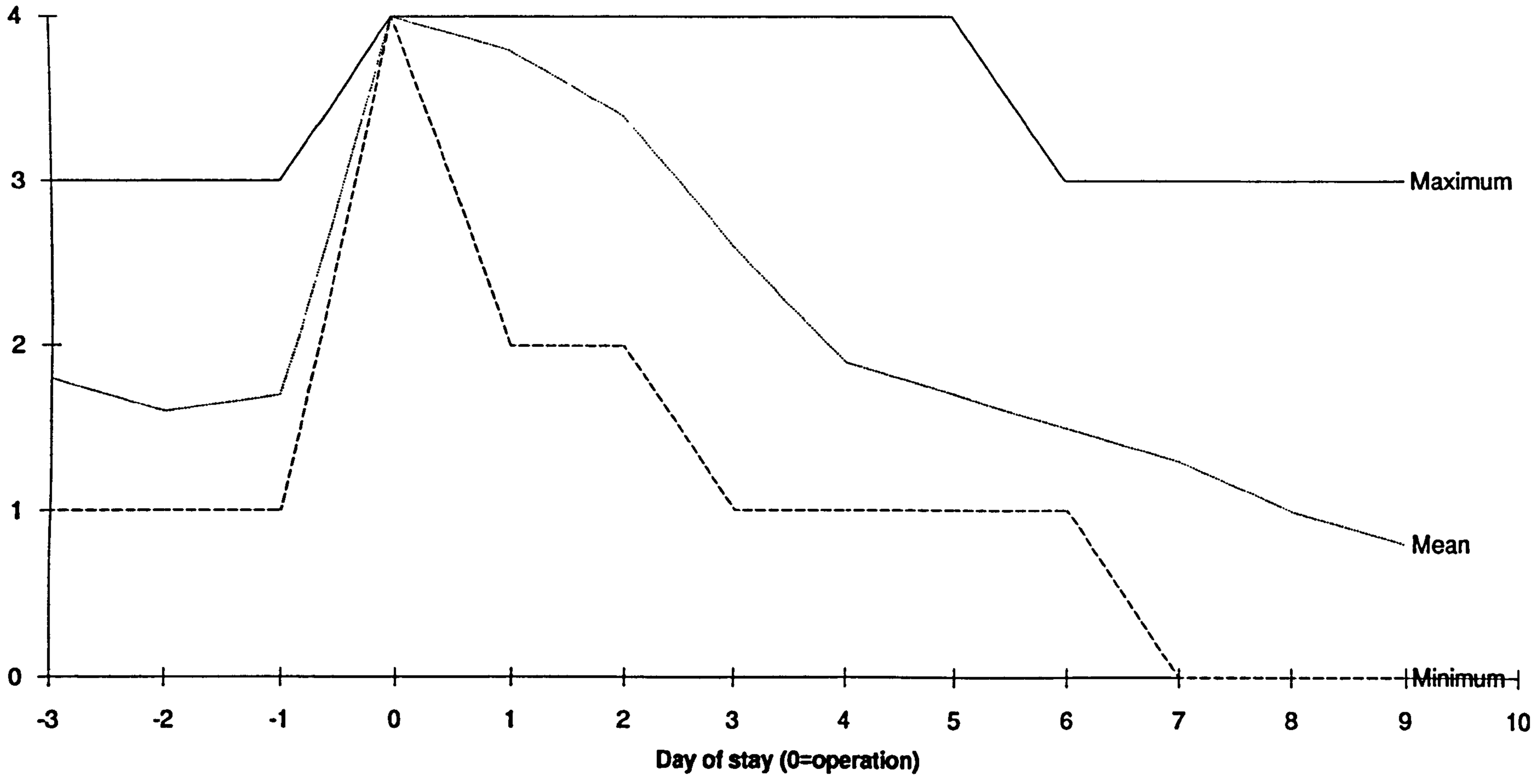
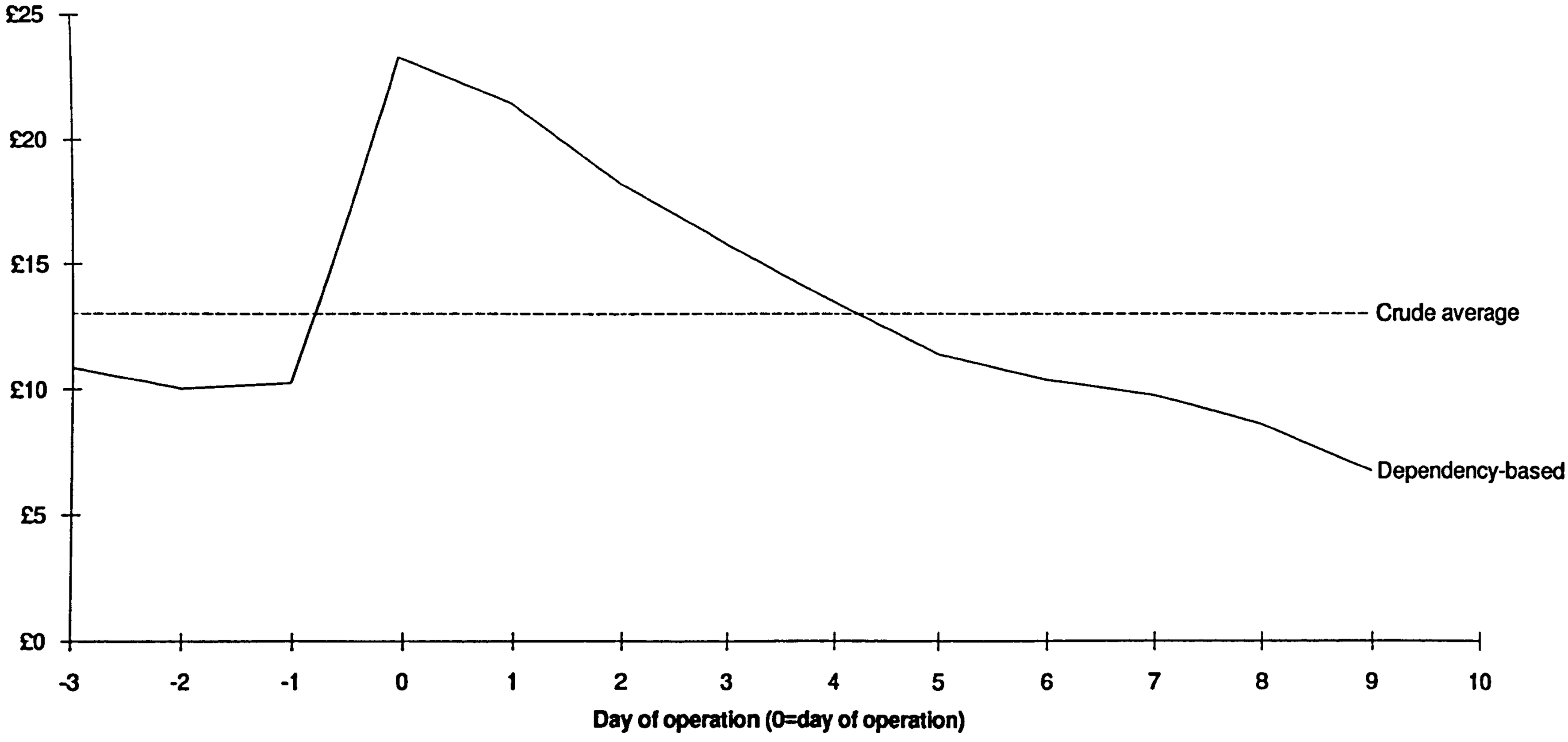


Figure Two - Cost per day by each method



owing to the strong association between length of stay and patient-specific workload units ($r = 0.82$, $p<0.01$).

Figure Three shows the difference between the cost of each method for individual patients. In 12 cases out of 65 the two costing methods produce the same result within 2.5%. In more than one third of cases (23/65), the two estimates differ by more than 7.5%, although the number of over-estimates above 2.5% (29/65) is approximately equal to the number of under-estimates below 2.5% (24/65). These differences are associated with the number of patient-specific workload units ($r = 0.32$, $p<0.01$) but not with dependency-based costs, average costs or with length of stay.

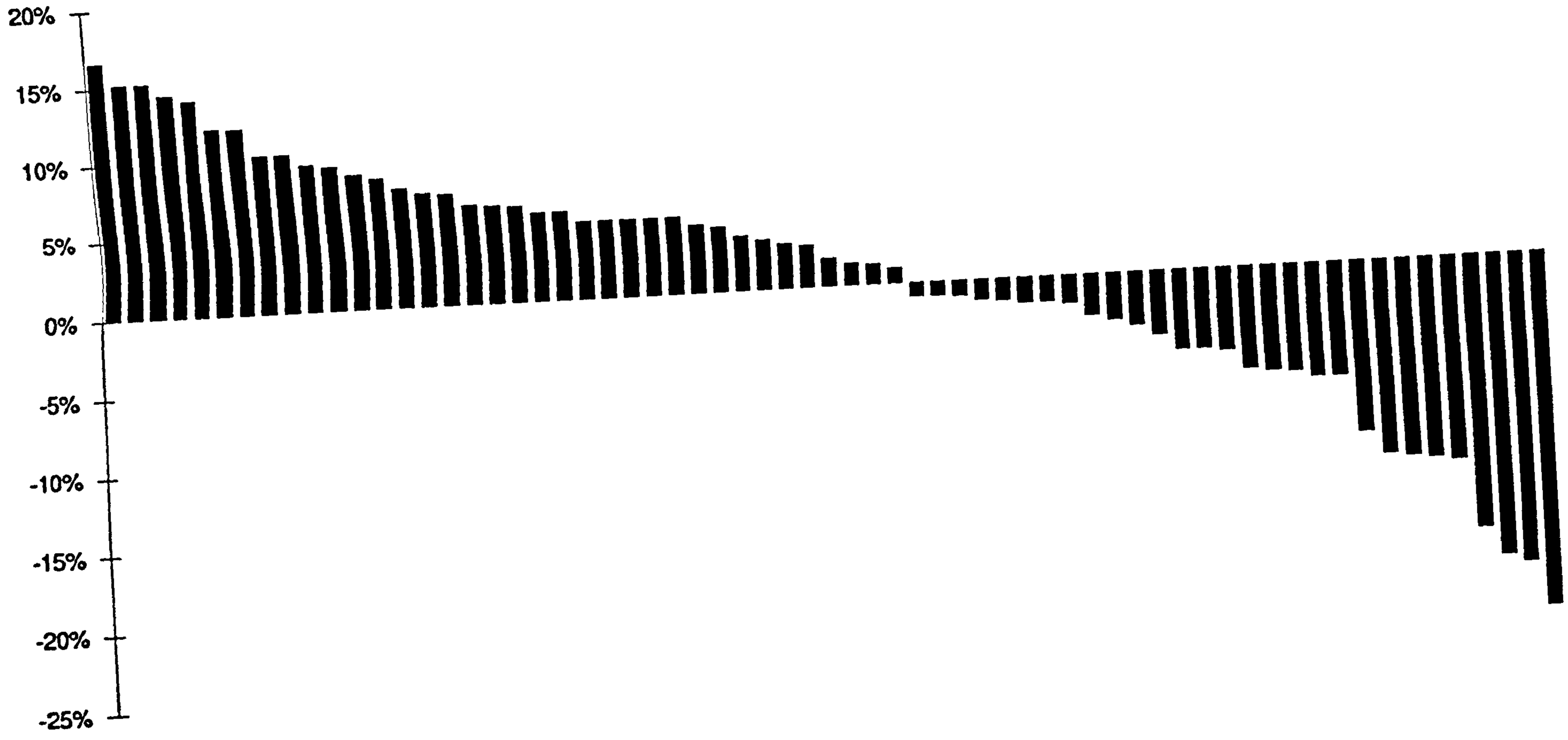
Discussion

The dependency-based approach offers a superior measure of the day-to-day variance in resource consumption. The fundamental problem is how accurate does a cost figure have to be? Three-quarters of the estimates derived above are equivalent (within plus or minus 12.5%), implying that the refinement is unlikely to affect the evaluation results. A re-working of Rees's data show that a simple average cost would also have been adequate in that case.

Given the increasing availability and routine use of dependency measures, data are easily available and have been used in other studies (Brooks (1981)). However, a number of assumptions are required which would ideally be tested in a sensitivity analysis. More importantly, the scales used measure dependency alone, which is a subjective concept, influenced by the experience of the nurse making the judgement (Proctor (1993)). There is some evidence that dependency does not correlate with time spent with a patient (and thus resource consumed) (Bagust (1990)); empirical evidence of this is provided by the heart transplant study which found large differences within dependency categories (Buxton et al. (1985)). This was acknowledged in the peptic ulcer study: "The ward staff costs covering the salaries of medical and nursing personnel are in fact medium term marginal costs, and the element of variability in response to changes in the level of patient care is more apparent than real." (Rees (1987)). In other words, the dependency approach is a model which is not based on observed behaviour. The dependency-mix of the ward influences the average cost per workload unit: if the ward were full of people in the least dependent state the cost per workload unit would rise and care would appear to be more expensive, whereas in practice nursing staff would be reallocated to busier wards. The model has nothing to say about the quality of care: if total staff time is fixed in the short term, a rise in workload means quality may suffer.

Another problem is the allocation of 'shared care', which is equivalent to a ward-specific overhead cost. The allocation is inevitably arbitrary; for example, the peptic ulcer study (implicitly) and the heart transplant study (explicitly) allocate shared care in direct proportion to differences in individual

Figure Three - Difference betwen total nursing cost by patient



care. There is no 'right' answer, hence these issues will remain a matter for researcher judgement.

Implications for this study

Variations in the costs of treating individual patients are important, particularly if they can be predicted in advance, e.g. by linking them to some observable characteristic when the patient is admitted. However, the implication of the results is that these are unlikely to be significant for care following a bowel resection and that the cost included in the specialty total in hospital accounts is adequate.

ALLOCATING OTHER COSTS

This means that the allocation of two major items in the cost of treating colorectal cancer will either be by specialty average figures (nursing) or on an arbitrary basis (overheads). This implies there is little point in using detailed refinement methods for other categories.

1. Medical staff

Problems in allocating medical staff time to individual patients include the element of shared care and other duties. Prospective data collection is unpopular but there are no other sources available. Various proxy measures have been proposed but the situation is very complex: the stage of the patient's treatment, type of admission and operation, diagnosis and prognosis, and the number and nature of complications are all relevant. A scoring system using some of these factors was attempted in the peptic ulcer study but the approach was not validated (Rees (1987)). For this study, medical staff time is thus allocated on the basis of the specialty average per day.

2. Operating theatre

In the past evaluations such as the heart transplant study calculated a staff cost per hour and a standard cost per operation for equipment. This was used to value the duration of operations which recorded in patient notes and theatre records. A detailed study of theatre costs estimated an average cost per hour of £216.50 (NHS Management Executive (1989)). While these figures exclude gases, operation-specific items of equipment and other support services such as X-ray, they have been used in the present study.

3. Blood transfusion

The peptic ulcer study based cost on the charge made by the NHS blood transfusion service to the private sector (Rees (1987)): this is intended to reflect the long-run marginal cost of provision per unit of blood. The figures used were £20 per unit plus £3 per unit for grouping and cross-matching, and these have been updated.

4. Drugs

Drug dosage is recorded in case notes and can be valued according to the standard figures compiled in the British National Formulary. The recording of drug consumption for individual patients can be extremely laborious, however. A small pilot study confirmed that the costs of largely generic versions of drugs made little difference to the total costs of treatment. Further research could give more attention to a few expensive drugs or to define a standard regime to derive an average cost per patient based on a small sample. In the present study, however, drugs and fluids are ignored.

5. Pathology and radiology

Radiology and pathology tests are production processes in themselves, bringing together staff, equipment and raw materials: costing individual tests represents a major study in itself (Stilwell (1981)). Using published figures is dangerous if work practice, staff levels and available equipment vary but in colorectal cancer care the costs are not large enough to matter. The existing costing system groups procedures under a small number of headings which are (broadly) homogeneous in terms of resource use. In a similar manner to the dependency approach for nursing, each group is assigned a score which can be used to derive a total number of workload units for the department. The cost per point can be calculated and hence the average cost for each of the groups of procedures. This does not reflect the cost of any particular procedure, still less the resources consumed by an individual patient. In the present study the figures for UHN are used.

6. Overhead costs

General service departments either support the work on the wards (laundry and catering) or relate to the running of the hospital itself (administration and maintenance). Studies have attempted to devise more appropriate ways to allocate overhead costs to departments, wards and ultimately to patients, but, "The main point to note at the outset is that there is no unambiguously right way to apportion such costs." (Drummond, Stoddart and Torrance (1987)). Examples of enthusiasm outrunning common-sense abound. For example, the report on the cost of operating theatres mentioned earlier calculated the proportion of windows in each of the hospitals in its sample which were in the theatre block and apportioned the total window cleaning budget accordingly. Another methodological review considers allocating patient catering costs on the basis of a sample of menus offered to patients for meals retained for more detailed costing, as well as allocating more catering costs to male wards than to geriatric wards as the former are given larger portions (Hillman and Nix (1984)). Thus an accounting cost allocated equally over all the work of the hospital seems the only satisfactory method. In this study the method described in the Appendix to Chapter Four was used.

PREVIOUS STUDIES OF THE COST OF TREATING COLORECTAL CANCER

Despite growing interest in the costs of all forms of health care activity, the literature on the costs of treating colorectal cancer is sparse. This is more surprising in view of analysis of American Medicare data which reveals that colorectal cancer is one of the most expensive cancers to treat (Baker et al. (1989)).

Some of the findings were as follows:

- Treatment costs in the five years from diagnosis were £3,531 for colon cancer and £4,462 for rectal cancer (1990 figures) (Wessex Cancer Intelligence Unit (1990)). In the case of rectal cancer 55% of costs were due to hospital stay, 15% were costs of theatre time and 9% were the costs of colostomy supplies after discharge; all other categories represented less than 5% of the total.

- Mean length-of-stay is about 15 days for a resection. There is common agreement that post-operative complications are an important factor (Arabi et al. (1980); Payne et al. (1987); Tartter (1988)). Other factors considered include: need for pre-operative assessment, performing a two-stage operation, and need for blood transfusion. None of the studies found stage at diagnosis to have a significant effect.

- Only the Danish screening trial has reported the impact of screening on length-of-stay: the data imply a median reduction of four days (Kronborg et al. (1989)). Breast cancer screening forms an interesting comparison, however. Mathematical models predict that following a rise of about 15% in workload following the initial screening, the overall number of operations then resumes its previous level and trend (de Koning et al. (1990); de Koning et al. (1992); Holmberg et al. (1986)). Net treatment cost savings represent one-third of the costs of the screening programme.

THE IMPACT OF SCREENING ON TREATMENT COSTS IN THE MRC TRIAL

To examine the impact of screening on treatment costs in the MRC trial, data were gathered for all cases with a minimum of three years of follow-up since their diagnosis. This interval was chosen to maximise the sample size but also to allow time for disease to recur and follow-up expenses to be incurred. MRC trial records identified 371 such cases and provided basic information on date of birth, date of entry to trial, date of death and cause of death (as recorded on the death certificate).

The main data source was the patients' hospital case notes. Clinical data collected included the cancer stage, site, differentiation, size and fixity, as well as symptom duration, type of admission, type of operation, number of post-operative complications and clinical assessment of prognosis at the time

of treatment. Resource-use data related to diagnostic investigations, length of stay in hospital, length of operation, radiology, pathology and ECG requests, and units of blood ordered.

Many patients were treated at University Hospital, Nottingham, a teaching and research centre, and hence procedures related to research were excluded wherever possible. For example, a comparative trial of pre-operative methods of imaging rectal cancer was in progress during part of the study period: all resources used as a result have been excluded. This creates particular problems with regard to post-discharge follow-up, and treatment of recurrent and of terminal disease. There appears to be considerable variation in practice even within a centre: some patients receive chemotherapy, others radiotherapy, while surgeons in Nottingham have started to use cryotherapy as an aid to the resection of liver metastases. (While the current focus is on the costs of treatment, this variation may also be reflected in the success of that treatment, of course). Practice outside of Nottingham is likely to show similar variations. These items were excluded from the analysis awaiting clearer statements on what represents 'good practice' in these areas. The costs used relate only to in-patient care for surgery or for other forms of patient care such as symptom control and non-surgical treatment such as endoscopic investigation.

Costing method

As noted above, the perspective of the costing is the use of hospital resources since these will have a direct bearing on opportunity cost of a screening programme to purchasers. This excludes costs incurred in the community (either by the NHS or other agencies); an example of a category excluded is the on-going cost of supplying colostomy equipment, although an estimate can be made from the literature (see Wessex Cancer Intelligence Unit (1990)) noted above).

Costs for hotel services, out-patient clinics, pathology, ECG and X-ray requests were derived from the Hospital Cost Returns (FR11) for University Hospital, Nottingham. Note that the hotel services excluded theatres, diagnostic and support services and overheads. General service overhead costs are derived using the method outlined in Chapter Four. X-ray costs use the radiology points system with chest, abdominal and pelvic X-rays counting as one point and ultrasound and upper gastrointestinal investigations as two points. The costs of other diagnostic investigations are as in Chapter Four. The duration of an operation was recorded from the anaesthetist's record and valued according to the average cost per hour calculated in the Bevan Report. (One type of rectal examination, the examination under anaesthetic (EUA), is common but rarely has a duration recorded: it was assumed that the duration was thirty minutes on average). The figure for a pathology request is the average of the costs of haematology, microbiology and clinical chemistry requests. The cost per ITU day is calculated from the Wessex figures assuming a similar ratio between ward and ITU costs; other recent

estimates include £399 per day (Ridley et al. (1991)) and £500 per day as an average across England (Shiel et al. (1989)).

Description	Unit cost (£)
Out-patient clinic	18.02
Hotel stay (per night)	76.10
Overheads (per night)	40.56
ECG	5.22
X-ray (per point)	5.11
Theatre (per minute)	4.47
Histopathology (per request)	36.42
Pathology	3.77
Flexible sigmoidoscopy	47.81
Gastroscopy	47.81
BEXR	78.87
Colonoscopy	105.10
ITU (per night)	332.90

Drugs and fluids were excluded.

As noted in Chapter Three, the follow-up period of three years for the sample is insufficient to justify discounting of future effects, although this would be important once long-term follow-up data are available.

Results

Of the 371 patients, eleven (3%) were excluded (nine treated privately, two hospital notes could not be found). This sub-set is sufficiently small that the results are not felt to be seriously affected.

Patient characteristics

The age and sex distribution for each group were:

Table Four

Control group	Number	Mean age
Male	90	67.6
Female	62	66.9
All	152	67.3

Study group	Number	Mean age
Male	119	67.4
Female	89	66.7
All	208	67.1

In the sample as a whole men outnumber women by 1.45:1. The male preponderance is not as large in the group offered screening. The study group can be sub-divided into three further groups: screening detected cases ('positives'), cases presenting symptomatically within two years of a negative

screening test ('intervals') and people who did not return a completed test ('refusers'). Of the 208 cases in the study group 77 were screening detected (38%), 26 were intervals (12%) and 105 were refusers (51%). The sex distribution is similar to the study group as a whole. Refusers are the oldest group (average age 68.4); positives and intervals have similar mean ages (65.8 and 65.6 respectively) but the variation in interval cases is higher (standard deviation of 9.06 versus 7.04 for positives).

Tumour characteristics

Synchronous cancers were found in four control cases (3%) and in two study group cases (1%); the convention of classifying these patients according to the more advanced stage is followed here. Table Five shows the staging distribution by trial group:

Table Five

Stage	Control	Study	*Positive*	*Interval*	*Refuser*
A	15	60	*39*	*7*	*14*
B	42	58	*18*	*4*	*36*
C	49	41	*15*	*7*	*19*
D	43	49	*5*	*8*	*36*
Missing	3	0			

Screening detects cases at an earlier stage but the refuser group nullifies much of the advantage over the control group. The 56 excess cases in the study group represent a combination of the early detection of future cases and over-diagnosis of 'lifetime latent' cancers; given the very similar numbers of stage C and D cases in each group, over-diagnosis may be an important problem.

The distribution of cases within the large bowel* is as follows:

Table Six

Site	Control	Study	*Positive*	*Interval*	*Refuser*
Rectum	74	86	*28*	*10*	*48*
Left colon	33	66	*33*	*5*	*28*
Mid colon	13	19	*8*	*2*	*9*
Right colon	27	30	*7*	*8*	*15*
Missing	5	7	*1*	*1*	*5*

Cases in the sigmoid colon were twice as common in the study group: this is almost entirely due to screening detected cases, and mirrors the results for adenomas presented in Chapter Seven. Cases can also be compared in terms of the size and differentiation of the tumour:

---

* Site distributions are defined as follows: rectum includes rectosigmoid junction; left colon includes the sigmoid and descending colon; mid colon includes the transverse colon and both the splenic and hepatic flexures; right colon includes the ascending colon and the caecum.

Table Seven

	Control	Study	*Positive*	*Interval*	*Refuser*
Mean size (cm.)	3.6	3.4	*2.9*	*3*	*3.8*
Differentiation					
Good	12	26	*18*	*3*	*5*
Moderate	86	126	*45*	*18*	*63*
Poor	24	19	*7*	*2*	*10*
Missing	30	37	*7*	*3*	*27*

Thus screening conforms to all expectations: in comparison with those presenting with symptoms, asymptomatic cases are smaller, are detected at an earlier stage, and are better differentiated.

As noted in Chapter Three, many American evaluations assume that all symptomatic cases are homogenous. This assumption can now be considered further. Interval cases have proportionately more very early and very late stage cases than controls, although in size and differentiation they are similar to screening positives. One way of explaining this would be that some interval cancers are aggressive and grow quickly before presenting at a late stage, even though they are no larger than early stage cases. Others grow at a slower rate but are detected earlier either because the person has been alerted to the significance of bowel symptoms or because those who complete a screening test (even if it is negative) are more aware of their health and have a lower 'trigger' threshold for consulting their GP. It is also possible that GPs involved in the screening programme have different referral patterns to those who are not; this could also influence the (comparatively) high proportion of stage A cases in the control group.

The same pattern of extremes is found among screening refusers in comparison with controls. The incidence rate is higher among refusers than controls, probably because the former is an older group on average. The site distribution shows a higher proportion of refuser cases in the left side colon than would be expected from the control distribution. Differentiation and size are similar, but refuser cases are slightly older. How can this be interpreted? It has been suggested previously that people who do not accept a screening invitation are not homogeneous: some have a very poor opinion of their own health and their ability to influence it while others do not perceive personal vulnerability to colorectal cancer. It may be the former group who represent the large number of advanced stage cases and the latter who represent the early stage cases. However, if this is correct then why are the same trends not present in the control group? Assuming that the two groups are truly identical, one explanation would be that the counterparts of control group cases in the study group are the symptomatic presenters (refusers plus intervals) and the cases of advanced disease in the positive group. The staging distributions would then be virtually identical. This implies that the benefits of screening (of whatever magnitude) have not yet had time to reveal

themselves. At this early stage screening detected cases would all still have been 'silent', prevalent cases.

Resource use

Before calculating the costs of treating cases from each group some details of the resources used is provided. As this information is untainted by local cost factors, it is more 'generalisable' to other locations.

Unlike the surgical series discussed in Chapter One, the present sample is drawn from a whole population rather than just those admitted to a surgical ward. In the trial as a whole, 3% of cases were not admitted to hospital: some were post mortem findings while others were treated by polypectomy on a day-case basis. Of those admitted 22% of controls were emergencies compared to 18% of the study group. Only 4% of screen-detected cases were emergencies compared to 23% of intervals and 26% of refusers.

As noted in Chapter One, surgery remains the principle treatment for colorectal cancer. Table Eight summarises the proportion of patients undergoing each type of operation as their initial treatment:

Table Eight

Operation	Control	Study	*Positive*	*Interval*	*Refuser*
None	21	12	*0*	*1*	*11*
Resection	112	154	*56*	*22*	*76*
(colostomy)	(28)	(36)	*(7)*	*(5)*	*(24)*
Endoscopy	10	20	*15*	*1*	*4*
Other	9	19	*3*	*2*	*14*
(colostomy)	(6)	(10)	*(1)*	*(0)*	*(9)*

(Note that the 'other' category includes laparotomy only, colostomy only and complicated palliative procedures). The main features of this table are the high proportion of screen positive cases amenable to polypectomy (in line with the finding that they are smaller tumours) and the high proportion of refusers who do not undergo a full resection owing to the presence of very advanced disease.

Variations in treatment are reflected in the average length of time spent in theatre (in minutes):

Table Nine

	Control	Study	*Positive*	*Interval*	*Refuser*
Initial operation	106	117	*115*	*120*	*118*
All operations	118	136	*133*	*135*	*139*

While screening-detected cases are smaller and more amenable to polypectomy, it appears that they take more theatre time. One explanation is that the perceived chances of cure are much greater in these cases, hence

they merit fuller attention. Another possibility is that the majority of screening-detected cases were treated in University Hospital, hence the operation may have been prolonged for teaching purposes.

A similar table shows the average length-of-stay in days surrounding the main initial form of treatment and over the whole course of treatment:

Table Ten

	Control	Study	*Positive*	*Interval*	*Refuser*
Initial stay	20	20	*15*	*18*	*24*
Total stay	26	28	*20*	*30*	*32*

Screening detected cases have a generally shorter length of stay than any other group, reflecting the proportion treated by endoscopy alone. This advantage does not extend to the study group as a whole because of the comparatively long stays in the refuser group.

Costs

Table Eleven presents the average cost per case in the study and control groups:

Table Eleven

Category	Control (£)	Study (£)
Investigation	103	121
Pre-operative	248	192
Main treatment	2,698	2,717
Readmit	150	271
Other follow-up care	805	1,010
TOTAL	4,005	4,311

These data are presented graphically in Figure Four. As the standard deviation for each sub-category is larger than the mean none of these differences achieves statistical significance. While the analysis has concentrated on the average costs of treatment, there are 37% more cases in the study group than in the control group implying a higher total cost. The total cost of treating the control group was £608,760 compared to £896,688 for the study group. The total cost is 47% higher in the study group. It is interesting to note that a similar excess has been observed in the Danish trial (see Chapter Two).

The results for the study group can also be analysed according to response to the offer of screening:

Figure Four - Cost of treatment by screening group

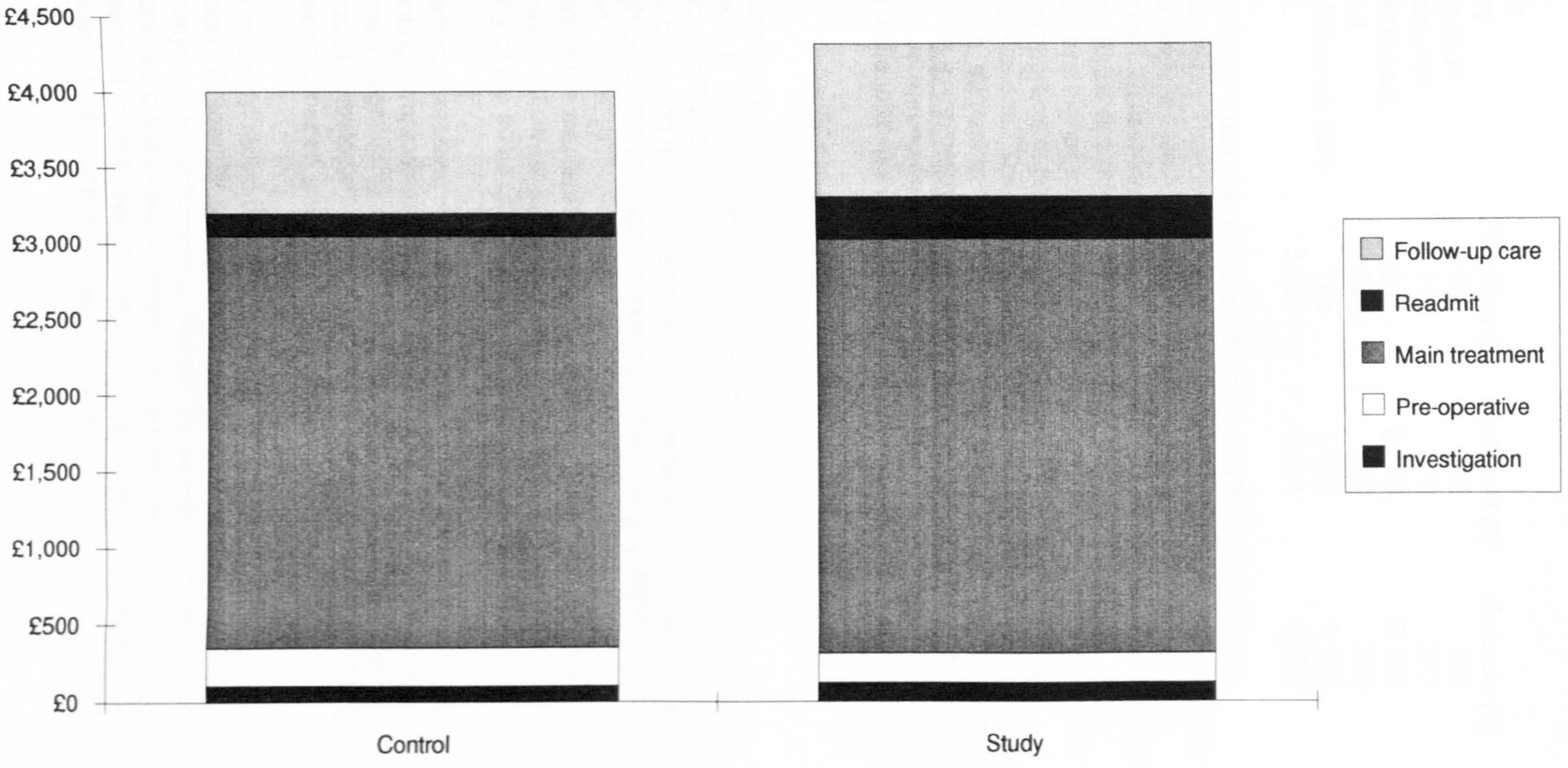


Table Twelve

Category	*Positive (£)*	*Interval (£)*	*Refuser (£)*
Investigation	162	86	100
Pre-operative	115	72	277
Main treatment	2,241	2,376	3,150
Readmit	218	447	266
Other follow-up care	734	1,440	1,106
TOTAL	3,471	4,421	4,899

Once again, standard deviations are so large that differences are not statistically different. The components of treatment costs in each sub-group is illustrated in Figure Five. The results reflect variations in resource use described earlier. Screening detected cases are more expensive to diagnose as there are no signs or symptoms to give a clue to the primary site. Pre-operative costs are higher in the groups that present symptomatically owing to the higher proportion of patients who require stabilisation following emergency admissions. Treatment costs are notably higher in the group who refuse screening, stemming from the longer time spent in hospital. Follow-up costs are highest in the interval and refuser groups; this is consistent with the hypothesis that screening detected cases are from the less aggressive end of the spectrum while intervals and refusers are more prone to recurrent disease.

Discussion

The results have several strengths:

i) the sample is large, giving a good chance of showing any significant differences in the data;

ii) the data collection methods adopted in the research trial mean that a very high proportion of cases occurring in the trial population have been identified;

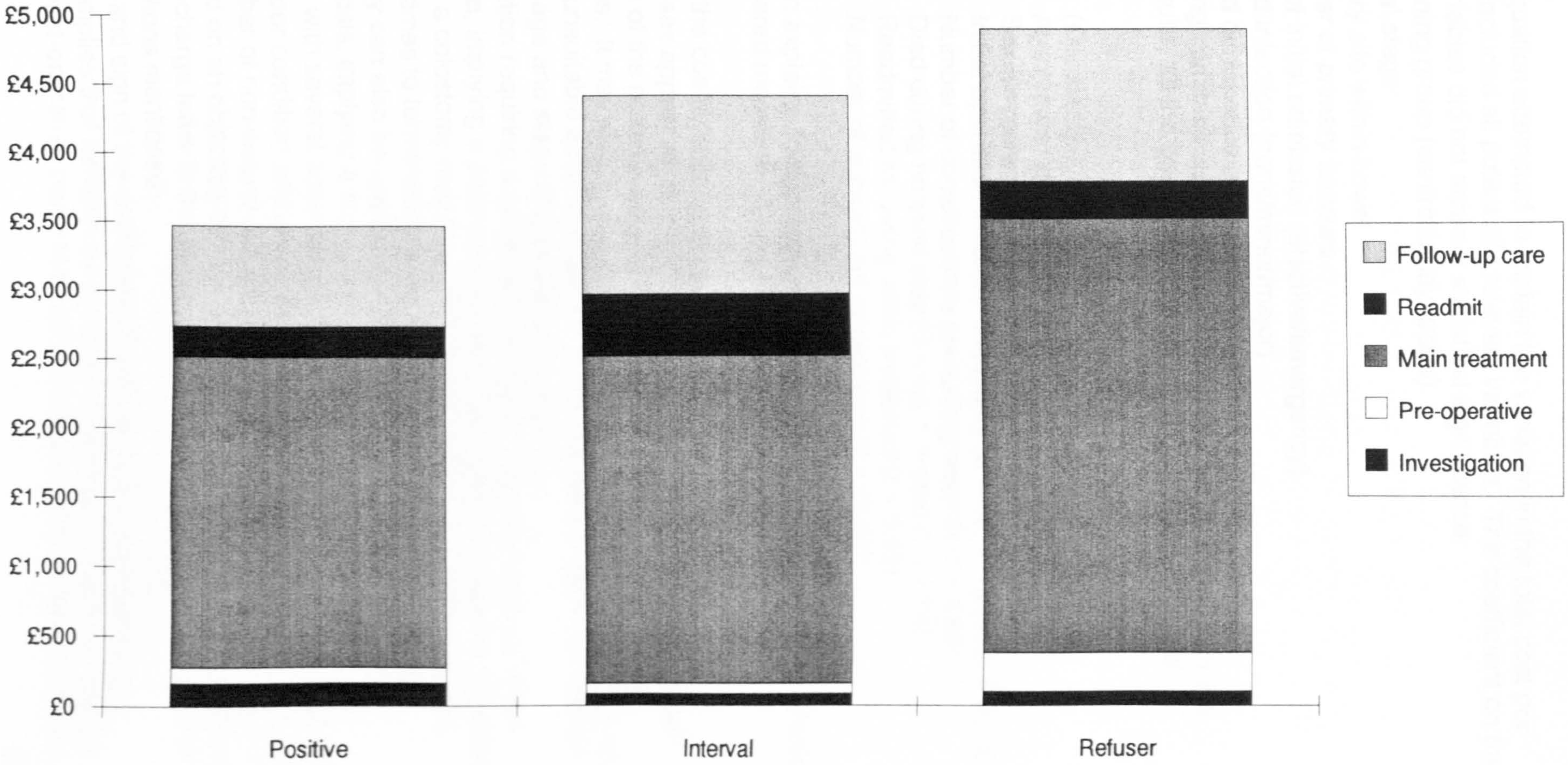
iii) almost all of the available hospital case notes were tracked down;

iv) the minimum follow-up period is three years from the operation, implying that many of the costs of recurrent and terminal disease are included.

In absolute terms, screening detected cancers are the cheapest to treat. Cases presenting symptomatically in the group offered a screening test, either following a negative test or having refused a test, are the most expensive groups. However, the standard deviations are so large that none of the differences between groups are statistically significant.

Why is there no significant difference between the groups in terms of treatment costs? To identify the factors which determine treatment costs a regression analysis was performed on the combined data sample. A total of 85 observations were excluded because the relevant data were not complete: these cases tended to be those who do not undergo surgery, resulting in uncertainty over the site, size and stage of the primary tumour.

Figure Five - Cost of treatment within group offered screen



The initial equation attempted to explain the variation in the total cost per patient and included all potentially relevant variables. The coefficient on the following variables did not achieve statistical significance:

- screening group (control/study group)
- Dukes' stage
- primary site within bowel
- number of primary tumours
- type of initial admission (elective/emergency)
- wound infection (none/minor/major)
- wound dehiscence (none/minor/major)

The remaining variables were used to re-estimate the equation, with the following results, all the coefficients being significant at the 5% level:

Total cost =

- £541	(Constant)
+ £41	Age (years) (t-statistic = 2.17)
+ £818	Sex (0=male, 1=female) (t = 2.95)
+ £1,199	Initial operation involved colostomy (0 = no, 1 = yes) (t = 3.73)
+ £1,079	Number of complications (excluding wound) (t = 4.96)
+ £797	Died during hospital stay (0 = no, 1 = yes) (t = 2.14)
+ £982	Readmitted following initial treatment (t = 3.65)
+ £1,322	Number of subsequent admissions (t = 11.56)

The equation explains 51% of the variation in total cost per patient, according to the R-squared measure.

How should the coefficients be interpreted?

- Age and sex appear to be standing as proxies for the general physical condition of the patient and possibly for the number of concurrent conditions. It may also be harder to discharge elderly patients if they live alone in unsuitable accommodation. There appears to be no correlation between age and stage of disease at diagnosis.
- An operation requiring a colostomy can be an indication of a palliative procedure, implying a poor general condition, but even curative resection involving a colostomy may have prolonged post-operative stay as the patient comes to terms with and learns to manage their stoma. A colostomy can also be used as a temporary measure to cover an anastamosis, implying a subsequent readmission for closure.
- A patient with several complications (e.g. cardiovascular, respiratory) will be in a poor condition and unsuitable for discharge until they are resolved. The number of non-wound complications is a subjective element since it is not based on an objective clinical standard. Instead, it is coded according to the discharge letter to the patient's GP and simply counts the number of complications mentioned.
- The size and sign of the coefficient on the variable coded for death in hospital implies that terminal care is often prolonged rather than a short sudden post-operative death from complications. This variable may also

be correlated with the number of subsequent admissions. Clearly, this may be correlated with both the age of the patient and the number of complications.

- The readmission rate following the initial operation can be either to complete a two-stage operation or as an emergency; in either case the costs will be higher.
- Subsequent hospital admissions may be either for routine surveillance (e.g. of an excised rectal tumour) but may be to treat recurrence or for terminal care.

Do these variables help to explain the higher costs in the group refusing an offer of screening (and, to a lesser extent, among interval cases)?

- It has already been shown that those refusing screening are the oldest age group on average, being three years older than the other sub-sets of the study group and one year older than controls.
- The proportion of women in each group is similar, although the difference between screening detected cases (40% women) and interval cases (46% women) is more marked.
- There is a difference between groups with regard to the proportion requiring a colostomy during their initial treatment. Refusers and the control group have a higher proportion of operations requiring a stoma than do cases which accept screening (29% and 28% for refusers and controls, compared to 19% and 13% for intervals and screening detected cases).
- The average number of complications for the control group was 0.23. Figures for screening detected and interval cases in the study group were much lower (0.17 and 0.15 respectively), but the average number per refuser case was much higher at 0.31.
- The figures for deaths in hospital (during any admission) are similar for refusers and controls (30% and 25% respectively) but the interval case fatality rate is also high at 23% while the rate for screening detected cases is low (9%).
- The readmission rate is consistently higher in the study group than in the control group. The proportion of controls being readmitted was 16% while the figures for the screening detected, interval and refuser groups were 21%, 23% and 22% respectively.
- The number of subsequent admissions was also higher in the study group, although the average number among screening positive cases (at 0.58) was comparable with that of the control group (at 0.61). The interval and refuser groups both had an average of 0.85 subsequent admissions.

In summary, cases in the group refusing a screening test are more expensive because the people are older when they are diagnosed, are more likely to require a colostomy (as a palliative procedure), suffer complications to other body systems, require subsequent hospital care and are more likely to have prolonged admissions for terminal care.

EFFECTS ON GP WORKLOAD

The above analysis has concentrated on the impact of screening on use of hospital resources by changing treatment patterns; another potentially significant impact is on the workload of GPs. Unfortunately, it is a characteristic of the primary health care sector that there is little routinely available data on activity levels to use to make such calculations. However, the national census of GP workload conducted every decade indicates that approximately a quarter of a million GP consultations each year relate to cancer of the digestive tract (Royal College of General Practitioners (1986)); this implies a total cost of just over £1 million using a standardised value of GP time (Hughes (1991)).

It has been established that a postal invitation to complete an FOB test has no significant impact on GPs' workload (Pye et al. (1988)). Enquiries related to the appropriate diet, where to return the test to, and why the individual was singled out for testing. The receptionist had been briefed before hand and was able to deal with all of these problems. There are similar findings for breast cancer screening (Ashby et al. (1990)) and education in breast self-examination (Nicholls et al. (1982)). Note, however, that this does not include the longer term effects of prompt cards being placed in the notes of screening refusers, as described in Chapter Six.

To assess the impact of screening on those with neoplasia, a study of the records of two GP practices was undertaken. Patients were identified from MRC trial records and, with the permission of the GPs, their notes were studied. Six cases of cancer in the control group generated a total of 42 consultations, whereas nine cases in the test group resulted in 22 consultations (only two of these were among the three screening detected cases). Three patients found to have adenomas in the control group generated four consultations, while 15 test group patients made ten consultations. Given the number of cases studied any conclusions drawn must be heavily qualified; however, there appears to be a shift from consultations for patients with advanced cancer in the control group to more reassurance for patients with screening detected adenomas in the screened group. Overall there is a fall in the number of consultations, but this is not large in the context of the weekly workload of the two practices.

In summary, the findings for primary care mirror those for the acute sector: screening does not result in large-scale savings in terms of GPs' time.

CONCLUSION

It is possible to estimate the resource use of a sample of patients undergoing surgery from their case notes, although care should be taken to avoid over-elaboration in measuring and valuing resource use.

It is too early to assess the full impact of screening on treatment costs in the MRC trial. At this stage, however, screening-detected cases are cheaper to treat than cases detected by any other means but there are two problems. Firstly, there is an excess of cases in the group offered screening, partly reflecting the earlier detection of cases, but also attributable to over-diagnosis on screening. Secondly, cases presenting symptomatically in the study group are very expensive to treat, especially those who refuse screening. This is not attributable to any one factor, but to a generally unfavourable set of variables for these groups. The net effect is that no treatment cost saving can be expected from screening, at lest in the early years of the operation of a programme.

Chapter Eleven

COMPARING SERVICES FOR COLORECTAL CANCER

Introduction

The technical efficiency of colorectal cancer screening, in terms of the cost per cancer detected of various protocols, has been thoroughly explored in the chapters above. Without trial data to convert this into health gain, the allocative efficiency of screening cannot be assessed. As an interim step this chapter draws together the evidence on various treatments and interventions for colorectal cancer alone. Given that cancer treatment is likely to remain a high priority treatment it seems reasonable to suppose that the policy issue faced by health service purchasers is *how* to treat a particular form of cancer rather than *whether* to treat. No previous research has tried to draw together the evidence in this way.

Service options

The main areas in which services for colorectal cancer could be extended are:

- primary prevention, by reducing exposure to factors which increase the risk of developing the disease;
- screening various groups at increased risk of having asymptomatic or pre-malignant disease;
- treating symptomatic disease, particularly treatments which complement potentially curative surgery.

From these categories a list of options to compare can be drawn up. These include:

- faecal occult blood (FOB) testing of the general population aged 50-74;
- flexible sigmoidoscopy screening of the general population on a one-off basis between the ages of 55 and 60;
- systematic follow-up of adenoma patients by endoscopy;
- radiotherapy following potentially curative surgical resection of rectal cancer;
- chemotherapy following potentially curative resection of stage C cancer of the colon; and
- surgical resection of hepatic liver metastases for advanced disease (subject to tightly defined criteria).

Primary prevention is not considered because of the uncertainty regarding the causes of colorectal cancer, specifically the interaction between environmental and genetic components. While estimates of the effects exist, the cost involved (and the feasibility of such a dietary change) and the timing of the benefits are very uncertain.

The other main area excluded from the analysis is the follow-up of patients who have had a supposedly curative resection with aggressive treatment of

recurrence (as opposed to palliation). Despite the lack of evidence of proof of benefit from randomised trials this practice is widely established, albeit with considerable inter-surgeon variation in protocol. The problem is in two stages: is it worth screening asymptomatic people for recurrence, and is it worth treating recurrence once it has been identified? The second issue is harder to tackle than the first since it requires a precise estimate of the numbers who will benefit plus the extent of the benefit; surgical series tend to confine their comments to the operative mortality rate. An additional problem for evaluation in this area is that much of the benefit of follow-up may come from the reassurance provided merely by 'going through the motions' of monitoring progress. Other than noting the need for better evidence in this area, follow-up of curative resections is thus not considered further.

The evidence on the health service costs and health benefits of each of the options is now considered in turn in order to produce comparable data.

1. FOB testing of asymptomatic 50-74 year olds

Survival data from the MRC trail will not be available for several years; results cannot be presented or discussed in advance of the first official report. The only randomised trial evidence available to date relates to screening with the Haemoccult test using rehydration prior to development. On this basis colorectal cancer mortality is reduced by 30% with annual screening and by 10% for two-yearly screening (Mandel et al. (1993)). An American case-control study also claimed a mortality reduction of 31% for those completing an FOB test although the 'protective' effect lasted for less than three years (Selby et al. (1993)). However, doubt has subsequently been cast on the validity of these figures (Simon (1993)). In this country, one group of experts concluded that the mortality reduction form screening was likely to be less than 20% (Chamberlain and Miller (1988)); even this seems optimistic in the light of the Minnesota findings of a 10% reduction even with a more sensitive test.

Where will this benefit from screening be observed? Allowing for lead-time effects, there should be little reduction in mortality before the age of 55, given that screening commences at 50. Similarly, it is assumed that there will be no mortality reduction after the age of 80, given that screening stops at age 75. The figures quoted in Table One of Chapter Two can be used to estimate the life-years gained. If the mortality reduction is 20% then 2,092 lives and 25,686 life-years would be saved. Reductions of 10% and 5% are also considered. While they are small the health losses attributable to investigations and surgery should also be calculated for completeness.

The UK population in the target age range is 12 million people (OPCS (1993)), half of whom are screened each year. The 'steady-state' cost of screening 100,000 people given in Chapter Four was £184,324, the figure for the third round. However, those in the youngest age groups will be screened for the first time at a higher cost (£381,105 in Chapter Four). The cost of a

screening round over two years would, therefore, be £23,060,640, or £11,530,320 per annum. Average costs are shown in Table One:

Table One

Mortality reduction	£ per life saved	£ per LY saved
20%	5,512	449
10%	11,023	898
5%	22,047	1,796

Note that cases detected and costs incurred are not the only effects of screening: as noted in Chapter Four, people will test falsely positive, while others will suffer complications following investigation.

Chapter Ten showed no difference in treatment cost per case as a result of screening, but 37% more cases were detected in the group offered screening. Some of these are expected to represent the early detection of asymptomatic cases, but a proportion will be due to detection of 'lifetime-latent' cases. Making the arbitrary assumption that 50% of the excess cases are lifetime latents, and costing these cases as if they were screening detected cases (the cheapest to treat) implies an extra cost of £8,253,361. This would add 72% to the above results.

While, Chapter Seven demonstrated the potential cost savings from adenoma excision, a mortality reduction from colorectal cancer of 10% or less implies that lower estimates of the numbers of cancer prevented by adenoma excision is appropriate. This would imply that the cost savings would be lower than the figure given in the final section of that chapter.

2. *FOB testing of first degree relatives of patients diagnosed as having cancer*

There have been no randomised trials to establish the efficacy of screening relatives of colorectal cancer patients for the disease. The rationale for screening is based on the observed higher incidence and mortality in this group (Dunlop (1992)); the effect is particularly marked for first-degree relatives (children, siblings, parents). Some research indicates that the risk is highest for relatives of a newly diagnosed cases in younger people (aged 45 or under).

The following estimates are based on the data of Houlston et al. (1990). In England and Wales, 664 cases of colorectal cancer occurred in patients aged less than 45, generating 2,445 first degree relatives (60% parents, 40% siblings). Regular colonoscopy screening could prevent 244 cases. These people would be colonoscoped every five years, except for the 11% in whom adenomas are detected, requiring three-yearly follow-up.

Thus, 1,007 siblings with an average life-expectancy of 40 years (assuming them to be aged 40 on average) and 1,438 parents with a life-expectancy of

15 years (assume average age 65) would require surveillance. Of the 269 found to have adenomas, an even distribution between these two groups is assumed. Multiplying out gives a total of 13,241 colonoscopies in all at £105.10 each (from Chapter Four).

Of the 244 cases detected only 65% represent lives saved, since a proportion of these people would have undergone a curative resection even if presenting symptomatically; only 159 lives will be saved. The cost per life saved is thus £8,752. The calculation of the number of life-years saved is more problematic. The clinical data give the total number of cases that might be detected but give no idea of when these will occur. There is also no information on the distribution of cases between the younger and older age groups, yet this will have an important bearing on life-expectancy.

While there are costs to identifying, contacting and counselling the relatives, these are unlikely to be great relative to the costs of investigation. Potentially more serious is the problem of getting relatives to undergo surveillance. The figures above assume full participation, yet one study found that only 30% of those identified accepted an offer of screening (Armitage et al. (1986)). While this restricts the impact of the screening programme on the disease it also reduces the cost, hence for the current purpose the average cost per life saved is hardly affected. The calculations also assume no net effects on treatment costs.

3. Flexible sigmoidoscopy for age group of general population

There are no 'gold standard' data from randomised trials to support such a programme, but some researchers have enthusiastically advocated implementation in this country (Atkin et al. (1993)). The basis for their calculations is an American case-control study of rigid sigmoidoscopy screening which calculated that mortality was reduced by 60% within the range of the sigmoidoscope used. If this was the case, then the longer flexible instrument could reduce total mortality by 30% (Selby et al. (1992)). Supporting evidence has been provided by a similar American study (Newcomb et al. (1992)).

The flaws in the calculations of Atkin et al. (1993) have been pointed out in Chapter Five, where the option was shown to be inferior to some FOB screening protocols. While it was more expensive than unhydrated three-day testing (option 1 above), however, it was also more accurate, hence it is included here.

Transferring the data from Chapter Five to a national population means 600,000 people aged 57 would be screened each year. The costs of invitation plus investigation, including colonoscopy, give a total cost of £15,148,458. The number of cases of disease detected would be 804, but 35% of these would be cured in any event; 523 people thus benefit from screening. The cost per life saved is £28,965.

4. Follow-up of patients with a previous adenoma excision

The potential significance of the costs of adenoma follow-up are demonstrated by the finding of one evaluation that 25% of the total costs of a screening programme were due to adenoma follow-up alone (Wagner et al. (1991)). However, there is no randomised trial evidence on the health gain of adenoma follow-up by any protocol.

One study calculated that 226 colonoscopic investigations would be necessary to save one life (Ransohoff et al. (1991)). This was based on 5-yearly colonoscopy follow-up of a 50-year old who has had an adenoma excised; this protocol averts 75% of the colorectal cancer deaths which would have resulted from recurrence. If colonoscopy costs £105 the cost per life saved is £23,753. The results in the original study were sensitive to the assumption of 75% effectiveness and the cumulative risk of colorectal mortality after polypectomy of 2.5%. If the true risk is half of this figure then the cost per life saved is doubled. The study concluded that in low risk cases 3-yearly follow-up (as widely recommended in America) is "excessively costly".

5. Radiotherapy for rectal cancer

While a randomised trial of pre- and post-operative radiotherapy as a complement to curative surgical resection for rectal cancer is in progress, no details are yet available.

However, a course of out-patient radiotherapy of 20 fractions for stage B and stage C rectal cancer will result in an extra 10% of patients being cured (Dr. Fearon, consultant oncologist, Western General Hospital, Edinburgh, personal communication). About 10% of patients suffer some adverse reaction to therapy and require a short in-patient admission for monitoring: an average of five days in hospital is assumed.

On the basis of figures in Chapter One, there 10,581 cases of rectal cancer are diagnosed each year, with 66% being stages B or C. A radiotherapy fraction costs £70 (Fearon, as above) and an in-patient day costs £116.66 (hotel costs plus other overheads) as in Chapter Ten. The total cost for 6,983 patients is £10,183,343. Given the survival data at five years quoted in Chapter One 2,870 would be cured by resection alone; an extra 698 lives are saved as a result. The average cost per life saved is £14,589.

6. Chemotherapy for stage C colon cancer

An American randomised trial implies that chemotherapy for stage C colon cancer reduces colorectal cancer mortality rate in that group by 10% (Moertel et al. (1990)).

Based on the figures in Chapter One, 4,428 stage C colon cancers present each year. A six-month course of chemotherapy on an out-patient basis

costs £1,600 (Fearon, as above), but trial data are based on a twelve-month course of treatment hence the cost is doubled. Side-effects of chemotherapy are generally more serious: the American trial found 30% of patients experienced significant side-effects and it is assumed that all of these require admission. The same cost and length-of-stay assumptions are used as for radiotherapy above.

The total cost of treatment is £14,944,455. Of these people, 22% would be cured without chemotherapy on the basis of data in Chapter One. An extra 10% are cured as a result of chemotherapy, so 443 lives are saved at an average cost of £33,735 each.

7. Aggressive surgical and chemotherapeutic treatment of hepatic metastases

There is no evidence of benefit for this intervention from randomised controlled trials. However, expert opinion suggests that 5% of cases with metastases could be cured as a result of resection if guidelines for case selection were carefully adhered to.

From Chapter One 6,714 people with advanced disease present each year. The mortality rate following initial admission is 31% (Umpleby et al. (1984)); these people are not considered further. The 4,632 survivors are assumed to undergo a CT scan to establish the extent of their disease: this costs £193 (Hutton (1988) updated). Only 5% will be eligible i.e. 232 people. Based on observations in University Hospital, Nottingham, in the course of the screening study, the operation lasts four hours with in-patient stay of 15 days. Using the cost figures above, with theatre time valued at £4.47 per minute (as in Chapter Ten), the total cost is £1,548,888. If 20% of those undergoing the operation are cured as a result (Fearon, as above) then 46 lives are saved; the cost per life saved is £33,671. Once again, the losses to operative mortality will be small but should be calculated for completeness.

Summary

The above results can be drawn together in Table Two:

Table Two

Intervention	Cost per life saved (£)
Screening first degree relatives	8,752
FOB screening (10% mortality reduction)	11,023
Radiotherapy for rectal cancer	14,589
FOB screening (5% mortality reduction)	22,047
Flexible sigmoidoscopy screening	23,965
Adenoma follow-up	23,753
Resection of liver metastases	33,671
Chemotherapy for colon cancer	33,735

Discussion

These results have a number of weaknesses. Many assumptions are based on poor quality data, particularly those relating early detection to health gain. The effectiveness measure used is the number of lives saved as a result of treatment, which is unsatisfactory since it gives no indication of the length of life thus gained: many people would not regard saving the life of a 40-year old as having the same value as saving the life of an 80-year old, for example. Aside from these extreme cases the age group affected is relatively homogeneous, hence a comparison on this basis has some value. To the extent that this is important it would tend to reinforce the ranking since a younger age group will tend to benefit from screening, especially if the screening of first degree relatives focuses upon siblings in preference to parents. However, the measure also excluded aspects which affect health status such as post-operative complications and the psychological impact of screening.

Potential users of the results must make a judgement about the quality of the data. Given the lack of reliable epidemiological evidence, are the assumptions used good enough to base purchasing decisions upon? How conclusive does evidence have to be before it can be considered proof? FOB screening provides a good example: if a 10% reduction in mortality can be achieved as a result then it represents good value for money, yet a more pessimistic (but possibly more realistic) 5% reduction casts serious doubt on this use of resources.

One important element excluded from the calculations by virtue of the effectiveness measure used is the health status of patients with, during, and without treatment. One thorough Dutch study of the breast cancer screening literature concluded that quality-adjustment of life-year gains made little difference to the results i.e. rankings based upon life-year gains alone were unaffected (de Haes et al. (1991)). Can this finding be generalised to colorectal cancer? This is considered under a number of headings.

- The main concern regarding quality of life in colorectal cancer treatment relates to adjustment to a colostomy. A study of the post-operative life of such patients found them to be 'normal' on a variety of indicators although they are more prone to depression and anxiety (MacDonald and Anderson (1985)). When asked to value their quality of life, a sample of people with a colostomy gave values of 80 to 90% of normal (Boyd et al. (1990)).

- The ability to return to work is another indicator of post-operative health. In the MRC study found that of 18 people in paid employment previous to their operation, 11 had decided to return and 7 had decided to retire; many of this latter group were close to retirement age already. Similar findings are reported for patients requiring an ileostomy for inflammatory bowel disease; while more reported lethargy, the operation is more radical than many used to treat cancer (Wyke et al. (1988)).

- The quality of life of people with terminal colorectal cancer has not been widely studied: the only available data showed that patients with a life expectancy of a few months rated their own health at 56% of normal (McGowan et al. (1989)).

In summary the evidence for colorectal cancer appears to broadly confirm the Dutch finding of the limited quality of life implications of screening. Only during terminal disease are the effects of the disease likely to be so unpleasant as to be important. As the resource use patterns in Chapter Ten indicate, the terminal stages of the disease are mercifully brief for many patients. This implies that lives saved are a reasonable basis for comparing the benefits of alternative interventions.

CONCLUSION

Given the poor quality of the data used, many of the conclusions must be heavily qualified. While this may hamper choices between two interventions with small differences in cost-effectiveness ratios, the differences between screening relatives of young cancer patients and performing surgical resections on liver metastases is striking and cannot be explained by erroneous assumptions alone. If care is taken over interpretation, the results can be a guide to economic efficiency for purchasers in this field.

CONCLUSION

The conclusions of this evaluation can be divided into those relating to policy and those relating to methodology.

Policy conclusions

The questions posed at the start of this evaluation were as follows:

i) what are the costs and benefits to NHS purchasers of screening for colorectal cancer in an asymptomatic population aged 50-74 by an offer of Haemoccult II testing every two years?
ii) is this the most efficient way to screen for colorectal cancer?
iii) is this the most efficient way to reduce the health loss of colorectal cancer?
iv) can such a policy be justified in comparison with other uses of health service resources?

The first question has been considered in detail, first by a simple model (Chapter Four) and then by detailed analysis of aspects which may affect these results. The findings of the relevant chapters are considered below.

Chapter Four established the cost per case detected by three-day Haemoccult screening using the MRC trial protocol as being £3,347 at the initial round of screening and £2,989 in the 'steady-state' round. Given the likely order-of-magnitude of health gains resulting from screening this implies that FOB screening is worth considering in more detail: it cannot be ruled out as an efficient use of resources on the basis of this data alone. However, the results are sensitive to the cancer detection rate of the test and to the participation rate, and these were noted as areas requiring further investigation.

Participation is one of the key variables in determining the clinical effectiveness and the economic efficiency of a programme. Chapter Six showed that, given the costs and the number of extra responders, the MRC trial protocol is justified on economic grounds, although this remains subject to the relationship between numbers screened and health benefits. Some methods suggested for distributing FOB tests such as a GP-based system are expensive by comparison; the same is true of pursuing persistent non-participants in postal-based FOB screening.

Given the importance of the detection rate, the significance of the adenoma to cancer sequence was then assessed. Unfortunately, this is probably the area in which existing knowledge is weakest: this was fully revealed by a comparison of methods of calculating the preventive effects of adenoma excision. On the basis of literature estimates, between 5 and 86% of adenomas excised on screening may have become cancers within the lifetime of the patients. Previous evaluations have shown this to be a crucial

variable in determining the efficiency of screening; this finding is likely to exacerbate the problem. If the reduction in mortality from colorectal cancer screening is less than 10%, as seems likely on the basis of the American data, then it is possible that none of the model presented is correct: all are too optimistic about the eventual impact of adenoma excision. The predictions of large savings from excision can only be squared with the modest mortality reduction if either the cancers prevented from excision would have been easily cured in any event or the Minnesota trial has insufficient follow-up to capture any effect.

The detection rate and participation rates can also be influenced by the age group chosen for screening. Chapter Eight considered the effects of varying the screening protocol for various age ranges, including extensions to younger and older age groups. The rationale for screening the younger age group would be to excise adenomas but, given the uncertainty noted above, this is difficult to evaluate. Nevertheless, trial data indicate the yield is very low in this group. Screening older people gives a high yield but there are doubts about their ability to undergo surgery if a resection is required; nevertheless, there is some evidence that post-operative mortality in this age group is falling over time. In terms of the 50-74 year age group, two-yearly screening appears justified for the older group but not for the younger part of the population since detection rates fall to zero. This suggests a one-off screen for younger people at some point in their fifties with two-yearly screening commencing at about 65 years of age. Unfortunately, the data are not available to evaluate such a protocol

The calculations so far have accepted the MRC trial protocol for the follow-up of positive screening results, but Chapter Nine explores the efficiency of these aspects. The retesting of initial FOB positives observing dietary restrictions is shown to have minimal effect on yield while reducing the costs of diagnostic investigations quite considerably. Trial data indicate that the barium enema X-ray alone is insufficiently sensitive to small cancers in the sigmoid colon to investigate a positive FOB test. The addition of 60cm. flexible sigmoidoscopy makes an attractive combined test which is comparable to colonoscopy as the test of choice in most respects. However, colonoscopy is shown to fractionally more cost-effective so long as the investigation is complete first time in at least 70% of cases. This is feasible in the hands of a well-trained operator, but is unlikely to be the case for unsupervised junior doctors.

The final factor to consider is the impact of screening on the costs of treating colorectal cancer. It is possible that the trial is at too early a stage to consider this question fully, since screening has the initial effect of clearing out all prevalent cases; this would imply offsetting cost savings occurring at some point in the future. Considering only cases with at least three years of follow-up shows that the total cost of treating the group offered screening is 47% higher than the cost for the control group. This is mainly explained by

the excess of cases in this group: there is no significant difference between average costs in each group. In general, there is much more variation within the study and control groups in terms of treatment costs than there is between them. An attempt to explain variation in cost per case (for any screening group) indicated a range of factors as having some explanatory power. These include the age and gender of the patient, whether the initial operation involved a colostomy, the number of complications in post-operative recovery (including whether the patient died before being discharged), plus the need for a second (and subsequent) admissions. These offer at least a partial explanation of the higher costs of study group cases and particularly of those who refuse an offer of screening and are subsequently diagnosed as having cancer.

The second question related to the best test to use in colorectal cancer screening. To address this the model constructed in Chapter Four was used. Chapter Five compared Haemoccult screening with a variety of other screening tests for colorectal cancer; a set of options which dominated the others in the sense of having higher yield and lower costs was identified. This included FOB screening on the MRC protocol, but also FOB testing with rehydration of samples, the Hemeselect immunological test and colonoscopy. While colonoscopy can probably be ruled out as infeasible given the availability of resources in this country, the decision rule of cost per case does not allow purchasers to choose between the remaining options. To do so, further information on the health benefits of detecting a case and the willingness of purchasers' to pay for extra health benefits is required.

The third study question was considered in Chapter Eleven, with an evaluation of various ways of reducing the health cost of colorectal cancer. The available evidence for a range of interventions was assessed and found to be of generally poor quality. 'Order-of-magnitude' estimates of the cost per life saved were made for each option identified. Caution must be used in interpreting and using the results given the quality of the data used, but some broad conclusions can be drawn. For example, screening the relatives of people aged less than 45 who are diagnosed as having colorectal cancer should be implemented before surgical resection of liver metastases in advanced disease is considered. Better data are urgently required in order to place more certainty on results.

The same comment applies to addressing the final question, with the added requirement that some common measure of outcome be developed for comparing widely differing programmes in terms of their costs and benefits. A drawback in both instances is the lack of data from the MRC trial. Much of the research carried out for this work was funded by the MRC, who request that trial results with respect to survival are not analysed in advance of the publication of the first set of survival figures. This has enforced the present evaluation to a (possibly excessive) reliance on interim measures of the output of screening such as cases detected.

What advice can currently be offered to purchasers? In the light of the evidence from the Minnesota trial the MRC trial is unlikely to show a large mortality difference, especially in its first mortality report due in the mid-1990s. This implies that health gains from the detection of cancer at an asymptomatic stage are likely to be modest. The potential gains from adenoma excision are potentially larger but are also less certain; considerable patience will be required from all concerned in order to assess this effect, if any, since about 15 years of follow-up will be required. The net cost of screening is also uncertain at this stage of the MRC trial. The survival gains necessary to justify the costs of screening and diagnosis are not implausible. The crucial issue on this side of the 'cost-benefit equation' is the net effect on treatment costs. Screening increases the total cost in the early years, as would be expected, but adenoma excision could potentially outweigh this effect in the long-term, even with discounting to a net present value.

In summary, the policy conclusion is rather disappointing: it is simply too early to say. On a more optimistic note, the crucial pieces of missing data can eventually be observed from the MRC trial: will we have the patience to wait that long?

Methodological conclusions

Apart from the conclusions relating to the policy questions posed, there are a number of implications of this study for the evaluation of mass screening.

One problem rarely addressed in the literature relates to the timing of the evaluation. A central requirement of information for purchasers is that information is available promptly to make a decision. In this case the results of the evaluation must be available soon after evidence on efficacy if efficiency is to be considered. Even a limited evaluation such as this can take a considerable amount of time, however, hence the need to commence the work *before* efficacy data are available. There is a risk, however, that the economics research will be in vain if the treatment has no benefit over the alternative. In addition there are great difficulties in coping with the uncertainty regarding the magnitude of health gains as a result of screening.

More generally, the role of clinical trials lasting 15 to 20 years is at issue. Such trials provide 'gold standard' data of the type needed to evaluate many of the competing claims for scarce resources; however, the data arising often relate only to one test and to one protocol. Evidence from other sources comes to light while the trial is in progress, placing pressure upon purchasers to act before the results become available. Recent case-control study results and the support achieved by advocates of endoscopic screening are examples of this. A related problem is that by the time the efficacy (or otherwise) of the screening test is established new and better tests will be available. Clinical trials of preventative interventions may need rethinking,

possibly using the mathematical models of the disease process described in Chapter Three. Trials would then be designed specifically to address gaps in the current medical knowledge. Such models are often so sophisticated in terms of their statistical structure as to be inaccessible to the non-specialist, fostering suspicion of their results. This complexity is necessary, however, to address the screening problem.

Other methodological developments will be of interest to economists carrying out evaluations. The level of detail required in the costing of hospital procedures is an example. Chapter Ten demonstrated that individual patient-based costing of nursing time is unlikely to add greatly to the accuracy of average nursing costs on a surgical ward. This is of particular interest as hospitals develop individual patient-billing systems as part of the move towards disaggregated contracts in the internal market. Much effort may well be expanded on this, as on other aspects of costing, in the unthinking pursuit of yet greater precision.

One particular area of screening evaluation advanced here is the development of methods for determining the economic optimum participation rate. It is interesting to note that in this, as in other aspects of the evaluation, the MRC trial protocol for screening the study group does very well in comparison with the alternatives: the choice of screening test, the retesting protocol for positive FOB results and the choice of diagnostic work-up all appear technically efficient. The only variation uncovered by the economic evaluation would be in terms of varying the frequency of screening for different age groups. The trial protocol was devised to minimise research funding requirements: one interpretation, therefore, would be that medical staff are capable of devising efficient practice protocols when explicitly faced with the constraint of limited resources. It would interesting to gather experience from other trials to find whether this is generally true.

The model has identified many of the 'plethora' of margins referred to elsewhere (Shackley and Cairns (1993)), but this is often at the expense of sacrificing some of the (spurious) certainty which other evaluations convey. This certainty often results from the use of simplistic decision criteria such as preferring the option with the lowest cost per case detected. The correct comparison is in terms of the additional cost per additional unit of output or outcome, whether this is over and above doing nothing or in comparison with an alternative intervention. This principle has been applied in the evaluation of colorectal cancer screening wherever possible.

REFERENCES

Abrams J. 'A second look at colonoscopy' Archives of Surgery 1982 117 913-917.

Adams M., McCall N., Gray D., Orza M. and Chalmers T. 'Economic analysis in randomized control trials' Medical Care 1992 30 231-243.

Adamsen S. and Kronborg O. 'Acceptability and compliance in screening for colorectal cancer with the faecal occult blood test' Scandinavian Journal of Gastroenterology 1984 19 531-534.

Agrez M., Evans D. and Duggan J. 'Faecal occult blood testing for colorectal cancer: luxury or necessity?' Australian and New Zealand Journal of Surgery 1990 60 451-454.

Ahlquist D. 'Occult blood screening: obstacles to effectiveness' Cancer 1992 70 Supplement 1259-1265.

Ahlquist D., McGill D., Fleming J., Schwartz S., Wieand S., Rubin J. and Moertel C. 'Patterns of occult bleeding in asymptomatic colorectal cancer' Cancer 1989 63 1826-1830

Ahlquist D., McGill D., Schwartz S., Taylor W. and Owen R. 'Fecal blood levels in health and disease: a study using Hemoquant' New England Journal of Medicine 1985 312 1422-1428

Aldridge M. and Sim A. 'Colonoscopy findings in symptomatic patients without X-ray evidence of neoplastic disease' Lancet 1986 ii 833-834.

Allison J. and Feldman R. 'Cost benefits of Hemoccult screening for colorectal carcinoma' Digestive Diseases and Sciences 1985 30 860-865.

Allison J. and Feldman R. 'Hemoccult sensitivity in detecting colorectal neoplasm: sensitivity, specificity and predictive value' Annals of Internal Medicine 1990 112 328-333.

American Cancer Society 'Cancer of the colon and rectum: summary of a public attitude survey' Ca 1983 33 359-365.

Anderson J., Hole D. and McArdle C. 'Elective versus emergency surgery for patients with colorectal cancer' British Journal of Surgery 1992 79 706-709.

Anon 'Molecular secrets of colorectal cancer' Lancet 1991 338 1363-1364

Anon 'Colorectal cancer: new evidence for the adenoma/carcinoma sequence' Lancet 1992 340 210-211.

Applegate W. and Spector M. 'Colorectal cancer screening' Journal of Community Health 1981 7 138-151.

Arabi Y., Hamilton J. and Keighley M. 'Factors influencing the cost of elective colorectal operations' Journal of the Royal College of Surgeons of Edinburgh 1980 25 279-285.

Armitage N. 'Intervention studies in adenoma patients' World Journal of Surgery 1991 15 29-34.

Armitage N., Farrands P., Mangham C and Hardcastle J. 'Faecal occult blood screening of first degree relatives of patients with colorectal cancer' International Journal of Colorectal Disease 1986 1 248-250.

Ashby J., Buxton M. and Gravelle H. 'Will a breast screening programme change the workload and referral patterns of general practitioners?' Journal of Epidemiology and Community Health 1990 44 36-38.

Ashworth T. 'Breast screening: a response to Dr Maureen Roberts' British Medical Journal (letters) 1989 299 1337.

Atkin W., Morson B and Cuzick J. 'Long-term risk of colorectal cancer after excision of rectosigmoid adenomas' New England Journal of Medicine 1992 326 658-662.

Atkin W., Cuzick J., Northover J. and Whynes D. 'Prevention of colorectal cancer by once-only sigmoidoscopy' Lancet 1993 341 736-740.

Bagust A. 'Dispel that old myth' Health Service Journal 1990 July 5th p.1000-1001.

Baker M., Kessler L and Smucker R. 'Site-specific treatment costs for cancer: an analysis of the Medicare Continuous History Sample File' in R. Scheffler and N. Andrews "Cancer Care and Costs: DRGs and Beyond" (Health Administration Press, Ann Arbor, Michigan 1989).

Bech K., Kronborg O. and Fenger C. 'Adenomas and hyperplastic polyps in screening studies' World Journal of Surgery 1991 15 7-13.

Bingham S. 'Diet and large bowel cancer' Journal of the Royal Society of Medicine 1990 83 420-2.

Black D. and Pole J. 'Priorities in biomedical research: indices of burden' British Journal of Preventive and Social Medicine 1975 29 222-7

Blades C., Culyer A. and Walker A. 'Health service efficiency: appraising the appraisers. A critical review of economic appraisal in practice.' Social Science and Medicine 1987 25 461-472.

Blalock S., de Vellis B., Sandler R. 'Participation in fecal occult blood screening: a critical review' Preventive Medicine 1987 16 9-18.

Bolt R. 'Sigmoidoscopy in detection and diagnosis in the asymptomatic individual' Cancer 1971 28 121-122.

Boulos P., Cowin A., Karamanolis D. and Clark C. 'Diverticula, neoplasia, or both? Early detection of carcinoma in sigmoid diverticular disease' Annals of Surgery 1985 202 607-609.

Bowling A. and Jacobsen B. 'Screening: the inadequacy of population registers' British Medical Journal 1989 298 545-546.

Box V., Nichols S. Lallemand R, Pearson P. and Vakil P. 'Haemoccult compliance rates and reasons for non-compliance' Public Health 1984 98 16-25.

Boyd N., Sutherland H., Heasman K., Tritchler D and Cummings B. 'Whose utilities for decision analysis?' Medical Decision Making 1990 10 58-67.

Brooks R., Brown M., Marsh J. and Woodbury J. 'Costs of managing patients in a Canadian rheumatic diseases unit' Journal of Rheumatology 1981 8 937-948.

Brown K. and Burrows C. 'The sixth stool guaiac: $47 million that never was' Journal of Health Economics 1990 9 429-445.

Brown M. 'The national economic burden of cancer: an update' Journal of the National Cancer Institute 1990 82 1811-1814.

Brown S., Abraham J., Walsh S. and Sykes P. 'Risk factors and operative mortality in surgery for colorectal cancer' Annals of the Royal College of Surgeons of England 1991 73 262-272.

Buxton M. 'Problems in the appraisal of new health technology: the evaluation of heart transplants in the UK in M. Drummond (ed.) "Economic Appraisal of Health Technology in the European Community" (1987).

Buxton M., Acheson R., Caine N., Gibson S. and O'Brien B. "Costs and Benefits of the Heart Transplant Programmes at Harefield and Papworth Hospitals" (1985).

Cairns J. and Shackley P. 'Sometimes sensitive, seldom specific' Health Economics 1993 2 43-54.

Canon-Albright L., Skolnick M., Bishop T., Lee R. and Burt R. 'Common inheritance of colonic adenomatous polyps and associated colorectal cancers' New England Journal of Medicine 1988 319 533-537.

Castiglione G. and Ciatto S. 'Selection criteria in colorectal screening' Tumori 1988 74 451-456.

Center for Disease Control 'Trends in colorectal cancer incidence - United States, 1973-1986' Journal of the American Medical Association 1989 262 2664-2665.

Chamberlain J. and Miller A. (eds.) "Screening for Gastrointestinal Cancer" UICC, Hans Huber, 1988.

Chapuis P., Dent O., Fisher R., Newland R., Pheils M., Smyth E. and Colquhoun K. 'A multivariate analysis of clinical and pathological variables in prognosis after resection of large bowel cancer' British Journal of Surgery 1985 72 698-702.

Charny M., Lewis P. and Farrow S. 'Choosing who shall not be treated in the NHS' Social Science and Medicine 1989 28 1331-1338.

Chester J. and Britton D 'Elective and emergency surgery for colorectal cancer in a district general hospital: impact of surgical training on patient survival' Annals of the Royal College of Surgeons of England 1989 71 370-374.

Cole P. and Morrison A. 'Basic issues in population screening for cancer' Journal of the National Cancer Institute 1980 64 1263-1272.

Cook G. and Wald N. 'Can the coverage of screening for cancer of the cervix be improved using the electoral register ? Cancer Letters 1985 28 307-310.

CRESGE (Centre de Recherches Economiques Sociologiques et de Gestion) "Interet du Test Hemoccult II - Une Premiere Apprcohe Medicale et Economique" CRESGE 1987.

Culyer A. and Maynard A. 'Cost-effectiveness of duodenal ulcer treatment' Social Science and Medicine 1981 15C 3-11.

Culyer A., MacFie J. and Wagstaffe A. 'Cost-effectiveness of foam elastomer and gauze dressings in the management of open perineal wounds' Social Science and Medicine 1983 17 1047-1053.

Dales L., Friedman G. and Collen G. 'Evaluating periodic multiphasic health check-ups: a controlled trial' Journal of Chronic Diseases 1979 32 385-404.

Deans G., Parks T., Rowlands B. and Spence R. 'Prognostic factors in colorectal cancer' British Journal of Surgery 1992 79 608-613.

Dearnaley D. 'Current issues in cancer: cancer of the prostate' British Medical Journal 1994 308 780-784.

de Cosse J. 'Early cancer detection: colorectal cancer' Cancer 1988 62 1787-1790.

de Haes J., de Koning H., van Oortmarssen G., van Agt H., de Bruyn A. and van der Maas P. 'The impact of a breast cancer screening programme on quality-adjusted life-years' International Journal of Cancer 1991 49 538-544.

de Koning H., van Oortmarssen G., van Ineveld B. and van der Maas P. 'Breast cancer screening: its impact on clinical medicine' British Journal of Cancer 1990 61 292-297.

de Koning H., van Ineveld B., de Haes J., van Oortmarssen G., Klijn J and van der Maas P. 'Advanced breast cancer and its prevention by screening' British Journal of Cancer 1992 65 950-955.

Dent O., Bartrop R., Goulston K. and Chapuis P. 'Participation in faecal occult blood screening for colorectal cancer' Social Science and Medicine 1983 17 17-23.

Department of Health "Health of the Nation" (1991).

Department of Health and Social Security "Health and Personal Social Services Statistics for England 1991" HMSO, 1992.

Department of Transport "Values for journey time savings and accident prevention" (1987).

Devesa J., Morales V., Enriquez J., Nuno J., Camunes J., Hernandez M. and Avila C. 'Colorectal cancer: the basis for a comprehensive follow-up' Diseases of the Colon and Rectum 1988 31 636-652.

Dixon A., Maxwell W. and Thornton Holmes J. 'Carcinoma of the rectum: a 10-year experience' British Journal of Surgery 1991 78 308-311.

Dixon A., Thornton Holmes J. and Cheetham N. 'General practitioners' awareness of colorectal cancer: a 10 year review' British Medical Journal 1990 301 152-153.

Doll R. "The Causes of Cancer" (1990).

Drummond M. "Principles of Economic Appraisal in Health Care" (1980).

Drummond M., Stoddart G. and Torrance G. "Methods for the Economic Evaluation of Health Care Programmes" Oxford University Press, 1987.

Dunlop M. 'Screening for large bowel neoplasms in individuals with a family history of colorectal cancer' British Journal of Surgery 1992 79 488-494.

Durdey P., Weston P. and Williams N. 'Colonoscopy and barium enema as initial investigation of colonic disease' Lancet 1987 ii 549-551.

Eddy D. 'The economics of cancer prevention and detection: getting more for less' Cancer 1981 47 1200-1209.

Eddy D. 'Finding cancer in asymptomatic people: estimating the benefits, costs and risks' Cancer 1983 51 2440-2445

Eddy D. 'Secondary prevention of cancer: an overview' Bulletin of the World Health Organisation 1986 64 421-429.

Eddy D. 'Cost-effectiveness of colorectal cancer screening' in Levin B. and Riddell H. (eds) "Frontiers in Gastrointestinal Cancer" Elsevier, 1984.

Eddy D. 'Screening for colorectal cancer' Annals of Internal Medicine 1990 113 373-384.

Eddy D. 'Screening for colorectal cancer' (letter) Annals of Internal Medicine 1991 114 167.

Eddy D., Hasselblad V., McGivney W. and Hendee W. 'The value of mammography screening in women under age 50 years' Journal of the American Medical Association 1988 259 1512-1519.

Eddy D., Nugent W., Eddy J., Coller J., Gilbertsen V., Gottlieb L., Rice R., Sherlock P. and Winawer S. 'Screening for colorectal cancer in a high-risk population' Gastroenterology 1987 92 682-692.

Eide T. 'Risk of colorectal cancer in adenoma-bearing individuals within a defined population' International Journal of Cancer 1986 38 173-176.

Elkind A., Eardley A., Haran D. Spencer B. and Smith A. 'Computer-managed call and recall for cervical screening: a typology of reasons for non-attendance' Community Medicine 1989 11 157-162.

Elliott M., Levenstein J. and Wright J. 'Faecal occult blood testing in the detection of colorectal cancer' British Journal of Surgery 1984 71 785-786.

Ellman R., Angeli N., Christians A., Moss S., Chamberlain J. and Maguire P. 'Psychiatric morbidity associated with screening for breast cancer' British Journal of Cancer 1989 60 781-784.

England W., Halls J., Hunt V. 'Strategies for screening for colorectal cancer' Medical Decision Making 1989 9 3-13.

Fallowfield L., Rodway A. and Baum M. 'What are the psychological factors influencing attendance, non-attendance and reattendance at a breast screening centre?' Journal of the Royal Society of Medicine 1990 83 547-551.

Farrands P. "Screening for colorectal cancer using the faecal occult blood test" M.D. Thesis, University of Nottingham (1985).

Farrands P. and Hardcastle J. 'Accuracy of occult blood tests over a six-day period' Clinical Oncology 1983 9 217-225.

Farrands P., Hardcastle J. 'Colorectal screening by a self-completion questionnaire' Gut 1984 25 445-447.

Farrands P., Hardcastle J., Chamberlain J. and Moss S. 'Factors affecting compliance with screening for colorectal cancer' Community Medicine 1984 6 12-19.

Farrands P., Vellacott K., Amar S., Balfour T. and Hardcastle J. 'Flexible fibreoptic sigmoidoscopy and double-contrast barium-enema examination in the identification of adenomas and carcinoma of the colon' Diseases of the Colon and Rectum 1983 26 727-729.

Feczko P. and Halpert R. 'Reassessing the role of radiology in Haemoccult screening' American Journal of Radiology 1986 146 697-701.

Feldman A., Kessler L., Myers M. and Naughton D. 'The prevalence of cancer: estimates based on the Connecticut Tumor Registry' New England Journal of Medicine 1986 315 1394-1397.

Fielding P., Phillips R., Fry J. and Hittinger R. 'Prediction of outcome after curative resection for large bowel cancer' Lancet 1986 ii 904-907

Fink R. and Shapiro S. 'Significance of increased efforts to gain participation in screening for breast cancer' American Journal of Preventive Medicine 1990 6 34-41.

Finkler S. 'The distinction between costs and charges' Annals of Internal Medicine 1982 96 102-109.

Fleischer D., Goldberg S., Browning T., Cooper J., Friedman E., Goldner F., Keeffe E. and Smith L. 'Detection and Surveillance of colorectal cancer' Journal of the American Medical Association 1989 261 580-585.

Foley D., Dunne P., Ryan T., Codd M., Dervan P., Crowe J., O'Callaghan T. and Lennon J. 'Left-sided colonoscopy highlights the failure of Haemoccult to detect colorectal neoplasia' Gut 1988 29 A1483.

Forrest P. "Breast Cancer Screening: report to the Health Ministers of England, Wales, Scotland and Northern Ireland by a working group chaired by P. Forrest" HMSO 1986.

Francis A., Polissar L. and Lorenz A. 'Care of patients with colorectal cancer: a comparison of a Health Maintenance Organisation and Fee-for-Service practices' Medical Care 1984 22 418-429.

Frank J. 'Occult-blood screening for colorectal carcinoma: the benefits' American Journal of Preventive Medicine 1985 1 (3) 3-9.

Frank J. 'Occult-blood screening for colorectal carcinoma: the risks' American Journal of Preventive Medicine 1985 1 (4) 25-32.

Frank J. 'Occult-blood screening for colorectal carcinoma: the yield and costs' American Journal of Preventive Medicine 1985 1 (5) 18-24.

Gabrielsson N., Granqvist S. and Nilsson R. 'Guaiac test for detection of occult blood loss in patients with endoscopically verified colonic polyps' Scandinavian Journal of Gastroenterology 1985 20 978-982.

Gerard K., Turnbull D., Lange M. and Mooney G. 'Economic evaluation of mammography screening: information reassurance and anxiety' Health Economics Research Unit Discussion Paper 01/92, University of Aberdeen.

Gilbertsen V. and Nelms J. 'The prevention of invasive cancer of the rectum' Cancer 1978 41 1137-1139.

Gillam S. 'Understanding the uptake of cervical screening: the contribution of the health belief model' British Journal of General Practice 1991 41 510-513.

Goddeeris J. and Broncken T. 'Benefit-cost analysis of screening: a comparison of tests for gonorrhoea' Medical Care 1985 23 1242-1255.

GOICC (Gruppo Operativo Italiano per il Cancro Colorettale): Crespi M., Caperle M., Dellamano R., Aste H., Caprilli R., Pacini F. and Terranova O.

"Cost-Benefit Analysis of Colorectal Cancer Screening: A Mathematical Model" (1990).

Gravelle H, Simpson P and Chamberlain J. 'Breast cancer screening and health service costs' Journal of Health Economics 1982 1 185-207.

Greegor D. 'Detection of silent colon cancer in routine examination' CA - A Cancer Journal for Clinicians 1969 19 330-337.

Grey M. 'Screening for colorectal cancer' (letter) Annals of Internal Medicine 1991 114 166.

Gyde S. 'Screening for colorectal cancer in ulcerative colitis: dubious benefits and high costs' Gut 1990 31 1089-1092.

Haiart D., McKenzie L., Henderson J., Pollock W., McQueen D., Roberts M. and Forrest Sir P. 'Mobile breast screening: factors affecting uptake, efforts to increase response and acceptability' Public Health 1990 104 239-247.

Hall C., Dorricott N., Donovan I. and Neoptolemos J. 'Colon perforation during colonoscopy: surgical *versus* conservative management' British Journal of Surgery 1991 78 542-544.

Hardcastle J., Balfour T. and Amar S. 'Screening for symptomless colorectal cancer by testing for occult blood in general practice' Lancet 1980 i 791-793.

Hardcastle J., Farrands P., Balfour T., Chamberlain J., Amar S. and Sheldon M. 'Controlled trial of faecal occult blood testing in the detection of colorectal cancer' Lancet 1983 ii 1-4.

Hardcastle J., Armitage N., Chamberlain J., Amar S., James P. and Balfour T. 'Fecal occult blood screening for colorectal cancer in the general population: results of a controlled trial' Cancer 1986 58 397-403.

Hardcastle J., Thomas W.M., Chamberlain J., Pye G., Sheffield J., James P., Balfour T., Amar S., Armitage N., Moss S. 'Randomised controlled trial of faecal occult blood screening for colorectal cancer: results for the first 107344 patients' Lancet 1989 i 119-23

Harper D. 'Disease costs in a surgical ward' British Medical Journal 1979 i 647-649.

Haycox A. and Haran D. 'The role of incentives in the promotion of screening for cervical cancer' Paper presented at the Health Economics Study Group, Newcastle-upon-Tyne, January 1988.

Hennigan T., Franks P., Hocken D and Allen-Mersh T. 'Rectal examination in general practice' British Medical Journal 1990 301 478-480.

Hillman R. and Nix G. "DHSS Funded Research into Specialty Costing 1980-1982" (1984).

Hixson L., Fennerty M., Sampliner R., McGee D. and Garewal H. 'Prospective study of the frequency and size distribution of polyps missed by colonoscopy' Journal of the National Cancer Institute 1990 82 1769-1772.

Hobbs P., Kay C., Friedman E., St Leger A., Lambert C., Boggis C., Howard T. and Asbury D. 'Response by women aged 65-79 to invitation for screening for breast cancer by mammography: a pilot study' British Medical Journal 1990 301 1314-1316.

Hoff G. 'Colorectal polyps - clinical implications: screening and cancer prevention' Scandinavian Journal of Gastroenterology 1987 22 769-775.

Hoff G., Foerster A., Vatn M. and Larsen S. 'Epidemiology of polyps in the rectum and colon: recovery and evaluation of unresected polyps two years after detection' Scandinavian Journal of Gastroenterology 1986 21 853-862.

Holland W. and Stewart S. "Screening in Health Care: Benefit or Bane?" (1990)

Holliday H. and Hardcastle J. 'Delay in diagnosis and treatment of symptomatic colorectal cancer' Lancet 1979 i 309-311.

Holmberg L., Adami H-O., Persson I., Lundstrom T and Tabar L. 'Demands on surgical inpatient services after mass mammographic screening' British Medical Journal 1986 293 779-782.

Hoogewerf P., Hislop T., Morrison B., Burns S. and Sizto R. 'Health belief and compliance with screening for fecal occult blood' Social Science and Medicine 1989 30 721-726.

Houlston R., Murday V., Harcopos C., Williams C. and Slack J. 'Screening and genetic counselling for relatives of patients with colorectal cancer in a family cancer clinic' British Medical Journal 1990 301 366-368.

Howe G. and 25 others 'Dietary intake of fiber and decreased risk of cancers of the colon and rectum: evidence from the combined analysis of 13 case-control studies' Journal of the National Cancer Institute 1992 84 1887-1896.

Hughes D. 'Costing consultations in general practice: towards a standardised method' Family Practice 1991 8 388-393.

Hunt S., Alexander F. and Roberts M. 'Attenders and non-attenders at a breast screening clinic: a comparative study' Public Health 1988 102 3-10.

Hurley S., Huggins R., Snyder R and Bishop J. 'The cost of breast cancer recurrences' British Journal of Cancer 1992 65 449-455.

Hurley S. and Livingstone P. 'Personal costs incurred by women attending a mammographic screening programme' Medical Journal of Australia 1991 154 132-134.

Hurst J. 'Planning and hospital costs' in "Economics and Health Planning" K. Lee (ed.) (1978).

Hutton J. Economic evaluation of magnetic resonance imaging' Centre for Health Economics Newsletter, University of York, No. 3, June 1988.

Irvine E., O'Connor J., Frost R., Shorvon P., Somers S., Stevenson G. and Hunt R. 'Prospective comparison of double contrast barium enema plus flexible sigmoidoscopy v colonoscopy in rectal bleeding' Gut 1988 29 1188-1193.

Irving M. and Catchpole B. 'Anatomy and physiology of the colon, rectum and anus' British Medical Journal 1992 304 1106-1108.

Jansen J. 'Participation in the first and second round of a mass-screening for colorectal cancer' Social Science and Medicine 1984 18 633-636.

Jarvinen H. and Mecklin J.-P. 'Screening for hereditary colorectal carcinoma' Scandinavian Journal of Gastroenterology 1989 24 1153-1158.

Jass J. 'Do all colorectal carcinomas arise in pre-existing adenomas?' World Journal of Surgery 1989 13 45-51.

Jensen J., Kewenter J., Haglind E., Lycke G., Svensson C. and Ahren C. 'Diagnostic accuracy of double-contrast enema and rectosigmoidoscopy in connection with faecal occult blood testing for the detection of rectosigmoid neoplasms' British Journal of Surgery 1986 73 961-964.

Jensen J., Kewenter J., *Astely M., Lycke* G. and Wojciechowski J. 'Double contrast barium enema and flexible recto-sigmoidoscopy: *a reliable* diagnostic combination for detection of colorectal neoplasm' British Journal of Surgery 1990 77 270-272.

Johnson M. and Jolly P. 'Analysis of a mass colorectal cancer screening program for cost-effectiveness' American Journal of Surgery 1987 154 261-263.

Joseph A., Crowson T. and Rich E. 'Cost-effectiveness of HemoQuant versus Hemoccult for colorectal cancer screening' Journal of General Internal Medicine 1988 3 132-138.

Kalra L. and Hamlyn A. 'Comparative evaluation of investigations for colorectal carcinoma in symptomatic patients' Postgraduate Medical Journal 1988 64 666-668.

Kewenter J., Bjork S., Haglind E., Smith L., Svanvik J., Ahren C. 'Screening and rescreening for colorectal cancer: a controlled trial of fecal occult blood testing in 27,700 subjects' Cancer 1988 62 645-651.

Kewenter J., Engaras B., Haglind E. and Jensen J. 'Value of retesting subjects with a positive Hemoccult in screening for colorectal cancer' British Journal of Surgery 1990 77 1349-1351.

Kewenter J., Haglind E. and Svanvik J. 'Faecal occult blood screening for colorectal cancer: the Swedish experience' in Faivre J. and Hill M. (eds) "Causation and Prevention of Colorectal Cancer" Excerpta Medica, 1987.

King's Fund "Consensus Conference on Colorectal Cancer - Final Statement" (1990).

Klaaborg K., Stahl Madsen M., Sondergaard O. and Kronborg O. 'Participation in mass screening for colorectal cancer with faecal occult blood tests' Scandinavian Journal of Gastroenterology 1986 21 1180-1184.

Klein R. and Redmayne S. "Patterns of Priorities" National Association of Health Authorities and Trusts Research Paper Number 7 (1992).

Klemi P., Joensuu H., Toikanen S., Tuominen J., Rasanen O., Tyrkko J. and Parvinen I. 'Aggressiveness of breast cancers found with and without screening' British Medical Journal 1992 304 467-469.

Knight K., Fielding J. and Battista R. 'Occult blood screening for colorectal cancer' Journal of the American Medical Association 1989 261 587-593.

Koopmanschap M., Lubbe K., van Oortmarssen G., van Agt H., van Ballegooijen M and Habbema D. 'Economic aspects of cervical cancer screening' Social Science and Medicine 1990 30 1081-1087.

Koretz R. 'Malignant polyps: are they sheep in wolves' clothing?' Annals of Internal Medicine 1993 118 63-68.

Kristein M. 'The economics of screening for colo-rectal cancer' Social Science and Medicine 1980 14C 275-284.

Kronborg O., Fenger C., Sondergaard Q., Pedersen K. and Olsen J. 'Initial mass screening for colorectal cancer with faecal occult blood test: a prospective randomised study at Funen in Denmark' Scandinavian Journal of Gastroenterology 1987 22 677-686.

Kronborg O., Fenger C., Olsen J., Bech K. and Sondergaard O. 'Repeated screening for colorectal cancer with faecal occult blood test' Scandinavian Journal of Gastroenterology 1989 24 599-606.

Kronborg O., Fenger C., Worm J., Pedersen S., Hem J., Bertelsen K. and Olsen J. 'Causes of death during the first five years of a randomized trial of mass screening for colorectal cancer with fecal occult blood test' Scandinavian Journal of Gastroenterology 1992 27 47-52.

Leamer E. "Specification Searches: Ad Hoc Inference with Non-experimental Data" Wiley 1978.

Leicester R., Lightfoot A., Millar J., Colin-Jones D. and Hunt R. 'Accuracy and value of the Haemoccult test in symptomatic patients' British Medical Journal 1983 286 673-674.

Levin K. and Doizois R. 'Epidemiology of large bowel cancer' World Journal of Surgery 1991 15 562-567.

Lewis A. and Khoury G. 'Resection for colorectal cancer in the very old: are the risks too high?' British Medical Journal 1988 296 459-461.

Lieberman D. Colon cancer screening: the dilemma of positive screening tests' Archives of Internal Medicine 1990 150 740-744.

Lieberman D. 'Cost-effectiveness of colon cancer screening' American Journal of Gastroenterology 1991 86 1789-1794.

Lindsay D., Freeman J., Cobden I., Record C. 'Should colonoscopy be the first investigation for colonic disease?' British Medical Journal 1988 296 167-169.

McArdle C. and Hole D. 'Impact of variability among surgeons on postoperative morbidity and mortality and ultimate survival' British Medical Journal 1991 302 1501-1505.

MacDonald L. and Anderson R. 'The health of rectal cancer patients in the community' European Journal of Surgical Oncology 1985 11 235-241.

McEwan J., King E. and Bickler G. 'Attendance and non-attendance for breast screening at the South East London breast screening service' British Medical Journal 1989 299 104-106.

McGowan I., Barr H. and Krasner N. 'Palliative laser therapy for inoperable rectal cancer - dies it work? A prospective study of quality of life' Cancer 1989 63 967-969.

McNeil B, Weichelsbaum R. and Pauker S. 'Fallacy of five-year survival in lung cancer' New England Journal of Medicine 1978 299 1397-1401.

McNeil B, Weichelsbaum R. and Pauker S. 'Speech and survival: trade-offs between quantity and quality of life in laryngeal cancer' New England Journal of Medicine 1981 305 982-987.

Macrae F., Hill D., St. John J., Ambikapathy A. and Garner J. 'Predicting colon cancer screening behaviour from health beliefs' Preventative Medicine 1984 13 115-126.

Macrae F., St. John J., Ambikapathy A., Sharpe K. and Garner J. 'Factors affecting compliance in colorectal cancer screening' Medical Journal of Australia 1986 144 621-623.

Macrae F. and St John J. 'Relationship between patterns of bleeding and Hemoccult sensitivity in patients with colorectal cancers or adenomas' Gastroenterology 1982 82 891-898.

Macrae F., Tan K. and Williams C. 'Towards safer colonoscopy: a report on the complications of 5000 diagnostic or therapeutic colonoscopies' Gut 1983 24 376-383.

Mandel J., Bond J., Snover D., Williams S., Bradley M., Walker C., Schuman L and Gilbertsen V. 'The University of Minnesota's Colon Cancer Control Study: design and progress to date' in Chamberlain J. and Miller A. (eds.) "Screening for Gastrointestinal Cancer" (1988).

Mandel J., Bond J., Church T., Snover D., Bradley M., Schuman L. and Ederer F. for the Minnesota Colon Cancer Control Study 'Reducing mortality from colorectal cancer by screening for fecal occult blood' New England Journal of Medicine 1993 328 1365-1371.

Mant D., Fitzpatrick R., Hogg A., Fuller A., Framer A., Verne J. and Northover J. 'Experiences of patients with false positive results from colorectal cancer screening' British Journal of General Practice 1990 40 423-425.

Marteau T. 'Psychological costs of screening' British Medical Journal 1989 299 527.

Marteau T. 'Reducing the psychological costs' British Medical Journal 1990 301 26-28.

Maxfield R. 'Colonoscopy as a routine preoperative procedure for carcinoma of the colon' American Journal of Surgery 1984 147 477-480.

Meagher A. and Stuart M. 'Colonoscopy in patients with a family history of colorectal cancer' Diseases of the Colon and Rectum 1992 35 315-321.

Miller A. 'Introduction' in A. Miller "Screening for Cancer" (1985).

Moertel C., Fleming T., MacDonald J., Haller D., Laurie J., Goodman P., Ungerleider J., Emerson W., Tormey D., Glick J., Veeder M and Mailliard J. 'Levamisole and fluouracil for adjuvant therapy of resected colon carcinoma' New England Journal of Medicine 1990 322 352-358.

Morris J., Stellato T., Guy B., Gordon N. and Berger N. 'A critical analysis of the largest mass fecal occult blood screening program in the United States' American Journal of Surgery 1991 161 101-106.

Morson B. 'The evolution of colorectal carcinoma' Clinical Radiology 1984 35 425-431.

Morson B. and Bussey H. 'Magnitude of risk for cancer in patients with colorectal adenomas' British Journal of Surgery 1985 72 S23-S25.

Muto T., Bussey H. and Morson B. 'The evolution of cancer of the colon and rectum' Cancer 1975 36 2251-2270.

NHS Management Executive "Costing for Contracts"FDL(93)52, 1993.

NHS Management Executive Value for Money Unit "The Management and Utilisation of Operating Departments" ('The Bevan Report') (1989).

Neugut A. 'Screening for colorectal cancer' (letter) Annals of Internal Medicine 1990 113 899-900.

Neugut A. and Forde K. 'Screening colonoscopy: has the time come?' American Journal of Gastroenterology 1988 83 295-297.

Neugut A., Garbowski G., Loe W., Murray T., Nieves J., Forde K., Treat M., Waye J. and Fenogloi-Preiser C. 'Dietary risk factors for the incidence and recurrence of colorectal adenomatous polyps' Annals of Internal Medicine 1993 118 91-95.

Neuhauser D. 'The sixth stool guaiac test: a reply' Journal of Health Economics 1990 9 493-494

Neuhauser D. and Lewicki A. 'What do we gain from the sixth stool guaiac?' New England Journal of Medicine 1975 293 226-228.

Newcomb P., Norfleet R., Storer B., Surawicz T. and Marcus P. 'Screening sigmoidoscopy and colorectal cancer mortality' Journal of the National Cancer Institute 1992 84 1572-1575.

Nichols S. and Waters W. 'Effect on doctors' work load of a campaign encouraging early reporting of breast symptoms' Journal of Epidemiology and Community Health 1982 36 228-230.

Nichols S., Koch E., Lallemand R., Heald R., Izzard L., Machin D. and Mullee M. 'Randomised trial of compliance with screening for colorectal cancer' British Medical Journal 1986 293 107-110.

Nomura A. 'An international search for causative factors of colorectal cancer' Journal of the National Cancer Institute 1990 82 894-895.

O'Brien M., Winawer S., Graham A., Zauber G., Gottlieb L., Sternberg S., Fiaz B., Dickersin R., Ewing S., Geller S., Kasimian D., Komorowski R., Szporon A. and the National Polyp Study Workgroup 'The National Polyp Study: patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas' Gastroenterology 1990 98 371-379.

Office of Population Censuses and Surveys "Cancer Statistics: Registrations 1987" HMSO 1993.

Office of Population Censuses and Surveys "English Life Tables, No. 14" HMSO 1987.

Office of Population Censuses and Surveys "Mortality Statistics: Cause 1991" HMSO 1993.

Ott D., Gelfand D., Yu M. and Munitz A. 'Colonoscopy and the barium enema: a radiological viewpoint' Southern Medical Journal 1985 78 1033-1035.

Ozoux J., de Calan L., Perrier M., Favre J. and Brizon J. 'Surgery for carcinoma of the colon in people aged 75 years and older' International Journal of Colorectal Disease 1990 5 25-30.

Parsonage M. and Neuburger H. 'Discounting and health benefits' Health Economics 1992 1 71-79.

Payne J., Murdoch C., Dent O. and Chapuis P. 'The cost of resection for colorectal cancer' Australia and New Zealand Journal of Surgery 1987 57 627-633.

Phillips R., Hittinger R., Blesovsky L., Fry J. and Fielding P. 'Local recurrence following 'curative' surgery for large bowel cancer: I. The overall picture' British Journal of Surgery 1984 71 12-16.

Piachaud D. and Weddell J. 'The economics of treating varicose veins' International Journal of Epidemiology 1972 1 287-293.

Polednak A. 'Knowledge of colorectal cancer and use of screening tests in persons aged 40-74 years of age' Preventive Medicine 1990 19 213-226.

Pollard S., MacFarlane R. and Everett W. 'Surgery for recurrent colorectal carcinoma - is it worthwhile?' Annals of the Royal College of Surgeons of England 1989 71 293-298.

Pollock A. and Quirke P. 'Adenoma screening and colorectal cancer' British Medical Journal 1991 303 3-4.

Proctor S. 'It all depends' Health Service Journal 21.1.1993 27

Pye G., Christie M., Chamberlain J., Moss S. and Hardcastle J. 'A comparison of methods for increasing compliance within a general practitioner based screening project for colorectal cancer, and the effect on practitioner workload' Journal of Epidemiology and Community Health 1988 42 66-71.

Pye G., Jackson J., Thomas W.M. and Hardcastle J.D. 'Comparison of Coloscreen Self-Test and Haemoccult faecal occult blood tests in the detection of colorectal cancer in symptomatic patients' British Journal of Surgery 1990 77 630-631.

Pye G., Marks C., Martin S., Marks V., Jackson J. and Hardcastle J. 'An evaluation of Fecatwin/Feca EIA; a faecal occult blood test for detecting colorectal neoplasia' European Journal of Surgical Oncology 1989 15 446-448.

Ransohoff D., Lang C. and Kuo S. 'Colonoscopic surveillance after polypectomy: considerations of cost-effectiveness' Annals of Internal Medicine 1991 114 177-182.

Ransohoff D. and Lang C. 'Small adenomas detected during fecal occult blood test screening for colorectal cancer: the impact of serendipity' Journal of the American Medical Association 1990 264 76-78.

Ratcliffe J. 'The measurement of indirect costs in economic evaluation: a critical review' Paper presented to the Health Economists' Study Group, Strathclyde, July 1993.

Rawlinson J., Tate J., Brunton F. and Royle G. 'Radiological evidence for the polyp/cancer sequence in the colon' Clinical Radiology 1989 40 386-388.

Resabeck P. 'Colorectal cancer: the case for endoscopic screening' British Journal of Surgery 1987 74 12-17.

Rees M. "The costing of a control trial of the use of laser and thermal probe for the treatment of bleeding peptic ulcers" mimeo, 1987.

Reiertsen O., Bakka A., Tronnes S. and Gauperaa T. 'Routine double-contrast barium enema and fibreoptic colonoscopy in the diagnosis of colorectal cancer' Acta Chir. Scandinavia 1988 154 53-55.

Rex D., Wedge R., Lehman G., Pound D., O'Connor K., Hawes R., Dittus R., Lappas J. and Lumeng L. 'Flexible sigmoidoscopy plus air contrast barium enema versus colonoscopy for suspected lower gastrointestinal bleeding' Gastroenterology 1990 98 855-861.

Rex D., Lehman G., Hawes R., Ulbright T. and Smith J. 'Screening colonoscopy in asymptomatic average-risk persons with negative fecal occult blood tests' Gastroenterology 1991 100 64-67.

Rich G., Glass N. and Selkon B. 'Cost-effectiveness of two methods of screening for asymptomatic bacteriuria' British Journal of Preventive and Social Medicine 1976 30 54-59.

Ridley S., Biggam M. and Stone P. 'Cost of intensive therapy: a description of methodology and initial results' Anaesthesia 1991 46 523-530.

Ritchie J., Jacoby A. and Bone M. "Access to primary health care" (1981).

Roberts M. 'Breast screening: time for a rethink?' British Medical Journal 1989 299 1153-1155.

Roberts M., Alexander F., Anderson T., Chetty U., Donnan P., Forrest P., Hepburn W., Huggins A., Kirkpatrick A., Lamb J., Muir B. and Prescott R. 'Edinburgh trial of screening for breast cancer: mortality at seven years' Lancet 1990 335 241-246.

Robra B. 'The statutory health insurance cancer screening programme in the Federal Republic of Germany' in 'Screening for colorectal cancer' International Journal of Colorectal Disease 1986 1 63-78.

Rosenstock I. 'Patients' compliance with health regimens' Journal of the American Medical Association 1975 234 402-403.

Ross S. 'Cervical cytology screening and government policy' British Medical Journal 1989 299 101-104.

Royal College of General Practitioners, Office of Population Censuses and Surveys and Department of Health and Social Security "Morbidity Statistics from General Practice 1981/2 - Third National Study" Series MB5 No. 1, HMSO 1986.

Rumans M., Benner K., Keeffe E., Custis J., Lockwood D. and Craner G. 'Screening flexible sigmoidoscopy by primary care physicians: effectiveness and costs in patients negative for fecal occult blood' Western Journal of Medicine 1986 144 756-758.

St. John J. 'Will colorectal cancer screening be too costly?' Australia and New Zealand Journal of Surgery 1990 60 415-417.

St. John J., Young G., McHutchinson J., Deacon M. and Alexeyeff M. 'Comparison of the specificity and sensitivity of Hemoccult and HemoQuant in screening for colorectal neoplasia' Annals of Internal Medicine 1992 117 376-382.

Schofield P. and Jones D. 'Colorectal neoplasia - II: large bowel cancer' British Medical Journal 1992 304 1561-1563.

Scott N. and Quirke P. 'Molecular biology of colorectal neoplasia' Gut 1993 34 289-292.

Selby J. and Friedman G. 'Sigmoidoscopy in the periodic health examination of asymptomatic adults' Journal of the American Medical Association 1989 261 595-601

Selby J., Friedman G. and Collen M. 'Sigmoidoscopy and mortality from colorectal cancer: the Kaiser Permanente Multiphasic Evaluation Study' Journal of Clinical Epidemiology 1988 41 427-434.

Selby J., Friedman G., Quesenberry C. and Weiss N. 'A case-control study of screening sigmoidoscopy and mortality from colorectal cancer' New England Journal of Medicine 1992 326 653-657.

Selby J., Friedman G., Quesenberry C. and Weiss N. 'Effects of fecal occult blood testing on mortality from colorectal cancer: a case-control study' Annals of Internal Medicine 1993 118 1-6.

Shapiro S. 'Case-control studies of colorectal cancer mortality: is the case made for screening sigmoidoscopy?' Journal of the National Cancer Institute 1992 84 1546-1547.

Shiel A., Griffiths R., MacMillan R. and Atherton S. 'Pilot assessment of the cost-effectiveness of a British intensive care unit' British Journal of Hospital Medicine 1989 42 144-145.

Shike M., Winawer S., Greenwald P., Bloch A., Hill M., Swaroop S. and the WHO Collaborating Centre for the Prevention of Colorectal Cancer' Bulletin of the World Health Organisation 1990 68 377-385.

Silman A., Nicholls R., Mitchell P., Macrae F., Leicester R., Bartram C., Simmons M., Campbell P., Hearn C. and Constable P. 'Self-reported dark red bleeding as a marker comparable with occult blood testing in screening for large bowel neoplasms' British Journal of Surgery 1983 70 721-724.

Silman A. and Mitchell P. 'Attitudes of non-participants in an occupational based programme of screening for colorectal cancer' Community Medicine 1984 6 8-11.

Simon J. 'Occult blood screening for colorectal cancer: a critical review' Gastroenterology 1985 88 820-837.

Simon J. 'Colonic polyps, cancer and fecal occult blood' Annals of Internal Medicine 1993 118 71-73.

Skrabanek P. Review of "Screening in Health Care: Benefit or Bane?" Holland W. and Stewart S. (1990) Lancet 1991 337 100.

Slaney G. "Clinical Cancer Monographs: Large Bowel" (1991).

Spratt J. and Spratt J. 'Growth rates of benign and malignant neoplasms of the colon' in "Carcinoma of the Large Bowel and its Precursors" J. Ingall and A. Mastromarino (eds.) (1985)

Standing Medical Advisory Committee, Department of Health "Blood Cholesterol Testing: the Cost-Effectiveness of Opportunistic Testing" (1990).

Stilman A. 'Age-sex registers as a screening tool for general practice: size of the wrong address problem' British Medical Journal 1984 289 415-416.

Stilwell J. 'Costs of clinical laboratory tests' Journal of Clinical Pathology 1981 34 589-594.

Stower M. and Hardcastle J. 'The results of 1115 patients with colorectal cancer treated over an 8-year period in a single hospital' European Journal of Surgical Oncology 1985 11 119-123.

Stryker S., Wolff B., Culp C., Libbe S., Ilstrup D. and MacCarty R. 'Natural history of untreated colonic polyps' Gastroenterology 1987 93 1009-1013.

Svendsen L., Sorensen C., Kjersgaard P., Meisner S. and Kjaergaard J. 'The influence of age upon the survival after curative operation for colorectal cancer' International Journal of Colorectal Disease 1989 4 123-127.

Tartter P. 'Determinants of postoperative stay in patients with colorectal cancer' Diseases of the Colon and Rectum 1988 31 694-698.

Tate J., Northway J., Royle G. and Taylor I. 'Evaluation of a 'DIY' test for the detection of colorectal cancer' Journal of the Royal Society of Medicine 1989 82 388-390.

Tate J., Northway J., Royle G and Taylor I. 'Faecal occult blood testing in symptomatic patients: comparison of three tests' British Journal of Surgery 1990 77 523-526.

Thomas W.M., Pye G., Hardcastle J., Chamberlain J. and Charnley R. 'Role of dietary restriction in Haemoccult screening for colorectal cancer' British Journal of Surgery 1989 76 976-978.

Thomas W.M., Pye G., Hardcastle J. and Mangham C. 'Faecal occult blood screening for colorectal neoplasia: a randomised trial of three days or six days of tests' British Journal of Surgery 1990 77 277-279.

Thomas W.M., Jackson J., Lamont G. and Hardcastle J. 'Early incontinence following flexible sigmoidoscopy' British Journal of Surgery 1990 77 887.

Thomas W.M., Pye G., Hardcastle J. and Walker A. 'Screening for colorectal carcinoma: an analysis of Haemoccult sensitivity' British Journal of Surgery 1992 79 833-835.

Thomas M., Hardcastle J., Jackson J. and Pye G. 'Chemical and immunological testing for faecal occult blood: a comparison of two tests in symptomatic patients' British Journal of Cancer 1992 65 618-620.92

Thompson R., Michnich M., Gray J., Friedlander L. and Gilson B. 'Maximising compliance with Hemoccult screening for colon cancer in clinical practice' Medical Care 1986 24 904-914.

Thorson A., Christensen M. and Davis S. 'The role of colonoscopy in the assessment of patients with colorectal cancer' Diseases of the Colon and Rectum 1986 29 306-311.

Tuck J., Walker A., Whynes D., Pye G., Hardcastle J. and Chamberlain J. 'Screening and the costs of treating colorectal cancer: some preliminary results' Public Health 1989 103 413-419.

Umpleby H., Bristol J., Rainey J. and Williamson R. 'Survival of 727 patients with single carcinomas of the large bowel' Diseases of the Colon and Rectum 1984 27 803-810.

van Ness M., Chobanin S., Winters C., Diehl A., Esposito R. and Catiau E. 'A study of patient acceptance of double-contrast barium enema and colonoscopy: which procedure is preferred by patients?' Archives of Internal Medicine 1987 147 2175-2176.

Vernon S., Laville E. and Jackson G. 'Participation in breast screening programs: a review' Social Science and Medicine 1990 30 1107-1118.

Wagner J., Herdman R. and Wadhwa S. 'Cost-effectiveness of colorectal screening in the elderly' Annals of Internal Medicine 1991 115 807-817.

Warner K. and Luce B. "Cost-Benefit and Cost-Effectiveness Analysis in Health Care" (1982).

Wessex Cancer Intelligence Unit "Report on the Cost-Effectiveness of Treatment for Colorectal Cancer in three Health Districts" (1990).

Willett W. 'The search for the causes of breast and colon cancer' Nature 1989 338 389-394.

Williams A., Balasooriya B. and Day D. 'Polyps and cancer of the large bowel: a necropsy study in Liverpool' Gut 1982 23 835-842.

Williams A. 'The economics of coronary artery bypass grafting' British Medical Journal 1985 291 326-329.

Williams F. and Lloyd G. 'Gastric and colorectal cancers in Scotland: a study of the geographical distributions and selected associations' Scottish Medical Journal 1990 35 136-139.

Winawer S., Fath R., Schottenfeld D. and Herbert E. 'Screening for colorectal cancer' in "Screening for Cancer" A. Miller (ed.) (1985).

Winawer S., Schottenfeld D. and Flehinger B. 'Colorectal cancer screening' Journal of the National Cancer Institute 1991 83 243-253.

Working Party of the Clinical Services Committee of the British Society of Gastroenterology 'Provision of gastrointestinal endoscopy and related services for a district general hospital' Gut 1991 32 95-105.

Wright S. "Age, Sex and Health: a Summary of the Findings from the York Health Evaluation Survey" Centre for Health Economics Discussion Paper 15, University of York, 1986.

Wyke R., Edwards F. and Allan R. 'Employment problems and prospects for patients with inflammatory bowel disease' Gut 1988 29 1229-1235.