Effectiveness and cost-effectiveness of population screening for colorectal cancer

A systematic review of the literature

Jane Kerr, Marita Broadstock, Peter Day and Sarah Hogan

New Zealand Health Technology Assessment

Department of Public Health and General Practice
Christchurch School of Medicine and Health Sciences
Christchurch, New Zealand.
Effectiveness and cost-effectiveness of population screening for colorectal cancer

A systematic review of the literature

Jane Kerr, Marita Broadstock, Peter Day and Sarah Hogan
This report should be referenced as follows:


2007  New Zealand Health Technology Assessment (NZHTA)

ISBN  978-1-877455-14-8
ISSN  1174-5142
ACKNOWLEDGEMENTS

This systematic review was conducted by staff of NZHTA. It was prepared by Dr Jane Kerr (Research Fellow), Ms Marita Broadstock (Research Fellow), Mr Peter Day (Research Fellow) and Ms Sarah Hogan (Health Economist Consultant) who conducted the critical appraisals and prepared the report. Dr Kerr (and in the final stages, Ms Broadstock) coordinated the project. Dr Robert Weir (Senior Research Fellow and Acting Director) assisted with the meta-analysis and provided internal peer review of drafts. Ms Susan Bidwell (Information Specialist Manager) developed and undertook the search strategy and coordinated retrieval of documents and management of referencing. Ms Cecilia Tolan (Administrator) provided administrative support and Mrs Ally Reid (Administrative Secretary) provided document formatting. The Canterbury Medical Library assisted with inter-library loan. Ms Carol Webb provided sub-editing.

Acknowledgment is made of the contributions of Dr Susan Parry, Gastroenterologist, Middlemore Hospital; Associate Professor Ann Richardson, Epidemiologist, Department of Public Health and General Practice, Christchurch School of Medicine and Health Sciences; Dr Simon Baker, Public Health Physician, National Screening Unit; Ms Bronwyn Petrie, Analyst, National Screening Unit; and Dr Terri Green, Health Economist, University of Canterbury who provided valuable comments on the report. We are also grateful for editorial review provided by Dr Susan Parry, Associate Professor Ann Richardson, and Dr Terri Green.

NZHTA is a Research Unit of the University of Otago funded under contract to the New Zealand Ministry of Health.

This report was commissioned by Dr Ashley Bloomfield, Public Health Leader, National Screening Unit, of New Zealand’s Ministry of Health. We thank Dr Susan Parry and Associate Professor Ann Richardson for assisting in developing the scope of the review and providing background material. We also thank Mr Chris Sies, Medical Technologist, Canterbury Health Laboratories, for assistance with technical aspects of faecal occult blood testing.

DISCLAIMER

New Zealand Health Technology Assessment (NZHTA) takes great care to ensure the information supplied within the project timeframe is accurate, but neither NZHTA, the University of Otago, nor the contributors involved can accept responsibility for any errors or omissions. The reader should always consult the original database from which each abstract is derived, along with the original articles, before making decisions based on a document or abstract. All responsibility for action based on any information in this report rests with the reader. NZHTA and the University of Otago accept no liability for any loss of whatever kind, or damage, arising from reliance in whole or part, by any person, corporate or natural, on the contents of this report. This document is not intended as personal health advice. People seeking individual medical advice are referred to their physician. The views expressed in this report are those of NZHTA and do not necessarily represent those of the University of Otago or the New Zealand Ministry of Health.

NZHTA is a Research Unit of the University of Otago and is funded under contract by the New Zealand Ministry of Health.

COPYRIGHT

This work is copyright. Apart from any use as permitted under the Copyright Act 1994 no part may be reproduced by any process without written permission from New Zealand Health Technology Assessment. Requests and inquiries concerning reproduction and rights should be directed to the Director, New Zealand Health Technology Assessment, Christchurch School of Medicine and Health Sciences, P O Box 4345, Christchurch, New Zealand.
CONTACT DETAILS

New Zealand Health Technology Assessment (NZHTA)  
Department of Public Health and General Practice  
Christchurch School of Medicine and Health Sciences  
PO Box 4345  
Christchurch  
New Zealand  
Tel: +64 3 364 3696  Fax: +64 3 364 3697  
Email: nzhta@chmeds.ac.nz  
Web Site: http://nzhta.chmeds.ac.nz/
EXECUTIVE SUMMARY

Background
At present, there is no routine organised population screening for colorectal cancer (CRC) in New Zealand. In 2005, the New Zealand Ministry of Health’s National Screening Unit (NSU) plans to formally revisit the advice of the Working Party on Screening for Colorectal Cancer’s (1998) on population screening for CRC. The NSU asked NZHTA to undertake a systematic review of the evidence since 1997 regarding population screening for CRC, in order to update the 1998 review. The purpose of this systematic review is to provide information to assist the NSU to assess whether there should be an organised screening programme in New Zealand for CRC.

Aim
To systematically identify and appraise the international evidence for the effectiveness and cost-effectiveness of screening tests for colorectal cancer.

Data sources
A set of key terms relevant to the topic was developed and the literature was searched using major bibliographic databases, review databases, and health technology assessment resources. Internet websites were searched for clinical trial information, guidelines, screening programmes of other health systems, and details of the tests mentioned in the recent literature.

Full information on the data sources and search strategies is given in Chapter 2 and Appendices II and III.

Selection criteria
Inclusion criteria
Studies were included if they compared the clinical effectiveness or cost-effectiveness of:

- faecal occult blood test (FOBT) screening compared with no screening
- immunochemical FOBT screening compared with guaiac FOBT screening
- flexible sigmoidoscopy (FS) screening compared with no screening
- FS and FOBT combined screening compared with FOBT screening, FS screening, or no screening.

Publications included primary research (published as full original reports) and secondary research (systematic reviews and meta-analyses). Papers were included for review if they were published between January 1997 and November 2004 inclusive. Earlier papers were accessed where required to provide background material.

For primary research studies relevant to the effectiveness of screening tests for CRC only randomised controlled trials were included (with the exception of studies examining the accuracy of immunochemical FOBT screening compared with guaiac FOBT screening, where consideration was also made of study designs of a specific level of evidence and above to reflect the “best evidence” available on the issue).

For economic studies relevant to the cost-effectiveness of screening tests for CRC, study types considered for inclusion were cost-minimisation analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis (Drummond et al., 1997).

Secondary research studies reporting systematic reviews or meta-analyses of randomised controlled trials (RCTs) were included if they included a methods section describing how the relevant studies were identified. Secondary research papers were appraised principally as background information.
Exclusion criteria

Excluded studies included non-systematic reviews, correspondence, editorials, expert opinion articles, comments, articles published in abstract form, conference proceedings, studies that did not clearly describe their methods/results, non-English language articles as none were deemed necessary to the review, and animal studies.

Results

Of 1,986 articles identified by the search strategy, 220 articles were retrieved as full text. From these, a final group of 56 articles was identified as eligible for appraisal and inclusion in the review. Some of these articles included data relevant to multiple research questions, requiring more than one appraisal; appraisals therefore totalled 64:40 considering primary research articles and 24 considering systematic reviews.

Data extraction and synthesis

A systematic method of appraising and grading was employed in the preparation of this report.

The level of evidence presented in the selected studies of effectiveness was initially considered using the dimensions of evidence defined by the National Health and Medical Research Council (2000).

For primary and secondary studies of the effectiveness of screening tests for CRC, data relevant to study quality and statistical precision was extracted using design-relevant checklists (GATE: a Graphic Appraisal Tool for Evidence-based clinical practice. (This checklist is available from: http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/. Current as of May 2005).

Key outcomes for reports of clinical effectiveness included, where possible, diagnostic performance (sensitivity, specificity, positive predictive value, negative predictive value), health outcomes (stage of CRC at detection, CRC morbidity or mortality, overall mortality, incidence of CRC) and outcomes relating to screening (benefits, harms, acceptability).

Economic studies of cost-effectiveness were described and appraised in terms of their design, methods, data sources, key results, sensitivity of the model to value changes in variables, limitations and conclusions.

Key outcomes for reports of cost-effectiveness included, where possible: cost per life year saved, cost per disability-adjusted life year (DALY) saved, cost per quality-adjusted life year (QALY), cost per life saved, or cost per cancer detected. In comparing different screening modalities, outcomes also included, where possible, incremental measures of cost-effectiveness.

Key results

Clinical effectiveness of faecal occult blood test (FOBT) screening compared with no screening

Eleven eligible primary research papers and 13 eligible secondary research papers (representing nine reviews) were identified, providing evidence from four eligible RCTs that compared FOBT screening with no screening (level II evidence). Since 1997, ongoing follow-up data has become available from three major trials reported on by the Working Party on Screening for Colorectal Cancer (1998) that compared screening using the guaiac test Haemoccult/Haemoccult II with no screening. The fourth RCT appraised compared screening with an immunochemical test plus a health questionnaire to no screening.

Consistent with the findings of the Working Party on Screening for Colorectal Cancer (1998), high-quality evidence was found that FOBT screening with the guaiac FOBT Haemoccult reduces mortality from CRC. Evidence from ongoing follow-up for the three major RCTs suggested that this mortality reduction has been sustained for the populations in which screening has stopped, but decreased slightly for the population to whom screening has continued to be offered. The direct evidence available
concerning the efficacy of an immunochemical test to reduce CRC mortality was less robust, but did suggest that a reduction in rectal cancer may be achievable with the use of this test.

Some new evidence regarding the benefits and harms from CRC screening with FOBT has become available. However, little new data regarding possible psychological harms has been identified by this review since the Working Party on Screening for Colorectal Cancer (1998) report.

Accuracy of immunochemical FOBT screening compared with guaiac FOBT screening

Seven eligible primary research papers and five eligible secondary research papers were identified that evaluated a diverse range of both immunochemical and guaiac FOBTs. Design and quality flaws limited the conclusions that could be drawn from the published literature concerning comparative performance of immunochemical and guaiac FOBTs. However, evidence considered suggests that only HemeSelect (an immunochemical test) performs as well as, or better than, the guaiac tests HOII and HOS when compared head to head in cross-sectional studies.

There was good evidence that the simplified testing process and sampling kit of the immunochemical test Insure encouraged a greater proportion of people from the general population to participate when invited to complete FOB screening tests.

The evidence for this topic is consistent with that derived in other recent reviews. There is limited definitive evidence regarding superior immunochemical FOBT performance over the guaiac tests. HemeSelect is the immunochemical test that compares most favourably with the guaiac tests, but Insure may be more acceptable.

Clinical effectiveness of flexible sigmoidoscopy (FS) screening compared with no screening

One eligible primary research paper and two eligible secondary research papers were identified. No large-scale RCT providing incidence and mortality data relevant to the impact of flexible sigmoidoscopy screening has been completed. Only one RCT was identified in the current review, the Telemark Polyp Study. This study was graded as level III-1 evidence because the procedure for allocating the intervention was not truly random. Designed as a small feasibility study, there were only 400 people allocated to the screening group. Thirteen-year follow-up data were reported suggesting lower CRC incidence but higher mortality for the screening group compared to the control group. However, the study was limited by its relatively small size; differences in the method of selection for the intervention and control groups; and that 13-year follow-up colonoscopy was not offered to those non-attending for FS in the screening group. Two systematic reviews reporting on evidence from RCTs were identified, both discussing the Telemark Polyp Study (which was pseudo-randomised) and its limitations. One review concluded that definitive evidence from randomised controlled trials was required before widespread screening could be seriously considered.

Three well designed, large multi-centre studies were underway as this report was prepared which aim to provide reliable evidence concerning health outcomes following FS screening. Two of these studies offered one-time FS screening and expect to report incidence and mortality data in 2008 at the earliest. The other RCT underway is the US’s Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO), a community-based, multi-centre study evaluating the effectiveness of repeated cancer screening tests (including FS screening every 3-5 years) on site-specific cancer mortality. This trial will not report incidence and mortality data until 2010-2012. Baseline and procedural data from all three trials suggest that screening through flexible sigmoidoscopy is likely to be feasible and acceptable.

Clinical effectiveness of FS and FOBT combined screening compared with FOBT screening, FS screening, or no screening

Two eligible primary research papers and one eligible secondary research paper were identified. The two eligible RCTs compared combined FOBT and FS screening with FOBT screening. One RCT was graded as level II evidence and the other as level III-1 evidence.
The Working Party on Screening for Colorectal Cancer (1998) found there was a lack of evidence that first-line FS screening reduces mortality from CRC. Combined FOBT and FS screening was not considered in the 1998 Working Party report. The evidence on the effectiveness of combined screening considered in this review was insufficient to determine the superiority of combined screening over either screening strategy alone in terms of a reduction in CRC mortality. Combined FOBT and FS as a first-line screening strategy in asymptomatic middle-aged populations is not supported from the available literature. There was a lack of high quality RCTs evaluating the longer-term health outcomes of CRC incidence and mortality. The main focus of the studies considered was on screening compliance, diagnostic yield and colonoscopic utilisation, and findings from once-only combined screening compared to either modality alone. CRC mortality data from the ongoing NORCCAP trial is expected to become available in late 2007 and may provide greater clarity regarding the clinical effectiveness of combined screening compared to FS screening alone.

Additional to the three RCTs included in the section on clinical effectiveness of combined screening, three other eligible RCTs were also identified and appraised that provided data on the diagnostic accuracy of combined FOBT and FS screening compared to either modality alone. There were no CRC incidence or mortality data evaluated in these RCTs. Combination testing had significantly higher detection rates of neoplasms compared to FOBT alone, even with low FS compliance. There was insignificant additional diagnostic benefit from adding FOBT to FS, compared to FS alone. Participant compliance for FS in combination testing with FOBT was low compared to FOBT alone or FS alone. This low compliance was associated with acceptability issues and participants knowing their FOBT results prior to undergoing FS.

The evidence considered has not altered the conclusions made by the Working Party in regards to FS screening, which are also applicable to combined FOBT and FS screening as a first-line screening strategy. Such a strategy cannot be justified nor is it feasible based on the available evidence.

Cost-effectiveness of different screening modalities

Fifteen eligible primary research papers and three eligible secondary research papers were identified. The search identified eight eligible primary research papers comparing FOBT screening to no screening, three eligible primary research papers comparing guaiac-based FOBT to immunochemical FOBT, four eligible primary research studies comparing flexible sigmoidoscopy to no screening, and one eligible primary research study comparing flexible sigmoidoscopy plus FOBT to no screening. Most of these studies were heavily based on the three major RCTs but many also relied on previously published data, including results of autopsy studies. Only one study reported directly on the costs and outcomes of a RCT as recorded during the trial. Other studies relied on estimates, particularly of costs, and simulated longer term results using Marcovian and microsimulation models.

Most of the evidence currently available on the cost-effectiveness of CRC screening is based on comparing guaiac-based FOBT with no screening. The availability of major RCT data for this screening modality has led to a good number of economic studies with generally favourable cost-effectiveness estimates for guaiac-based FOBT. There was little evidence of the cost-effectiveness of other screening modalities: RCT data was lacking, published economic evaluations were scarce, economic models relied more heavily on assumptions, and results and conclusions vary widely.

The major weakness of the economic studies identified by this review was considerable uncertainty regarding the clinical data used to generate cost-effectiveness estimates. Uncertainty relating to the natural history of disease and the adenoma-carcinoma sequence, as well as uncertainty regarding the effectiveness of some screening modalities, resulted in significant limitations to the conclusions that could be drawn from the published literature. In addition to these problems, none of the studies provided a complete accounting of the costs of a CRC screening programme, including the costs of administrative overhead, health promotion, and recall systems. Another area of considerable uncertainty was the participation rate, although several studies demonstrated that achieving targeted participation may be key to a favourable cost-effectiveness ratio.
Conclusions

The following conclusions are based on the current evidence available from this report’s critical appraisal of literature published since 1997 on the published international evidence for the effectiveness and cost-effectiveness of screening tests for colorectal cancer:

High-quality evidence was found that FOBT screening with the guaiac FOBT Haemoccult reduces mortality from CRC. Specifically, evidence from ongoing follow-up for three major RCTs suggests that this mortality reduction has been sustained for the populations in which screening has stopped, but decreased slightly for the population to whom screening has continued to be offered. While less robust, the direct evidence available suggests that a reduction in rectal cancer may be achievable with the use of an immunochemical test. The evidence available since this topic was last considered in New Zealand by the Working Party on Screening for Colorectal Cancer (1998) has not changed substantially.

There is limited definitive evidence regarding superior immunochemical FOBT performance over the guaiac tests. However, evidence from cross-sectional studies suggests that the immunochemical test HemeSelect is comparable, or superior, to guaiac testing. Simplified FOBT tests may be more acceptable, as there was good evidence that the simplified testing process and sampling kit of the immunochemical test Insure encouraged greater participation rates. The conclusions on this topic should be revisited if further reliable evidence on the comparative performance of screening FOBTs becomes available.

International interest in establishing FOBT screening programmes for CRC remains high. Following a pilot screening programme conducted in five regions of the UK using a guaiac test, a national screening programme based on FOBT testing for CRC will be introduced by the National Health Service in England from April 2006. Australia is also evaluating FOBT screening in a pilot programme being conducted in three States using two types of immunochemical FOBTs.

No large sampled RCT has been completed that provided incidence and mortality data relating to flexible sigmoidoscopy screening. Three large ongoing trials are investigating flexible sigmoidoscopy as either one-off or repeated screening modalities for average-risk men and women aged from their mid-50s. Preliminary results are promising in terms of the feasibility and acceptability of this screening modality, but, long-term incidence and mortality data will not become available for these trials for three to seven years. While the introduction of FS in a national screening programme cannot currently be justified, this recommendation should be reviewed once the results of these trials are available. The generalisability of these findings to the New Zealand population would also need to be considered.

A combined screening strategy of FOBT and FS is not supported from the available literature published since 1997. There is insufficient evidence demonstrating that combined FOBT and FS as a first-line screening strategy in asymptomatic middle-aged populations brings about a greater reduction in CRC incidence and mortality compared to either modality alone. Data on CRC incidence and mortality from a Norwegian trial is expected to become available in late 2007 and this will provide much-needed evidence on the clinical effectiveness of combined screening compared to FS only screening strategies.

The strongest evidence of cost-effectiveness available is on the use of guaiac-based FOBT. This evidence is available from the results of major randomised controlled trials, including reports of directly observed cost information. It could be concluded from the low cost-effectiveness ratios estimated by these studies (generally less than $20,000 per life year saved) and the significant degree of consistency across studies in finding such a programme to be cost-effective that a screening programme based on annual or biennial guaiac-based FOBT is likely to be cost-effective by commonly accepted standards at a cost-effectiveness threshold of $50,000 per life year saved. However, true costs are likely to be higher than reported due to an incomplete accounting for programme costs in the studies. There is, however, uncertainty surrounding the natural history of disease, the likely participation rate, and certain components of programme costs. The cost-effectiveness of immunochemical FOBT and flexible sigmoidoscopy are subject to significantly more uncertainty than guaiac-based FOBT. These modalities have not been the subject of major randomised controlled trials and, consequently, evidence of their effectiveness and of the cost of running a screening programme based on these modalities is weak.
Although most cost-effectiveness studies suggested that screening for CRC is likely to be more cost-effective at saving life-years than cervical cancer screening or breast cancer screening, several studies pointed out that a CRC screening programme based on any screening modality will put additional pressure on the health system to increase the availability of colonoscopy and services used to investigate or follow-up positive results.

MeSH headings

Medical subject headings (MESH) used were exp colorectal neoplasms, mass screening, occult blood, sigmoidoscopy, exp immunological tests, guaiac, and immunochemistry. These were supplemented with free text to expand these concepts and ensure full coverage of variations in vocabulary and indexing between databases. Fuller details of search terms are given in the main body of the report and the complete search strategies in Appendix 2.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .........................................................................................................................1
DISCLAIMER ........................................................................................................................................1
COPYRIGHT ...........................................................................................................................................I
CONTACT DETAILS .............................................................................................................................II
EXECUTIVE SUMMARY ....................................................................................................................III
TABLE OF CONTENTS .........................................................................................................................IX
LIST OF TABLES ....................................................................................................................................XI
LIST OF FIGURES ..................................................................................................................................XII
LIST OF ABBREVIATIONS AND ACRONYMS .......................................................................................XIII
GLOSSARY .............................................................................................................................................XV

CHAPTER 1: INTRODUCTION

1

BACKGROUND .........................................................................................................................................1
SCREENING - GENERAL OVERVIEW .................................................................................................3
SCREENING FOR COLORECTAL CANCER ............................................................................................4
REVIEW SCOPE .......................................................................................................................................18
AIM ..........................................................................................................................................................19
REVIEW QUESTIONS ..........................................................................................................................19
STRUCTURE OF REPORT .....................................................................................................................20

CHAPTER 2: METHODOLOGY

21

2.1 STUDY SELECTION ..........................................................................................................................21
2.2 SEARCH STRATEGY ..........................................................................................................................24
2.3 STUDY SELECTION ..........................................................................................................................25
2.4 APPRAISAL OF STUDIES ................................................................................................................26
LIMITATIONS OF THE REVIEW ..........................................................................................................28

CHAPTER 3: EFFECTIVENESS OF FOBTS

29

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY ....................................................................29
SECONDARY RESEARCH ......................................................................................................................49
CONCLUSIONS .......................................................................................................................................51

CHAPTER 4: ACCURACY OF FOBTS

63

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY ....................................................................63
SECONDARY RESEARCH ......................................................................................................................74
CONCLUSIONS .......................................................................................................................................75

CHAPTER 5: EFFECTIVENESS OF FLEXIBLE SIGMOIDOSCOPY

81

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY ....................................................................81
SECONDARY RESEARCH ......................................................................................................................90
CONCLUSIONS .......................................................................................................................................91

CHAPTER 6: EFFECTIVENESS OF SCREENING WITH FOBT AND FS COMBINED

93

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY ....................................................................93
ONGOING TRIALS ..................................................................................................................................97
SECONDARY RESEARCH ......................................................................................................................101
CONCLUSIONS .......................................................................................................................................102

CHAPTER 7: COST-EFFECTIVENESS OF CRC SCREENING

105

PRIMARY RESEARCH STUDY DESIGNS AND QUALITY ....................................................................105
FOBT COMPARED WITH NO SCREENING: APPRAISAL OF STUDIES .............................................106
GUAIAC-BASED FOBT COMPARED WITH IMMUNOCHEMICAL FOBT: APPRAISAL OF STUDIES ....115
FLEXIBLE SIGMOIDOSCOPY COMPARED WITH NO SCREENING: APPRAISAL OF STUDIES ............120
CHAPTER 8: DISCUSSION
SUMMARY OF EVIDENCE
CONCLUSIONS
REFERENCES
APPENDIX 1
METHODOLOGY FOR CALCULATING DIAGNOSTIC TEST PERFORMANCE
APPENDIX 2
SEARCH STRATEGIES
SEARCHES FROM OTHER SOURCES
APPENDIX 3
SOURCES SEARCHED
APPENDIX 4
RETRIEVED STUDIES EXCLUDED FOR REVIEW
APPENDIX 5
RETRIEVED STUDIES APPRAISED FOR REVIEW
APPENDIX 6
ACCURACY OF FOBT AND FS COMBINED SCREENING
CONCLUSIONS
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Colorectal cancer data (per 100,000) in New Zealand, 2002</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>Criteria for assessing screening programmes</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>Duke’s classification</td>
<td>5</td>
</tr>
<tr>
<td>4.</td>
<td>PICO criteria for review questions</td>
<td>20</td>
</tr>
<tr>
<td>5.</td>
<td>Evidence dimensions</td>
<td>21</td>
</tr>
<tr>
<td>6.</td>
<td>Designations of levels of evidence</td>
<td>22</td>
</tr>
<tr>
<td>7.</td>
<td>Classification of economic evaluations of health care</td>
<td>22</td>
</tr>
<tr>
<td>8.</td>
<td>Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening</td>
<td>29</td>
</tr>
<tr>
<td>9.</td>
<td>Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening</td>
<td>53</td>
</tr>
<tr>
<td>10.</td>
<td>Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC</td>
<td>67</td>
</tr>
<tr>
<td>11.</td>
<td>Secondary research appraised relevant to accuracy of immunochemical FOBT screening for CRC</td>
<td>76</td>
</tr>
<tr>
<td>12.</td>
<td>Primary research studies appraised investigating the effectiveness of FS screening on outcomes from CRC</td>
<td>83</td>
</tr>
<tr>
<td>13.</td>
<td>Secondary research appraised relevant to the effectiveness of FS screening for CRC compared with no screening</td>
<td>92</td>
</tr>
<tr>
<td>14.</td>
<td>Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT screening alone</td>
<td>99</td>
</tr>
<tr>
<td>15.</td>
<td>Secondary research appraised relevant to effectiveness of FOBT and FS combined screening on outcomes for CRC</td>
<td>104</td>
</tr>
<tr>
<td>16.</td>
<td>Primary research appraised relevant to cost-effectiveness of FOBT screening compared with no screening</td>
<td>127</td>
</tr>
<tr>
<td>17.</td>
<td>Primary research appraised relevant to cost-effectiveness of Guaiac-based FOBT screening compared with immunological FOBT screening</td>
<td>129</td>
</tr>
<tr>
<td>18.</td>
<td>Primary research appraised relevant to cost-effectiveness of flexible sigmoidoscopy alone</td>
<td>130</td>
</tr>
<tr>
<td>19.</td>
<td>Primary research appraised relevant to cost-effectiveness of flexible sigmoidoscopy with FOBT</td>
<td>132</td>
</tr>
<tr>
<td>20.</td>
<td>Assessment of validity of a diagnostic test</td>
<td>151</td>
</tr>
<tr>
<td>21.</td>
<td>Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT only screening</td>
<td>182</td>
</tr>
<tr>
<td>22.</td>
<td>Primary research studies appraised investigating FOBT and FS combined screening for CRC</td>
<td>185</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. The colorectal screening pathway .....................................................................................17
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>Chi squared test</td>
</tr>
<tr>
<td>95% CI</td>
<td>95 percent confidence interval</td>
</tr>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality (formerly AHCPR) (USA)</td>
</tr>
<tr>
<td>AHTAC</td>
<td>Australian Health Technology Advisory Committee</td>
</tr>
<tr>
<td>AUD</td>
<td>Australian dollars</td>
</tr>
<tr>
<td>CDN</td>
<td>Canadian dollars</td>
</tr>
<tr>
<td>CEA</td>
<td>cost-effectiveness analysis</td>
</tr>
<tr>
<td>CG</td>
<td>control group</td>
</tr>
<tr>
<td>CINAHL</td>
<td>Cumulative Index to Nursing and Allied Health Literature</td>
</tr>
<tr>
<td>CRC</td>
<td>colorectal cancer</td>
</tr>
<tr>
<td>CT</td>
<td>computerised tomography</td>
</tr>
<tr>
<td>Cum Se</td>
<td>cumulative sensitivity (after primary screening and re-screening)</td>
</tr>
<tr>
<td>DALY</td>
<td>disability-adjusted life years</td>
</tr>
<tr>
<td>DCBE</td>
<td>double contrast barium enema</td>
</tr>
<tr>
<td>DKK</td>
<td>Danish Krone</td>
</tr>
<tr>
<td>DPCP</td>
<td>detectable pre-clinical phase</td>
</tr>
<tr>
<td>FNR</td>
<td>false negative rate</td>
</tr>
<tr>
<td>FOBT</td>
<td>faecal occult blood test</td>
</tr>
<tr>
<td>FPR</td>
<td>false positive rate</td>
</tr>
<tr>
<td>FS</td>
<td>flexible sigmoidoscopy</td>
</tr>
<tr>
<td>GP</td>
<td>general practitioner</td>
</tr>
<tr>
<td>HTA</td>
<td>health technology assessment</td>
</tr>
<tr>
<td>ICER</td>
<td>incremental cost-effectiveness ratio</td>
</tr>
<tr>
<td>ICT</td>
<td>immunochemical test</td>
</tr>
<tr>
<td>IHC</td>
<td>ischaemic heart disease</td>
</tr>
<tr>
<td>ITT</td>
<td>intention to treat analysis</td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Health (NZ)</td>
</tr>
</tbody>
</table>
MRI – magnetic resonance imaging
NHC – National Health Committee (NZ)
NNS – number needed to screen
NHS – National Health Service (UK)
NPV – negative predictive value
NSU – National Screening Unit (NZ)
NZ – New Zealand
NZD – New Zealand dollars
NZHTA – New Zealand Health Technology Assessment
OECD – Organisation for Economic Cooperation and Development
R – odds ratio
p.a. – per annum
PCR – polymerase chain reaction
PPV – positive predictive value
QALY – quality-adjusted life years
RCT – randomised controlled trial
RR – relative risk
RRR – relative risk reduction
SD – standard deviation
SEER – Surveillance, Epidemiology and End Results registry
Se – sensitivity
SG – screening group
Sp – specificity
TPR – true positive rate
UK – United Kingdom
US – United States
WHO – World Health Organization
USD – United States dollars
GLOSSARY

**Asymptomatic** - Asymptomatic people are those who do not have one or more symptoms (e.g., rectal bleeding) that may be due to a disease (e.g., colorectal cancer).

**Before-and-after study** - A situation in which the investigator compares outcomes before and after the introduction of a new intervention.

**Bias** - Deviation of results or inferences from the truth, or processes leading to such deviation. Any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth.

**Biopsy** - In a bowel lesion biopsy, a sample of tissue is removed to be examined under a microscope, as an aid to diagnosis.

**Blinded study** - A study in which observers and/or subjects are kept ignorant of the group to which they are assigned. When both observers and subjects are kept ignorant, the study is referred to as double blind.

**Cancer** - A general term for a large number of diseases that all display uncontrolled growth and spread of abnormal cells (also called malignant tumours). Cancer cells have the ability to continue to grow, invade and destroy surrounding tissue, and leave the original site and travel via the lymph or blood systems to other parts of the body where they may establish further cancerous tumours.

**Case control study** - An epidemiological study involving the observation of cases (persons with the disease, such as colorectal cancer) and a suitable control (comparison, reference) group of persons without the disease. The relationship of an attribute to the disease is examined by comparing the past history of the people in the two groups with regard to how frequently the attribute is present.

**Cohort study** - An epidemiological study in which subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed in different degrees, to a factor or factors hypothesised to influence the probability of occurrence of a given disease or other outcome. Studies usually involve the observation of either a large population, or for a prolonged period (years), or both.

**Confidence interval** - The computed interval with a given probability, e.g. 95%, that the true value of a variable such as a mean, proportion, or rate is contained within the interval. The 95% CI is the range of values in which it is 95% certain that the true value lies for the whole population.

**Confounder** - A third variable that indirectly distorts the relationship between two other variables, because it is independently associated with each of the variables.

**Con founding** - A situation in which the measure of the effect of an exposure on risk is distorted because of the association of exposure with other factor(s) that influence the outcome under study.

**Cost-benefit analysis** - A form of economic evaluation in which an attempt is made to value the consequences or benefits of a medical intervention in monetary terms so that these may be compared with the costs.

**Cost-effectiveness analysis (CEA)** - A form of economic evaluation in which the consequences or benefits of medical interventions are measured in terms of an appropriate health effect, such as life years saved, without placing a monetary value on such effects. These are balanced against the monetary cost of the intervention.

**Cost-minimisation analysis** - A form of economic evaluation in which it can be shown that outcomes are identical and, therefore, only costs are compared.
**Cost-utility analysis** - A form of economic evaluation in which the consequences or benefits of medical interventions are adjusted by health state preferences or utility weights, such as in quality-adjusted life years (QALYs) or disability-adjusted life years (DALYs).

**Coverage** - The number, percent, or proportion of eligible people reached by a programme.

**Cross-sectional study** - A study that examines the relationship between diseases (or other health related characteristics), and other variables of interest as they exist in a defined population at one particular time.

**Day patient** - A person who is admitted and discharged from hospital on the same day.

**Descriptive study** - A study concerned with, and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses.

**Diagnosis** - The process of identifying a disease by its characteristic signs, symptoms and findings on investigation.

**Diagnostic test efficacy** - The impact and usefulness of a diagnostic test expressed in terms of its technical properties.

**Effectiveness** - A measure of the extent to which a specific intervention, procedure, regimen, or service, when deployed in the field in routine circumstances, does what it is intended to do for a specified population.

**Efficiency** - The effects or end results achieved in relation to the effort expended in terms of money, resources and time. The extent to which the resources used to provide a specific intervention, procedure, regimen, or service of known efficacy and effectiveness are minimised.

**Elective services** - Non-urgent services for conditions which do not need immediate treatment. This includes services for patients with semi-urgent or non-life-threatening chronic conditions that tend to be stable or slowly deteriorate over time.

**Epidemiology** - The study of the distribution and determinants of health-related states or events in specified populations.

**Evidence-based** - Based on valid empirical information.

**Evidence table** - A summary display of selected characteristics (e.g., methodological design, results) of studies of a particular intervention or health problem.

**False negative result** - A negative test result in a person who does have the condition being tested for.

**False positive result** - A positive test result in a person who does not have the condition being tested for.

**Final truth determination** - Use of a reference standard to provide an accurate or “truth” diagnosis for verification of positive and negative diagnoses by a screening or diagnostic test (see also “reference standard”).

**Generalisability** - Applicability of the results to other populations.

**Grey literature** - That which is produced by all levels of government, academics, business and industry, in print and electronic formats, but which is not controlled by commercial publishers.

**Haustra** - Sacculations in the wall of the colon.

**High risk groups** - Usually refers to groups that have been identified as having a higher than average incidence of the disease in question.
**Histology** - The microscopic study of the minute structure and composition of tissues.

**Incidence** - The number of new events (cases, e.g. of disease) occurring during a certain period, in a specified population.

**Indicator** - An item of quantitative or qualitative information reported to enable the monitoring of a condition or the performance of an organisation.

**Intention to treat** - A method for data analysis in a randomised controlled trial in which individual outcomes are analysed according to the group to which they were randomised, even if they never received the treatment to which they were assigned.

**Māori** - The indigenous people of New Zealand.

**Matching** - The process of making a study group and a comparison group comparable with respect to extraneous factors.

**Mean** - Calculated by adding all the individual values in the group and dividing by the number of values in the group.

**Median** - Any value that divides the probability distribution of a random variable in half. For a finite population or sample the median is the middle value of an odd number of values (arranged in ascending order) or any value between the two middle values of an even number of values.

**Meta-analysis** - The process of using statistical methods to combine the results of different studies. The systematic and organised evaluation of a problem, using information from a number of independent studies of the problem.

**Misclassification** - The erroneous classification of an individual, a value, or an attribute into a category other than that to which it should be assigned.

**Morbidity** - Illness.

**Mortality rate** - The number of deaths from a specified disease that are diagnosed or reported during a defined period of time in a given population.

**Multiple regression** - Analysis of data that takes into account a number of variables simultaneously.

**Natural history** - The course of a disease from onset to resolution.

**Negative predictive value (NPV)** - The probability a person does not have the disease when the screening test is negative.

**Number needed to screen (NNS)** - The number of patients who would need to be screened, for a given period of time, in order to prevent a single event (i.e., death from colorectal cancer). The smaller the NNS is, signifies that fewer people need to be screened to prevent an event. The NNS often varies markedly with risk factors such as age.

**Odds ratio (OR)** - A measure of the degree or strength of an association. In a case control or a cross-sectional study, it is measured as the ratio of the odds of exposure (or disease) among the cases to that among the controls.

**Opportunistic screening** - The key feature that distinguishes opportunistic screening from screening programmes is the lack of a quality process, including routine monitoring and evaluation. Opportunistic screening usually occurs when a person who is presenting to the health system for another reason is asked a question or offered a test in order to detect the presence or confirm the absence of a specific condition. Opportunistic screening may be organised to a greater or lesser degree. However, because there are no attendant quality processes, its safety, effectiveness and cost-effectiveness cannot be assessed and guaranteed.
Outpatient - A person who goes to a health care facility for a consultation, and who leaves the facility within three hours of the start of the consultation. An outpatient is not formally admitted to the facility.

Population-based screening programme - A population-based screening programme is one in which screening is systematically offered by invitation to a defined, identifiable population: this requires a means of identifying and inviting the target population, for example through a population register.

Population screening programmes - Population screening programmes involve screening entire populations or a large and easily identifiable group within a population. The target population group for screening may be defined geographically or by some other characteristics such as gender, age or ethnicity. The New Zealand cervical and breast screening programme are examples of population screening programmes.

Positive predictive value (PPV) - The probability that a person actually has the disease when the screening test is positive.

Power - The ability of a study to demonstrate an association if one exists.

Prevalence - The number of events in a given population at a designated time (point prevalence) or during a specified period (period prevalence).

Primary care - First contact, continuous, comprehensive and coordinated care provided to individuals and populations undifferentiated by age, gender, disease or organ system.

Providers - Organisations and health professionals providing health services.

Random sample - A sample that is arrived at by selecting sample units such that each possible unit has a fixed and determinate probability of selection.

Randomised controlled trial - An epidemiologic experiment in which subjects in a population are randomly allocated into groups to receive or not receive an experimental preventive or therapeutic procedure, manoeuvre, or intervention. Randomised controlled trials are generally regarded as the most scientifically rigorous method of hypothesis testing available in epidemiology.

Reference standard - An independently applied test that is compared to a screening or diagnostic test being evaluated in order to verify the latter’s accuracy. A reference standard, therefore, provides an accurate or “truth” diagnosis for verification of positive and negative diagnoses. It is sometimes described as providing “final truth determination”.

Relative risk (RR) - The ratio of the risk of disease or death those exposed to the risk compared to the risk among those unexposed. It is a measure of the strength or degree of association applicable to cohort studies and RCTs.

Relative risk reduction (RRR) - The proportional reduction in rates of events between experimental and control participants in a trial. If there was an increase in the rate of events in the experimental group, the term would then be relative risk increase.

Risk factor - An exposure or aspect of personal behaviour or lifestyle, which on the basis of epidemiologic evidence is associated with a health-related condition.

Screening - Screening is the examination of asymptomatic people in order to classify them as likely or unlikely to have the disease that is the object of screening. The aim of screening is to detect disease before it is clinically apparent, and for this to improve the outcome for people with the disease.

Secondary care - Surgical and medical services that are generally provided in a hospital setting. In many cases, access to these services is by referral from a primary care health professional such as a general practitioner.
Selection bias - Any error in selecting the study population such that the people who are selected to participate in a study are not representative of the reference population or, in analytic studies the comparison groups are not comparable.

Sensitivity analysis - A method to determine the robustness of an assessment by examining the extent to which results are affected by changes in methods, values of variables, or assumptions.

Sensitivity (Se) - Sensitivity is the proportion of truly diseased persons in a screened population who are identified as diseased by a screening test. Sensitivity is a measure of the probability of correctly diagnosing a case, or the probability that any given case will be identified by the test.

Specificity (Sp) - The proportion of truly non-diseased persons who are so identified by a screening test. It is a measure of the probability of correctly identifying a non-diseased person with a screening test.

Surveillance, Epidemiology and End Results (SEER) registry - A set of geographically defined, population-based, central cancer registries in the United States, operated by local non-profit organisations under contract to the National Cancer Institute (NCI). Registry data are submitted electronically without personal identifiers to the NCI on a biannual basis, and the NCI makes the data available to the public for scientific research.

Symptomatic - Symptomatic people are those who have one or more symptoms (e.g., rectal bleeding) that may be due to a disease (e.g., colorectal cancer).

Systematic review - Literature review reporting a systematic method to search for, identify and appraise a number of independent studies.

True negative - A test correctly identifies a person without the disease.

True positive - A test correctly identifies a person with the disease.

Tumour - An abnormal growth of tissue.
Chapter 1: Introduction

BACKGROUND

Need for the proposed systematic review

At present, there is no routine organised population screening for colorectal cancer (CRC) in New Zealand. In 2005, the New Zealand Ministry of Health’s National Screening Unit (NSU) plans to formally revisit the advice of the Working Party on Screening for Colorectal Cancer’s (1998) on population screening for CRC. The NSU asked NZHTA to undertake a systematic review of the evidence since 1997 regarding population screening for CRC in order to update the 1998 review. The purpose of the systematic review is to provide information to assist the NSU in considering whether there should be an organised screening programme in New Zealand for colorectal cancer.

Colorectal cancer in New Zealand

Colorectal cancer, often referred to as large bowel cancer, is a malignant disease that starts in the colon or rectum. It is an important health concern in New Zealand. For the year 2002, the age-standardised incidence and mortality rates of CRC, as well as one-year and five-year period prevalence rates, are shown in Table 1 below. In recent health statistics (2000) colorectal cancer was the second most common cause of cancer registrations for both males (after prostate cancer) and females (after breast cancer). Furthermore CRC was the second most common cause of cancer death in males (after lung cancer) and the third most common cause of cancer death in females (after breast cancer and lung cancer) (New Zealand Health Information Service, 2004). At present, New Zealand’s death rate from colorectal cancer is the highest of all the OECD1 countries (Ministry of Health, 2003). The lifetime risk of developing CRC is 5.9 percent by age 75 years in New Zealand (New Zealand Guidelines Group [NZGG], 2004).

Table 1. Colorectal cancer data (per 100,000) in New Zealand, 2002

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence [Age-standardised rate (world)]</td>
<td>53.0</td>
<td>42.2</td>
</tr>
<tr>
<td>Mortality rate [Age-standardised rate (world)]</td>
<td>23.2</td>
<td>18.6</td>
</tr>
<tr>
<td>1-year period prevalence</td>
<td>972</td>
<td>908</td>
</tr>
<tr>
<td>5-year period prevalence</td>
<td>3974</td>
<td>3707</td>
</tr>
</tbody>
</table>


It has been predicted that CRC incidence and mortality in New Zealand would peak in the late 1990s and then decline from the second decade of the 21st century (Cox, 1995; Hodgen et al., 2002). The predicted decline is explained by a reduced risk of colorectal cancer in younger generations of New Zealanders, which may be due to factors such as dietary changes (Cox, 1995; Working Party on Screening for Colorectal Cancer, 1998). However, for both males and females, the actual number of colorectal cancer registrations is predicted to increase, due to a continuing growth in population size and an aging population (Hodgen et al., 2002).

---

1 Organisation for Economic Co-operation and Development. The 24 OECD countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and the United States.
New Zealand has a heterogeneous population. Reported ethnicity groups (based upon single ethnic group identification) at the last census in 2001, resident population counts were approximately: 70 percent European; 7.9 percent Māori; 4.5 percent Pacific Peoples; 5.7 percent Asian; 0.5 percent ‘other’; 7.5 percent more than one ethnic group from previous categories, and; 3.8 percent ethnicity not stated (http://www.stats.govt.nz/census.htm, accessed 06 September 2004). Similar to other colonised countries, the indigenous people have poorer health status than others, including for many types of cancer (Ministry of Health, 2003). Interestingly, registration rates of colorectal cancer for Māori (20.8 per 100,000 population) are lower than for non-Māori (53.6 per 100,000 population), but some caution must be taken in interpreting this data, as ethnicity is not specified accurately on all cancer registrations (New Zealand Health Information Service, 2004). Additionally, data from the 1990s indicated that although the likelihood of Māori being diagnosed with CRC compared to non-Māori was lower, the risk of dying from this disease was not significantly different between the two groups (Hodgen et al., 2002). This suggests that factors such as more advanced stage of disease at diagnosis (with corresponding poorer prognostic outcome) may contribute to the risk of Māori dying from CRC.

There is a governmental obligation in New Zealand to ensure that Māori health needs are considered with respect to the Treaty of Waitangi (Ministry of Health, 2003), therefore issues specific to Māori health needs and outcomes must be clarified when planning strategies to tackle CRC in New Zealand. Despite lower Māori registration rates for CRC, the mortality data clearly suggest that this disease is of concern to both Māori and non-Māori population groups.

**New Zealand Working Party on Population Screening for CRC**

In late 1996, after two population-based randomised controlled trials (RCTs) demonstrated a reduction in mortality from CRC as a result of screening with faecal occult blood tests (FOBTs), the National Health Committee of New Zealand convened an independent working party. The objectives of this working party were to:

- review the evidence for benefits and risks associated with the introduction of population screening for CRC
- identify the economic and resource implications of introducing a CRC screening programme and its likely acceptability
- report to the National Health Committee on issues surrounding population screening for CRC and make recommendations on the introduction of a screening programme in New Zealand or other actions that should be taken to reduce deaths from CRC in New Zealand.

To meet these objectives, the working party critically evaluated the evidence from the published literature up to May 1998. This process was aided in part by a systematic review of colorectal cancer screening published by the Australian Health Technology Advisory Committee (AHTAC, 1997). In November 1998, the working party published a comprehensive report with evidence-based recommendations. The main recommendations included that:

- Given the modest potential benefit, the considerable commitment of health sector resources and the small but real potential for harm, population-based screening for CRC with FOBTs is not recommended in New Zealand.

- Population-based screening for CRC with other modalities, such as flexible sigmoidoscopy, colonoscopy or double-contrast barium enema, is also not recommended as there is not yet evidence from randomised controlled trials that screening with any of these modalities reduces a reduction on CRC mortality.

- These decisions should be reviewed as evidence of benefit from new faecal occult tests and other screening modalities becomes available (Working Party on Screening for Colorectal Cancer, 1998).
The Working Party on Screening for Colorectal Cancer (1998) also identified that the implications for New Zealand, particularly those for colonoscopy services, of introducing an FOBT screening programme in New Zealand were beyond resources at that time.

Groups in New Zealand at increased risk of colorectal cancer

Another recommendation of the working party was that wider consultation should be undertaken and further consideration should be given to develop appropriate surveillance recommendations for groups identified to be at increased risk of CRC (Working Party on Screening for Colorectal Cancer, 1998). Following that recommendation, the New Zealand Guidelines Group published an evidence-based best practice guideline for the surveillance and management of such groups in May 2004. This guideline defined the four major groups at increased risk for CRC as those with:

1. personal history of CRC
2. colorectal adenoma
3. inflammatory bowel disease
4. family history of CRC (this may increase an individual’s lifetime risk of developing CRC, and is determined by the number of affected first-degree relatives and the age at which they were diagnosed with CRC) (NZGG, 2004).

Since the evidence for the surveillance and management of higher-risk groups has been recently reviewed, it has not been considered within the current systematic review. The current review focuses on the population screening (rather than surveillance) of asymptomatic, average-risk individuals.

SCREENING - GENERAL OVERVIEW

The focus of this review is secondary prevention, specifically screening. The National Health Committee (NHC) of New Zealand defines screening as “… a health service in which members of a defined population, who do not necessarily perceive they are at risk of, or are already affected by, a disease or its complications, are asked a question or offered a test to identify those individuals who are more likely to be helped than harmed by further tests or treatments to reduce the risk of disease or its complications” (NHC, 2003, p29). Reduction in mortality as a result of a screening programme partly depends on high levels of coverage of the population. Criteria to inform the assessment of screening programmes in New Zealand have been developed by the NHC (2003). These criteria are listed in Table 2 below.
Table 2. Criteria for assessing screening programmes

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The condition is a suitable candidate for screening.</td>
</tr>
<tr>
<td>2.</td>
<td>There is a suitable test.</td>
</tr>
<tr>
<td>3.</td>
<td>There is an effective and accessible treatment or intervention for the condition identified through early detection.</td>
</tr>
<tr>
<td>4.</td>
<td>There is high quality evidence, ideally from randomised controlled trials, that a screening programme is effective in reducing mortality or morbidity.</td>
</tr>
<tr>
<td>5.</td>
<td>The potential benefit from the screening programme should outweigh the potential physical and psychological harm caused by the test, diagnostic procedures and treatment.</td>
</tr>
<tr>
<td>6.</td>
<td>The health care system will be capable of supporting all necessary elements of the screening pathway, including diagnosis, follow-up and programme evaluation.</td>
</tr>
<tr>
<td>7.</td>
<td>There is consideration of social and ethical issues.</td>
</tr>
<tr>
<td>8.</td>
<td>There is consideration of cost-benefit issues.</td>
</tr>
</tbody>
</table>

SCREENING FOR COLORECTAL CANCER

To provide an overview of the issues relating to screening for CRC, screening for this condition has been considered with reference to each of the NHC (2003) criteria for assessing screening programmes (Table 2).

Criterion 1: The condition is a suitable candidate for screening

To justify any screening programme, the first prerequisite is that the disease being considered should be medically important, well-defined, and its prevalence reasonably well known (Grimes and Schulz, 2002). As discussed in the introductory section above, colorectal cancer in New Zealand appears to fulfil these criteria, and the burden of this disease affects both individuals and the greater community.

The opportunity for early detection and treatment of a malignant condition depends on the natural history of the disease. It is hypothesised that most colorectal cancers begin as adenomatous polyps, and progress (usually slowly, over many years) to carcinoma through an adenoma-carcinoma sequence. This hypothesis is supported by a large amount of circumstantial evidence (see Chapter 4, Working Party on Screening for Colorectal Cancer, 1998, for overview). Therefore, there does exist an early stage at which most CRC could, in theory, be detected or prevented from developing.

Although most colorectal adenomas do not progress to become malignant (AHTAC, 1997), adenomas are quite common. NZGG (2004) noted that in autopsy studies they have been detected in 30-40 percent of individuals aged over 60. Only a small minority of CRCs do not arise from adenomatous polyps, developing instead either de novo or from “flat” adenomas. This mainly occurs in individuals who are already at higher risk for CRC (Working Party on Screening for Colorectal Cancer, 1998). Hyperplastic bowel polyps, which are small lesions that can be distinguished microscopically (but not macroscopically) from adenomatous polyps, are generally considered as benign, with no direct implications for progression to CRC.

The most important factor in determining prognosis for patients with CRC is the stage at which the disease is diagnosed (Keating et al., 2003). The most widely used and simplest staging system for CRC is Duke’s classification (Table 3), which considers depth of bowel wall invasion, presence or absence of lymph node metastases, and presence or absence of distant metastases. Evidence from trials has shown that screening asymptomatic individuals can alter the distribution of stage at diagnosis (Working Party on Screening for Colorectal Cancer, 1998). Thus, documentation of the stage at diagnosis of cancer is an important surrogate marker for the success of screening (Keating et al., 2003). A recent local study by Keating et al. (2003) of the epidemiology of CRC in New Zealand found that 14 percent were staged as Duke’s A at diagnosis, 43 percent at Duke’s B and 43 percent at Duke’s C, and that this was in keeping with stage distributions found in other unscreened populations.
Table 3. Duke’s classification (adapted from Working Party on Screening for Colorectal Cancer, 1998)

<table>
<thead>
<tr>
<th>Duke’s Stage</th>
<th>Description</th>
<th>Five-year survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duke’s A</td>
<td>Tumour confined to the bowel wall with no lymph node metastases</td>
<td>90%</td>
</tr>
<tr>
<td>Duke’s B</td>
<td>Tumour penetrating through the bowel wall to serosa or perirectal fat with no lymph node metastases</td>
<td>60-70%</td>
</tr>
<tr>
<td>Duke’s C</td>
<td>Lymph node metastases present</td>
<td>30-40%</td>
</tr>
<tr>
<td>Duke’s D*</td>
<td>Distant metastases present</td>
<td>Rare</td>
</tr>
</tbody>
</table>

* Not in Duke’s original description, added subsequently.

**Criterion 2: There is a suitable test**

With reference to a suitable test for detecting colorectal cancer, the test should ideally be:

- **safe**: harm is minimised
- **simple**: easy to perform and interpret
- **reliable**: the test is repeatable, and gives consistent results
- **valid**: the test is capable of measuring what it sets out to measure
- **highly sensitive**: high probability that the test will give a positive result when the person being screened has the condition. Sensitivity should be sufficient to lead to a substantial impact on the disease from a population perspective
- **highly specific**: high probability that the test will give a negative result when the person being screened is disease-free. This is important to minimise harms from false positive screening results (NHC, 2003).

Measuring screening test accuracy using sensitivity (Se) and specificity (Sp) provides important information when considering its use. Traditional definitions of Se and Sp are usually based on evaluations of tests at one point in time, with a reference standard test then used to determine the true disease status.

Predictive value measures are also frequently used as an additional method of evaluating screening test validity, and can be particularly useful in clinical settings (Grimes and Schulz, 2002). Predictive values measure whether or not an individual actually has the disease, given the results of the screening test. The positive predictive value (PPV) is the probability that a person truly has the disease, given the results of the screening test, while the negative predictive value (NPV) is the probability that a person is truly free of disease given a negative test result (Hennekens and Buring, 1987). Grimes and Schulz (2002) emphasise the importance of remembering that predictive measures vary with the underlying prevalence of disease in the population. Therefore, PPV will be much higher for a test if, for example, a population of symptomatic patients is tested, compared to an asymptomatic general population.

For colorectal cancer, there are several tests that have been considered as screening tools. These tests are briefly described below, and the issues surrounding the use of each test as a screening tool, including the main advantages and disadvantages, are discussed.

**Faecal occult blood tests**

Faecal occult blood tests (FOBTs) are tests for blood or blood products, with the presence of blood taken as an indicator of neoplasia (Young, St John et al., 2002). They require the collection of faecal matter, which is then applied to a testing kit. Screening using this modality is based on the observation that CRCs (as well as larger polyps) bleed intermittently, and that follow up investigations of positive FOBTs, such as colonoscopy, will lead to the detection of cancer (Working Party on Screening for Colorectal Cancer, 1998). Approximately two thirds of CRCs bleed in the course of a week (which
EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER

does limit the potential sensitivity of FOBTs) and benign lesions can also bleed into the bowel lumen, affecting specificity. FOBT is a non-specific tool that is aimed at providing information about the likelihood of the presence of colorectal cancer. However it is limited by its inability to inform clinicians where a possible abnormality may lie and so while FOBT is a non-invasive test, an abnormal FOBT in a screening situation will lead to more invasive testing methods such as colonoscopy.

There are two main types of FOBTs, guaiac and immunochemical. Guaiac tests detect blood in the form of intact haem molecules and haemoglobin through a reaction involving peroxidase, which produces a blue colour if positive. Although they can in theory detect bleeding from any part of the alimentary tract, they are somewhat more selective for the large bowel over the upper gastrointestinal tract. False positives can occur with guaiac FOBTs as a result of dietary factors and ingestion of certain medications, although this can be limited by pre-test restrictions in tested subjects (Ransahoff and Lang, 1997; Working Party on Screening for Colorectal Cancer, 1998). In their position paper on the interpretation and follow-up of FOBTs, Ransahoff and Lang (1997) argue that any person with a positive result who did not restrict diet or medications pre-test should still undergo diagnostic work-up, rather than resubmitting repeat FOBTs after diet and medication restrictions.

Guaiac-type FOBT is the only modality to date for which there is direct evidence from prospective RCTs that mortality from CRC can be reduced if used as a screening tool. Of this type of FOBT, the Haemoccult (also spelt Hemoccult) test, which was available from 1970 until 1977 and has now become Haemoccult II (Beckman Coulter, Inc, Fullerton, CA), is the most studied (Allison, 2003). The Working Party on Screening for Colorectal Cancer (1998) considered the results of a meta-analysis conducted by AHTAC (1997) using the results available from the two population-based screening trials that used Haemoccult/Haemoccult II as a biennial screening tool: the Funen RCT (Kronberg et al., 1996) and the Nottingham RCT (Hardcastle et al., 1996), which had mean follow-up periods of 10 years and 7.8 years respectively. The meta-analysis revealed a 16 percent (95% CI 6%-25%) mortality reduction from CRC in the population offered screening. Concerns centred on the relatively low sensitivity (around 50% per screening round) of the Haemoccult tests for CRC. Studies that have measured the sensitivity and specificity of these FOBTs for “once-only” use in asymptomatic individuals have found results of approximately 40 percent and 98 percent respectively. Rehydrating the Haemoccult test cards prior to developing greatly increased test sensitivity (to 60%), but decreased specificity to around 90 percent (Ransahoff and Lang, 1997; Wagner et al., 1996, in Pignone et al, 2000a).

Rehydrating the Haemoccult test was adopted as a strategy in the other RCT that was considered by the Working Party on Screening for Colorectal Cancer (1998), which was conducted in Minnesota using a volunteer population (Mandel et al., 1993). After 13 years of follow-up, the Minnesota trial reported a statistically significant mortality reduction from CRC of 33 percent (95% CI 13%-50%) in those offered annual screening, but no significant risk reduction for those offered biennial screening. The main detrimental consequence of FOBT rehydration is a decreased positive predictive value, with the associated increased test positivity rate (number of positive tests per population tested) requiring an increased number of follow-up investigations. Therefore, the costs of increased screening (Scholtefeld and Moss, 2002; Walsh and Terdiman, 2003). The increased screening costs are not just financial; they also include potential harm (perforation or haemorrhage) resulting from the extra colonoscopies performed. At present, rehydration of guaiac based FOBTs is therefore not recommended by major organisations such as the World Health Organization (Young, St John et al., 2002), the US Preventive Services Task Force (Pignone, 2002b), or in the guidelines of the American Gastroenterological Association (Winawer et al., 2003). In addition, rehydration of the Haemoccult test has never been endorsed by the manufacturer (Allison, 2003).

Historical concerns that colonic neoplasms may bleed intermittently formed the basis of the development of the standard testing protocol for guaiac tests. Two smears of faecal material are applied to each FOB test card from two different areas of the same stool, and this is done on three consecutive days (Ransahoff and Lang, 1997). This results in completion of three test cards, comprising six samples in total. These cards can then be mailed to a central testing facility for development and interpretation. Test cards are ‘developed’ using the reagent applied to all six testing windows, which results in a blue colour in the presence of the peroxidase reaction. Interestingly, different definitions of a positive overall FOBT result were adopted in the Funen and Nottingham RCTs – i.e., how many slides out of six must turn blue, and whether to repeat tests with less than a certain number of blue slides. Ransahoff and Lang (1997) found little evidence to support particular
testing or re-testing strategies, and state clearly that a positive result should be defined as one or more positive windows.

Immunochemical FOBTs detect antibodies specific to human haemoglobin, albumin or other blood components in the faeces. In theory they eliminate false positives from dietary causes, and have a higher degree of selectivity for colorectal bleeding over gastric bleeding (Young, St John et al., 2002). In addition, they offer the attractive possibility of quantitated automated laboratory reading (Scholefield and Moss, 2002). Identified drawbacks include variation between tests, greater expense (two to five times more expensive) and the need for laboratory development resources (Working Party on Screening for Colorectal Cancer, 1998; Young, St John et al., 2002). Because there is direct RCT evidence for the guaiac FOBT ‘Haemoccult’, performance endpoints such as test sensitivity, specificity and predictive value could be used to compare available immunochemical tests with the guaiac type. To become acceptable, a clear demonstration would be required of an improved diagnostic profile. Young, St John et al. (2002) point out that such end points could be demonstrated in practice by a comparison of test positivity and lesion detection rates in screened populations.

Another strategy that received comment in the working party’s report was a two-step FOBT testing approach, using a newer, more sensitive guaiac test (HaemoccultSENSA), followed by the use of an immunochemical test called HemeSelect if the first test is positive. If both tests are positive, the overall test result is considered positive, and appropriate follow-up tests conducted. However, insufficient data was available at that time to the working party on the performance of the tests used in the two-step approach for a definitive judgement to be made about the efficacy of this method for CRC screening. The final conclusion made on FOBT performance was that at that time it remained to be determined which FOBT would achieve the optimal balance of sensitivity, specificity and cost in the context of population screening (Working Party on Screening for Colorectal Cancer, 1998).

Flexible sigmoidoscopy

Flexible sigmoidoscopy (FS) has emerged from the development of medical fibre optics. A 60cm flexible endoscope is used to enable direct visualisation of the rectum and left side of the colon to a maximum of 60cm from the anal verge (Working Party on Screening for Colorectal Cancer, 1998). Patient preparation typically involves the self-administration of two enemas on the morning of the examination, and the procedure is usually performed without sedation (Pignone et al., 2002a). As well as the advantage from being able to directly visualise the distal colon, FS also offers the potential for pathological sampling of lesions identified at the time of screening. However, it is a more invasive test than FOBT. Although flexible sigmoidoscopy is not currently widely used in New Zealand (NZGG, 2004), compared to standard colonoscopy the technique is simpler to learn, requires less bowel preparation, has lower risk of complications and is quicker to perform (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002; Atkin et al., 2001).

The main concern surrounding the use of FS is that only the distal portion of the colorectum is visualised, so only approximately 50-65 percent of lesions lie within the reach of the FS (Working Party on Screening for Colorectal Cancer, 1998; Atkin et al., 2001; Scholefield and Moss, 2002). However, up to 23 percent of patients with proximal cancers have concomitant adenomas or CRCs in the distal large bowel. If such distal lesions are identified by FS, a subsequent whole large bowel evaluation, for example by full colonoscopy, may increase detection of all asymptomatic CRCs to up to 70 percent (Working Party on Screening for Colorectal Cancer, 1998). Debates surround the question of exactly which distal colon findings on FS should trigger colonoscopy, and intervals for screening using this modality remain uncertain (Ransohoff, 2002; Pignone et al., 2002a; Working Party on Screening for Colorectal Cancer, 1998). With increasing age, the prevalence of advanced proximal neoplasia increases and therefore it is possible that strategies to increase detection of such abnormalities may be most cost effective when directed towards older populations (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002).

A small study conducted in Norway using FS as a CRC screening tool has shown a reduction in mortality (Hoff et al., 1996, in Working Party on Screening for Colorectal Cancer, 1998). It also may be more effective at preventing cancer than FOBT, because it can directly detect polyps (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002). Three major RCTs are currently underway investigating flexible sigmoidoscopy (see Chapter 5 for overview of these trials). These include two multicentre trials in the UK and Italy (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002;
Segnan et al, 2002) which are employing the same protocol to investigate a once only screen for people in the age range 55-64 years. Single screening by FS has been suggested as a potentially cost effective way to reduce incidence rates, because detection of distal adenomas increases with age until the late 50s before levelling off. This suggests that most people destined to develop distal colorectal cancer will develop distal adenoma before they reach 60 years of age (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002). The other trial, which is being conducted in the USA (Gohagan et al, 2000), is investigating five-yearly FS screening for people aged 55-74 years at entry as part of a wider screening programme that is also considering prostate, lung, and ovarian cancer screening.

Flexible sigmoidoscopy and faecal occult blood testing combined

To overcome the limitations of using either FS or FOBT alone to screen for CRC, combination FS and FOBT testing has also received attention as a screening strategy. A common recommendation for such combination testing is for participants to have an annual FOBT, and to undergo FS every five years (with the FOBT completed first in the year that FS is due) (Walsh and Terdiman, 2003). A non-randomised controlled trial has compared the annual use of rigid sigmoidoscopy (an older technique) to rigid sigmoidoscopy plus FOBT. This trial reported reduced mortality in the group screened with the combination of tests compared to the rigid sigmoidoscopy group after five to eleven years of follow up (Winawer et al., 1993, in Walsh and Terdiman, 2003). However, the evidence from this study was neither strong nor generalisable, as the size of the effect on the outcome of mortality was only marginally significant, and participant compliance was low. Despite this, interest has been sustained in a combined FOBT and FS screening strategy, and since 1997 several large RCTs have compared the performance characteristics and participant compliance of combined testing to either one or the other modality alone. Published studies have shown small improvements in diagnostic performance by adding FOBT to FS compared to FS alone, whereas the diagnostic performance of combined screening has improved significantly compared to FOBT alone. Participant compliance with FS screening in combined screening generally has been low. The effect of combined screening compared to either test alone on longer term health outcomes such as CRC incidence and mortality has not been adequately evaluated in RCTs. Because there is direct RCT evidence for the guaiac FOBT ‘Haemoccult’ test, performance endpoints could be used to compare combined FOBT/FS testing with FOBT alone. This can also be used to compare FOBT/FS testing with FS alone, once mortality data from current RCTs in the US and Europe becomes available, see previous section. For combined FOBT/FS testing to become acceptable a clear demonstration would be required of benefits to health outcomes with a reduction in CRC incidence and mortality compared to either test alone. Issues surrounding which polyps in the distal colon require colonoscopy at FS, the frequency of examinations, and compliance also remain to be resolved.

Colonoscopy

Colonoscopy uses a longer flexible fibre optic endoscope than FS, which enables direct visualisation of the entire large bowel, removal of samples for pathological assessment, and the option of polypectomy at time of examination. Patient preparation typically involves dietary restrictions and the administration of laxatives the day before the procedure, and the procedure is usually performed under conscious sedation (Pignone et al., 2002a). Colonoscopy is generally considered to be the gold standard investigation of the large bowel, and holds an established place in the diagnosis and treatment of CRC.

Arguments for the use of colonoscopy as a primary screening tool stem from its high sensitivity (AHTAC, 1997). Sensitivity of colonoscopy for CRC is as high as 95 percent and specificity up to 100 percent (Conseil d’Évaluation des Technologies de la Santé du Quebec, 2000; NZGG, 2004). However, the performance of this procedure is operator dependent; missed cancers may result if the examination does not view the entire colon when the operator does not recognise the examination was incomplete (Working Party on Screening for Colorectal Cancer, 1998; NZGG, 2004).

Indirect evidence from a re-analysis of data from the Minnesota FOBT RCT (that had a risk reduction in CRC mortality of 33 percent compared with control) where a high proportion (38%) of individuals underwent colonoscopy, suggested that some of the mortality reduction may have been attributable to chance selection of individuals for colonoscopy (Lang and Ransahoff, 1994, in Barkun et al., 2004). However, there have been no large population-based RCTs examining the efficacy of colonoscopy to reduce mortality from CRC when used as primary screening tool for this disease (Working Party on Screening for Colorectal Cancer, 1998; US Preventive Services Task Force, 2002; Barkun et al., 2004).
Data regarding efficacy and risks of colonoscopy came from studies of its use as either a diagnostic or therapeutic tool, which limits the direct relevance of such data to the screening situation (see Pignone et al., 2002a, for an overview). The risks of physical harm from colonoscopy, e.g., bowel perforation, colonic bleeding post-procedure or medication side effects, are higher than from either FOBT or FS, although risk decreases as colonoscopist experience increases (US Preventive Services Task Force, 2002).

Double contrast barium enema

To undertake a double contrast barium enema (DCBE), patients usually prepare with dietary restrictions plus enemas or laxatives the day before. On the day of the test, barium is instilled into the colorectum and then air insufflated, followed by x-ray examination in various positions (Conseil d’Évaluation des Technologies de la Santé du Quebec, 2000; Pignone et al., 2002a). It is one of the standard radiological techniques for investigating the large bowel. DCBE as a potential screening test for CRC was considered by the Working Party on Screening for Colorectal Cancer (1998). Potential advantages of this method are: an increased sensitivity over FOBT to detect polyps as well as cancers; an ability to visualise the entire colon compared to FS; and increased safety and lower cost compared to colonoscopy. The inability to take samples is one of the disadvantages of DCBE compared to colonoscopy, meaning that colonoscopic follow-up of suspicious lesions identified by DCBE is required.

As others have recently found (Conseil d’Évaluation des Technologies de la Santé du Quebec, 2000; Pignone et al., 2002a; Barkun et al., 2004), there have been no published RCTs on the efficacy of DCBE for reducing mortality from CRC in a screening situation. Studies to measure the accuracy of DCBE in diagnostic settings, which are of only limited applicability to a screening situation, have found sensitivity levels of 80-90 percent for cancer (see Pignone et al., 2002a, for overview).

Virtual colonoscopy

Virtual colonoscopy, using either computed tomography (CT) or magnetic resonance imaging (MRI) scanners, is another modality that has been considered more recently as potentially suitable for CRC screening (Pignone et al., 2002a; Walsh and Terdiman, 2003). This technique involves a similar bowel preparation to standard colonoscopy. Air or carbon dioxide is insufflated into the colon through a rectal tube and then data are acquired by the scanner which generates images of the colon. Dependent on the equipment used, the data from the scanners are presented as two-dimensional images, with three-dimensional images generated of areas identified as suspicious (Levin et al., 2003). Sedation is not required, although some mild patient discomfort may be experienced from the air insufflation (Walsh and Terdiman, 2003). After radiologist review of results, patients noted to have suspicious-looking polyps or a colonic mass need to proceed to conventional colonoscopy, to enable diagnostic biopsies to be taken.

Virtual colonoscopy has several potential advantages over other tests considered for CRC screening. It potentially can identify large-bowel malignancies that are often poorly assessed by conventional colonoscopy, such as those located within haustral folds, as well as being able to view the entire colon. For this reason it has also been suggested as the examination of choice for failed or incomplete colonoscopies in many settings (Levin et al., 2003). Virtual colonoscopy may also allow for small (and probably hyperplastic, low-risk) colonic polyps to be left in-situ when detected, and for regular reassessment to be carried out to monitor them (Levin et al., 2003). Future potential developments for this modality include the ‘patient-friendly’ possibility of avoiding pre-imaging bowel preparation. This is attained using either stool labelling with oral contrast, or electronic cleansing where the stool density relative to other tissues is calculated and subtracted during image generation (Levin et al., 2003).

As with any test, virtual colonoscopy also has disadvantages. False positives can occur as a result of retained stool in the bowel, diverticular disease (which can produce poorly distensible areas of the colon), or thickened bowel folds. It has been speculated that virtual colonoscopy may not be able to detect relatively rare flat adenomas, and, as with DCBE, polyps cannot be removed during the procedure (Levin et al., 2003). Variable factors such as radiologist experience and training may also influence accuracy (Pignone et al., 2002a; Walsh and Terdiman, 2003).
Recent reviews of screening for colorectal cancer that have included this modality have found that several studies evaluating the accuracy of virtual colonoscopy compared to conventional colonoscopy per se have been undertaken, with varying results (Pignone et al., 2002a; Walsh and Terdiman, 2003). No population-based RCTs evaluating the efficacy of virtual colonoscopy screening on CRC outcomes have been completed to date, and its performance has not yet been studied in typical screening populations (Levin et al., 2003). Pignone et al. (2002a) and Walsh and Terdiman (2003) have commented on a lack of information for this modality regarding adverse effects and acceptability in the general population. Furthermore, high resource costs for equipment and radiologist training are currently well-recognised barriers to widespread use of virtual colonoscopy.

Other potential screening modalities

Stool sampling for DNA mutations associated with the process of colorectal carcinogenesis, using amplification tests such as the polymerase chain reaction (PCR) to assist detection, is an emerging technology that holds promise as a non-invasive test for the detection of colorectal cancer (Walsh and Terdiman, 2003). Unlike blood, DNA is stable in the faeces, shed continuously by developing malignant lesions, requiring only one sample, and can be specifically attributed to large bowel neoplasms (Levin et al., 2003). These advantageous features could endow DNA stool sampling with the attractive test characteristics of both high sensitivity and specificity. Another possibility is that as a result of the stool sampling process it may be possible to detect mutated DNA that enters the proximal gastrointestinal tract from other malignant lesions, e.g., lung or stomach (Levin et al., 2003). Noted disadvantages of DNA stool sampling are that the tests are expensive and time-consuming, and that there are limited data currently available regarding the use of this technology in screening populations (Levin et al., 2003; Walsh and Terdiman, 2003).

**Criterion 3: There is an effective and accessible treatment or intervention for the condition identified through early detection**

It is important when contemplating a screening programme to know that there is an effective treatment for the condition identified by the screening process, and that the required health system resources are available to provide such treatment to patients with screen-detected disease. Management of CRC depends on the stage and histology at presentation, and disease detected at an earlier stage is more straightforward to manage. An overview of current treatment strategies for CRC is presented here.

**Localised disease**

With regards to CRC, treatment depends on the stage and histology at presentation. Polyps that have undergone early malignant change can often be removed endoscopically, if they are on a stalk. Further follow-up for such polyps depends on pathological findings and a calculation of the risk of cancer recurrence (NZGG, 2004). For CRC that cannot be managed endoscopically, surgery is the definitive treatment for localised disease, and is the only potentially curative treatment (Leslie and Steele, 2002). For optimal disease management in this situation it is important to establish at the outset the extent of the local disease, as well as confirm that distant spread has not already occurred. Radiological imaging of the liver and lungs is considered to be a minimum requirement for any patient in which elective surgical removal is being considered (Leslie and Steele, 2002; Scottish Intercollegiate Guidelines Network [SIGN], 2003).

The standard treatment for localised tumours is to surgically remove the involved segment of the bowel and the regional lymph nodes. The actual surgical technique employed depends on whether the malignancy is located in the colon or the rectum, with rectal surgery being more technically challenging (see Leslie and Steele, 2002, for an overview of different techniques).

Unfortunately, significant proportions of patients with CRC (up to 30%) present as emergencies (with symptoms of bowel obstruction, perforation or haemorrhage) and require urgent surgical management. Post-operative morbidity (19%) and mortality (8%) rates for such patients are increased compared with those managed electively, and survival reduced to 30% at five-years (Royal College of Surgeons Ireland, 2002).

Variance in outcome measures following the surgical management of CRC, such as peri-operative morbidity, mortality and long-term survival also occurs independent of disease or patient factors. This
is considered as being related to variance in surgical practice (Working Party on Screening for Colorectal Cancer, 1998; McArdle, 2000a). Sub-specialisation in colorectal surgery has evolved internationally and in New Zealand over the last decade, and may be associated with better post-operative patient outcomes for rectal cancer (Connelly et al., 2002; SIGN, 2003). Laparoscopic surgery to remove colorectal tumours is also feasible, and international evidence regarding its place continues to accrue (Connelly et al., 2002). Laparoscopic surgery may be associated with an improvement in post-operative outcomes such as pain and length of hospital stay, but it is not yet widely recommended for the routine management of CRC (Royal College of Surgeons Ireland, 2002; SIGN, 2003).

Adjuvant therapy

After surgical removal of the bowel segment and lymph nodes, pathological examination is necessary to classify the stage of the disease. As well as estimating likely prognosis, stage of CRC determines whether adjuvant therapy should be used as an additional weapon to surgery. The aim of adjuvant therapy is to eradicate any occult cancer cells, undetectable by current imaging techniques, which may have metastasised before surgery (Midgley and Kerr, 2000). The current mainstays of adjuvant therapy are chemotherapy and radiotherapy, however a detailed discussion of the specific regimens for each of these treatment modalities is beyond the scope of this review.

For chemotherapy, there is no evidence to support its use as an adjuvant treatment for patients with Duke’s A colonic or rectal cancer, and evidence for its routine use in Duke’s B disease is inconclusive (Royal College of Surgeons Ireland, 2002). For patients with Duke’s C disease of the colon conclusive evidence exists that adjuvant chemotherapy improves survival (in those who are otherwise fit enough to receive such treatment), and it may also confer a survival benefit for people with Duke’s C rectal tumours (Royal College of Surgeons Ireland, 2002; SIGN, 2003). Given the morbidity associated with the use of abdominal radiotherapy, adjuvant radiotherapy is only advised for use in rectal tumours (Leslie and Steele, 2002). However, there is strong evidence that the pre-operative use of radiotherapy confers beneficial local disease control for patients with bulky Duke’s C rectal tumours. Post-operative radiotherapy is an alternative strategy, recommended for certain patients with Duke’s C disease (Royal College of Surgeons in Ireland, 2002; SIGN, 2003).

Follow up of CRC patients

The management of patients who have undergone apparently curative surgical resection of localised CRC usually includes regular review (by either the surgical or primary care team) on an ongoing basis, although there is lack of consensus regarding the optimal intensity for such follow-up. The aims of such follow-up are to facilitate the early detection and treatment of recurrent or metastatic disease, as well as providing psychological support for the patient and to enable audit (McArdle, 2000b). Meta-analyses of evidence for formal follow-up to detect metastatic disease indicate that a survival advantage may be conferred. Computed tomography scanning and serial carcinoembryonic antigen levels appear to be the most useful modalities for this purpose, endoscopic surveillance may be beneficial and FOBT is of no value (SIGN, 2003). A recent economic evaluation based on a meta-analysis of follow-up in five RCTs also suggested that intensive follow-up of people with CRC is cost-effective compared to conventional follow-up methods (Renahan, et al., 2004). However, the definitions of ‘intensive’ and ‘conventional’ follow-up as evaluated by each RCT were not specified in the meta-analysis report, and may have varied (Renahan, et al., 2002).

Advanced disease

Advanced colorectal cancer is defined as disease that is either metastatic, or so locally advanced that a curative surgical procedure is unlikely to be attainable at time of presentation or recurrence (Young and Rea, 2000). Surgical resection may still be attempted for locally advanced CRC in patients for whom it is anticipated that quality of life may be improved, and palliative surgical procedures for symptoms such as bowel obstruction may also provide benefit (Leslie and Steele, 2002; SIGN, 2003). There is in addition some low-grade evidence that disease survival can be increased for certain patients by resection of discrete metastatic lesions in the liver or even the lung (SIGN, 2003). For otherwise fit patients with tumour metastases, there is good evidence that chemotherapy can improve survival, although the improvement is usually measurable in months only (SIGN, 2003). Evidence is now available to support the benefit of certain second-line chemotherapeutic regimes for patients whose disease progresses or relapses despite initial treatment (Royal College of Surgeons Ireland, 2002;
SIGN, 2003). Radiation therapy for advanced CRC is limited to management of invasive rectal tumours, and may also help with symptom control of painful bone metastases (Young and Rea, 2000). Finally, access to specialist palliative care services is recommended for all patients with terminal disease, to provide aid with both physical and emotional needs (SIGN, 2003).

Summary

Effective treatments for CRC exist, and disease detected at an earlier stage is clearly more straightforward to manage and potentially amenable to cure. New CRC treatment modalities are still being developed, and ongoing evidence for these and for new combinations of existing treatment regimes continues to be generated (Slevin and Payne, 2004). Treatment strategies for CRC potentially involve many sectors of the health system, including medical and nursing specialties within primary care, surgery, radiology, oncology, pathology and palliative care. This means that resourcing issues for co-ordinating treatment may be complex. Thus, although effective treatments are possible, the accessibility of such treatments for people diagnosed with CRC will depend on the economic and workforce capacity of the health system on which they depend.

The issue of whether there is evidence that earlier treatment of screen-detected disease, compared to symptom-detected disease, leads to better outcomes is complicated by potential biases and confounding factors that may affect treatment outcomes. These are discussed further in the next section. Essentially, only an RCT study design can effectively deal with such biases, and therefore information on CRC morbidity and mortality or other health outcomes from long-term prospective RCTs will a priori provide information on whether earlier treatment as a result of screening for CRC is beneficial.

**Criterion 4: There is high-quality evidence, ideally from randomised controlled trials, that a screening programme is effective in reducing mortality or morbidity**

Because of the potential impact of a screening programme, to assess whether health benefits exist for any particular screening modality the NHC (2003) specifies the need for high-quality evidence that screening reduces disease morbidity or mortality. The best level of evidence for screening comes from randomised controlled trials (RCTs), in which study participants are randomly allocated into intervention and control groups, and screening is offered only to those in the intervention group. Participants in both groups are followed up over the same time period, and the study analysis compares differences that occur between the study groups with respect to specific outcomes of interest, which are defined at the outset of the study.

An RCT is the study design most able to control for critical potential confounding and biases. A confounding factor is a third variable that indirectly distorts the relationship between exposure and outcome. The unique strength of randomisation in an intervention study with an adequate sample size is that it will virtually ensure that all potential confounding factors, both known and unknown to the investigator, will be evenly distributed among the groups under scrutiny (Hennekens and Buring, 1987). Potential biases resulting in non-comparability of screened and symptom-diagnosed cases of diseases are the chief threat to validity in studies that assess the efficacy of screening (Hennekens and Buring, 1987). Such potential biases are:

- lead-time bias: identifying a disease by screening will advance the date of diagnosis, and may therefore confer an apparent prolonged survival time, even if time of death is not delayed
- length bias: tumours grow at different rates, and so remain in the pre-symptomatic screen-detectable phase for different lengths of time, so that the probability of detecting slow-growing tumours is greater at each screening round than of detecting faster growing ones. This may lead to apparent differences in outcomes between screened and unscreened populations, since slow-growing tumours are associated with improved prognosis
- overdiagnosis bias: screening for disease may detect very early lesions that may never affect a person in his/her lifetime and this may lead to apparent differences in outcome between screened and unscreened groups
- selection bias: people who are offered and accept screening may systematically differ in their underlying risk of developing the disease, and this may over-estimate or under-estimate the effects of screening (Working Party on Screening for Colorectal Cancer, 1998).
An RCT with colorectal cancer mortality as the specific outcome of interest is the only study design not affected by these biases (Working Party on Screening for Colorectal Cancer, 1998). However, it is also important that the general quality of such RCTs is good. This means that measures are taken in the conduct and analysis of the study to maintain high internal validity. In particular, although it may not always be possible to blind participants to their allocated group, the investigators should ensure that assessment of study outcomes is blinded. Within reason, effort should be made to ensure that as few study participants from both intervention and control groups as possible withdraw from the study or are lost to follow-up, and that unavoidable losses are explained where possible. Crucially, an analysis of results should be undertaken on an intention-to-treat basis. This means that all participants who are randomised at the outset of an RCT should be included in the final analysis of outcomes, and kept to their originally assigned group, whether or not they adhered to the study protocol (Jüni, Altman and Egger, 2001). Failure to undertake an intention-to-treat analysis may bias the results often in favour of the intervention. An intention to treat analysis is also important because it maintains control of confounding by maintaining randomisation.

With reference to this NHC criterion, the current review also regarded evidence from RCTs as providing the best evidence to support the efficacy of screening tests. Only RCTs were considered for inclusion in the review, when the direct evidence available concerning screening test efficacy was evaluated (Chapters 3, 5 and 6). Other study designs were considered for inclusion in Chapter 4 alone, when evidence regarding diagnostic accuracy of FOBTs was reviewed.

**Criterion 5: The potential benefit from the screening programme should outweigh the potential physical and psychological harm caused by the test, diagnostic procedures and treatment**

There are various ethical issues to be considered concerning the establishment of population screening programmes. Since the primary intention of screening is to reduce the risk of disease or its consequences, the potential benefits from screening should outweigh the potential physical and/or psychological harms. The most important feature that distinguishes screening from other health services is that the health service approaches apparently healthy individuals and encourages them to participate, rather than the usual sequence when those who perceive they need help approach the health system and seek it. Therefore, in the case of screening, the health system has a particular obligation to minimise harm, as any adverse outcomes that arise from screening are, essentially, iatrogenic (Grimes and Schulz, 2002).

Potential harms can result from the screening test itself, the subsequent diagnostic procedures or any required treatment for the disease in question (NHC, 2003). This issue is clearly a concern, as it has been suggested by authors such as Alquist (1997, in Robinson, 1999) that deaths prevented by screening (benefits) may be cancelled out by deaths actually caused by screening (harms). The best and least-biased evidence regarding physical harms resulting from screening comes from high-quality RCTs that prospectively examine outcomes such as morbidity and mortality in the long-term.

Physical harms associated with screening

When CRC screening is considered, although actual risks are low, the highest risk of immediate physical harm comes from more invasive tests such as colonoscopy (and to a lesser degree FS) if they are used as the primary screening tool. However, even for FOBT, a very low-risk test, there remains a risk of physical harm to those who test positive at the primary screening episode, as they will consequently undergo a more invasive diagnostic test, such as colonoscopy, flexible sigmoidoscopy or double contrast barium enema (DCBE). It is possible that studies reporting complication rates may be over-estimates due to early experience with a procedure rather than current clinical practice. Conversely, rates may be under-estimated given that retrospective reviews may suffer from under-reporting, variable complication rates across general practice and major health centres, and late or longer term complications may not be included (Winawer, et al., 1997). Possible harms from CRC treatment for disease-positive screenees include post-operative sequelae as well as possible side-effects from adjuvant or palliative therapies, if they are used. Thus, those invited to participate in screening should have access to all the information that they require to make an informed decision to participate. A more detailed discussion concerning informed consent for screening is provided by the National Health Committee (2003).
Colonoscopy is performed as a day procedure and usually requires sedation. Complications include perforation and haemorrhage, particularly in association with polypectomy, and in rarer instances respiratory depression from sedation, arrhythmia, transient abdominal pain, nosocomial infection and death (Winawer, et al., 1997). Data from a review of six prospective studies on colonoscopic complications reported perforation rates of 1 per 1000 patients, haemorrhage rates of 3 per 1000 patients and mortality rates of 1 to 3 per 10,000 patients (Winawer, et al., 1997). A review of eight studies on complications associated with diagnostic colonoscopy found rates of perforation of 1.7 per 1000 colonoscopies, haemorrhage 3 per 10,000 colonoscopies and mortality 2 per 10,000 colonoscopies (Habr-Gama and Waye, 1989). The same review with nine studies on complication rates associated with colonoscopy, including polypectomy, found higher complication rates with a perforation rate of 14 per 1000 colonoscopies, haemorrhage 3 per 1000 colonoscopies and mortality 3 per 10,000 colonoscopies (Habr-Gama and Waye, 1989). The risk of transmitting infection by colonoscopy is low with no reports in larger, more recent studies (Winawer, et al., 1997).

The Minnesota (US) RCT on FOBT screening for CRC reported safety data associated with a population screening program utilising guaiac Hemoccult testing and dietary restrictions. Colonoscopic complications reported were perforation, with a rate of 3 per 10,000 colonoscopies, and serious bleeding, with a rate of 9 per 10,000 colonoscopies and no deaths associated with colonoscopy. Surgery was required in all four perforation instances and in three of the 11 serious bleeding cases (Mandel, et al, 1993). The Nottingham (UK) RCT on FOBT screening for CRC utilising guaiac Hemoccult testing and dietary restrictions also reported complication rates, but on a smaller sample of colonoscopies. The complication rate for colonoscopy was 5 per 1000 colonoscopies. Of the seven patients experiencing complications associated with colonoscopy, six required surgery. There were no complications associated with DCBE and no mortality associated with either procedure (Robinson, et al., 1999). The UK Flexible Sigmoidoscopy Screening Trial from baseline data reported four perforations among 2377 people (0.2%) having a colonoscopy (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002).

Physical harms associated with sigmoidoscopy are reported to be extremely low, with colon perforation rates of less than 2 in 10,000 examinations but slightly higher when biopsy or polypectomy is performed. The risk of transmitting infection by sigmoidoscopy is low with no reports in larger, more recent studies. Additionally, no mortality associated with the procedure has been reported (Winawer, et al., 1997). In the Norwegian Colorectal Cancer Prevention (NORCCAP) trial there were no serious complications at FS and minor events were reported in 0.2 percent at baseline (Gondal, et al., 2003). There were also no serious complications resulting from the 2,235 sigmoidoscopies performed after a positive test in the Funen-2 trial (Rasmussen, et al., 1999). There was only one perforation in more than 40,000 flexible sigmoidoscopies (including over 19,000 polypectomies) and 12 people were admitted to hospital for bleeding after FS in the UK Flexible Sigmoidoscopy Screening Trial. Eighty percent reported on the day of their FS (98% response rate) that they experienced no or mild pain, while 3 percent experiencing severe pain (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002).

The double contrast barium enema procedure (DCBE) is performed as an outpatient procedure and sedation is not used. Colon perforation and cardiac complications, though rare, are the most serious complications associated with DCBE. Overall complications are reported to be extremely low with an estimated 3 per 10,000 tests and a death rate of 3 in 100,000 tests. There is also no evidence that radiation exposure during an examination causes clinically important increases in the risk of tissue damage or cancer (Winawer, et al., 1997).

Psychological harms associated with screening

Although not specific to CRC screening, psychological harms can affect apparently asymptomatic and healthy individuals as a result of their involvement in population screening programmes. Psychological harms can be more difficult to measure objectively, or count, than physical ones, as they vary between individuals and situations. Such harms include:

- the generation of fear and apprehension concerning a particular disease
- anxiety while completing the screening test itself (e.g., distaste involved in handling faecal material required for FOBT, or embarrassment during procedure for FS)
• stress occurring while waiting for test results for both primary screening tests and then further confirmatory diagnostic tests if required
• the consequences of either false alarm or false reassurance resulting from false-positive or false-negative screening test results respectively.

The psychiatric morbidity associated with FOB screening was investigated as part of the Nottingham FOBT Trial (Parker et al., 2002). FOB screening did not cause increased or sustained anxiety or psychiatric morbidity, and the existence of psychiatric morbidity did not affect screening compliance. Anxiety scores were highest in those participants receiving notification of a positive test but these fell markedly in participants with false-positive test results after colonoscopy and at one month follow-up. There was no association between depression and suicide and screening. In a retrospective analysis of participants receiving false-positive test results, 26 percent were ‘very’ or ‘quite’ distressed and the other 74 percent were either ‘slightly’ or ‘not at all’ distressed (Mant, et al., 1990). In participants with false-positive results 85 percent thought they would accept rescreening in two years if offered. A similar high proportion of participants (98.1%) of false-positive test recipients thought the experience to be worthwhile (Mant, et al., 1990). Despite false-positive harms most participants were still in favour of screening. In the UK Flexible Sigmoidoscopy Screening Trial only 5 percent of participants reported at baseline finding the test more than mildly embarrassing, 98 percent were glad that they had the test and 97 percent said they would encourage a friend to have a test if asked (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002).

In summary, identifying and trying to quantify or anticipate potential harms from screening tests being considered for colorectal cancer is as important as calculating expected benefits.

Criterion 6: The health care system will be capable of supporting all necessary elements of the screening pathway, including diagnosis, follow-up and programme evaluation

For a colorectal cancer screening programme in New Zealand to replicate a reduction in mortality demonstrated in RCTs, an organised screening programme would be required. In screening programmes, as opposed to opportunistic screening, all activities are planned, co-ordinated, monitored and evaluated. Opportunistic approaches rely on an individual’s contact with the health system, as well as the initiative of the health professional or patient themselves. However, an organised screening programme systematically identifies and approaches all those individuals who are eligible to undergo screening, and co-ordinates all aspects of the screening pathway (see Figure 1) from start to finish, including ongoing monitoring and evaluation processes (National Health Committee, 2003; BreastScreen Aotearoa, 2004). An organised screening programme must ensure that all population groups who are likely to benefit from screening are reached. In New Zealand there is a particular requirement that specific strategies to include Māori in screening programmes are developed, using consultative and culturally appropriate processes (National Health Committee, 2003).

Planning a screening programme must include some consideration of screening test intervals. As outlined under the various tests considered for CRC, different tests may be used at different intervals. Intuitively, using a test frequently (with short between-test intervals) has appeal, given a possible benefit from increasing the chance that any existing cancers will be detected at an early stage. However, the trade-offs in this situation include that the number of false-positive screening tests is likely to increase (thereby increasing the potential for physical and psychological harms from further diagnostic testing), and that there may also be a corresponding decrease in cost-effectiveness of the screening programme. Direct evidence from RCTs has shown that both annual screening (with rehydrated tests) and biennial FOBT screening (with both rehydrated and non-rehydrated tests) can reduce CRC mortality. It is less clear from evidence to date what the optimal interval between tests for other potential screening modalities should be.

Evaluation of a screening programme must include some measures of screening programme performance. As screening programmes for CRC may involve multiple tests undertaken over time for each participant, calculating performance parameters for a screening programme is slightly more complex than calculating performance of a once-only test. Interval cancers (cancers that are detected within a given period after a negative screen) must be counted to allow for estimates of programme
sensitivity. Programme sensitivity is traditionally calculated as the ratio of screen detected/(screen detected + interval cancers). However, screening may detect cancers that may never have become symptomatic (overdiagnosis bias) as well as those that are slow-growing (length bias). These factors can be taken into account by another method of assessing programme performance, the proportional incidence method. The proportional incidence method compares the incidence of interval cancers with an estimate of what the predicted cancer incidence would be in the absence of screening, to give the proportion of expected cancers in which diagnostic anticipation was obtained by screening (Day, 1985). Both these methods assume that all interval cancers were missed by screening (rather than arising after screening) although various mathematical models can account for such assumptions.

To calculate actual programme specificity would require some method of diagnostic verification, which would require further testing in screen-negative participants. This would be neither feasible nor desirable in the context of an established screening programme. However, specificity can be estimated reasonably accurately by dividing all negative tests by the sum of all negative and false positive tests. This gives a good estimate since true negatives comprise such a large proportion of all negative tests. A further discussion of other methods to assess screening programme performance is beyond the scope of this chapter. It is important to note, however, that estimating the potential performance of a screening programme’s performance is another factor that underlies the prerequisite for high-quality evidence from well-conducted studies of screening efficacy, before nationwide screening programmes are established. This is because after establishment of such programmes there is no unscreened group with which comparisons can be made of what population outcomes would be in the absence of screening.
Figure 1. The colorectal screening pathway

1. Health promotion
2. Identification, invitation of, and participation by eligible people

Screening test
- e.g. FOBT
- Two-yearly

Test result
- Positive
- Negative

Further assessment
- e.g. colonoscopy

Result
- Positive
- Negative

Treatment and follow-up
- e.g. surgical removal lesion

Ongoing monitoring and evaluation of screening programme

---

2 This theoretical screening programme assumes two-yearly FOBT screening, with colonoscopy follow-up for those with positive FOBT.
**Criterion 7: There is consideration of social and ethical issues**

The NHC (2003) states “that there should be evidence that a complete screening programme (identification and invitation, test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically understood and acceptable to health professionals and the wider public”. Information regarding aspects of screening programme management that have made screening more or less acceptable to other populations could help inform the development and implementation of a colorectal cancer screening programme, were New Zealand to adopt this as a public health strategy. For example, analysing factors that may have led to screening programme success or failure in other countries, such as methods of approach to potential screenees or publicity campaigns promoting screening, may provide useful background knowledge. Additionally, measurable variables such as participation rates of population members approached can give a quantitative indication of the acceptability when published reports of screening RCTs are evaluated.

However, even if it is decided that adequate high-quality evidence for a screening test exists, prior to actual establishment of CRC screening in New Zealand specific social and ethical issues would require further research, particularly concerning the acceptability of CRC screening. This is exemplified by FOBT (the least invasive screening test), which may be distasteful to many members of any population, but for New Zealand Māori may present significant cultural difficulties with regards to handling and storage of faecal material (Working Party on Screening for Colorectal Cancer, 1998). The Working Party on Screening for Colorectal Cancer (1998) pointed out that acceptability of CRC screening has not been extensively studied in New Zealand, and that information on acceptability would best be determined by pilot programmes in this country. Piloting the screening process would be necessary for any test considered for CRC screening.

**Criterion 8: There is consideration of cost-benefit issues**

Screening programmes tend to be resource intensive, therefore it is important to consider the economic impact on the health system of a screening programme. Funding a CRC screening programme has an opportunity cost, in that other potential health programmes may not be funded. Economic information allows the efficiency of screening programmes to be compared to that of other health interventions. It is important that the costs and consequences of other strategies for which there is evidence of efficacy in reducing the burden of disease are considered as alternatives. For colorectal cancer screening, such a comparison requires not only reviewing the available information on the costs and consequences of different options for a screening programme, but also reviewing the available information on the costs and consequences of the current management of colorectal cancer patients. Useful contextual information can also be derived from cost data on other screening programmes, which may have similar components.

**REVIEW SCOPE**

The review scope was developed with the assistance of Ministry of Health staff and experts from the fields of screening, epidemiology and gastroenterology. The initial work undertaken for this review revealed that there was no population-based RCT evidence that any CRC screening modality other than FOBT had demonstrated a reduction in mortality. The Working Party on Screening for Colorectal Cancer (1998) had previously stated that screening methods that were not supported by evidence from randomised controlled trials would not be recommended.

The aim of this systematic review was to identify and appraise the most relevant new evidence published regarding screening for CRC that had emerged since the Working Party on Screening for Colorectal Cancer (1998) report. When this systematic review commenced, more data regarding screening with FOBTs had become available since 1998, and preliminary effectiveness data from RCTs evaluating screening with FS alone and combined FS and FOBT was available. At the time that this systematic review was proposed, it was anticipated that mortality data from the UK flexible sigmoidoscopy RCT may become available within the timeframe available to conduct the review, and that the final review may include an appraisal of such evidence. No new RCT efficacy data concerning the clinical outcomes of interest to this review for CRC screening using DCBE or colonoscopy had become available since the 1998 report, and virtual colonoscopy and other newer modalities remained
in early stages of research. Therefore, to provide the NSU with a focused systematic review of relevant new data, this review examined the efficacy of screening with the modalities of FOBT (including evidence for both guaiac and immunochemical types), FS, and the combination of FOBT and FS together.

The main focus of the economic assessment was to identify a likely range for the cost (per patient cost and also total annual cost to the health system) of a screening programme for colorectal cancer based on the reported results of the reviewed studies of the screening modalities above. A full review of the economic literature concerning colorectal cancer was conducted. All economic evaluations of colorectal cancer screening and treatment were identified and assessed. However, for completeness and consistency, only evaluations that involved both costs and consequences for colorectal cancer screening using FOBTs, FS, or FOBT and FS combined were reviewed.

Because this review was intended to update the information previously considered in New Zealand, evidence produced from 1997 onwards was considered for inclusion. The NHC (2003) criteria also require high-quality evidence for health outcomes from screening, therefore only high-quality, rigorous study design types reflecting the best evidence available were included. Given the recent New Zealand report for the surveillance and management of groups at increased risk of CRC, evidence that considered the screening and/or surveillance for these groups was not considered in this review. Full details of inclusion and exclusion criteria are provided in Chapter 2.

**AIM**

To systematically identify and appraise the international evidence for the effectiveness and cost-effectiveness of screening tests for colorectal cancer.

**REVIEW QUESTIONS**

The PICO (Patients, Interventions, Comparisons, Outcomes) criteria (Richardson et al., 1995) were used to develop well-defined clinical questions for this review (see Table 4). This involved focusing the questions on the following four elements:

- the patients or problem to be addressed
- the intervention or test being considered
- the comparison test, if necessary
- the clinical outcomes of interest.

Therefore, this review examined the following questions:

**Question 1** - What is the effectiveness of FOBT screening on outcomes from colorectal cancer, compared to no screening, in people aged 50-75 years?

**Question 2** - What is the accuracy of immunochemical FOBT screening, compared to guaiac FOBT screening, in people aged 50-75 years?

**Question 3** - What is the effectiveness of FS screening on outcomes from colorectal cancer, compared to no screening, in people aged 50-75 years?

**Question 4** - What is the effectiveness of FS and FOBT combined screening on outcomes from colorectal cancer, compared to FOBT screening alone, FS screening alone, or no screening, in people aged 50-75 years?
Question 5 - What is the cost-effectiveness or incremental cost-effectiveness of:

- FOBT screening compared to no screening
- immunochemical FOBT screening compared to guaiac FOBT screening
- FS screening compared to no screening
- FS and FOBT combined screening compared to FOBT screening, FS screening, or no screening, in people aged 50-75 years.

Table 4. PICO criteria for review questions

<table>
<thead>
<tr>
<th>Patient</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>People aged 50-75 years who are asymptomatic for CRC. [This population would not include those already identified as being at greater risk for CRC, as defined by the NZGG (2004).]</td>
<td>Guaiac or immunochemical FOBT</td>
<td>Usual care</td>
<td>Diagnostic performance</td>
</tr>
<tr>
<td></td>
<td>Immunochemical FOBT</td>
<td>Guaiac FOBT</td>
<td>Clinical outcomes</td>
</tr>
<tr>
<td></td>
<td>FS</td>
<td>Usual care</td>
<td>Outcomes relevant to screening</td>
</tr>
<tr>
<td></td>
<td>FS and FOBT combined</td>
<td>Usual care</td>
<td>FOBT alone</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FS alone</td>
</tr>
</tbody>
</table>

STRUCTURE OF REPORT

This report includes eight chapters, divided into sections. In Chapter 2 the methods (including selection criteria, search strategy, and study selection and appraisal methods) and limitations of the review are detailed. Chapter 3 investigates the effectiveness of FOBTs for CRC screening, and Chapter 4 the accuracy of guaiac and immunochemical FOBTs. The effectiveness of FS screening alone, and combined screening using FOBT and FS, is investigated in Chapters 5 and 6 respectively. Each chapter includes a summary of relevant findings of secondary research (systematic reviews and meta-analyses) where available, as well as primary research considered. Detailed evidence tables are provided, presenting key features by papers appraised. Each chapter summarises results, briefly discusses methodological limitations in the area, and presents key conclusions. Chapter 7 presents the systematic review of analyses of cost-effectiveness appraised. In Chapter 8, an overall summary and review conclusions are provided.
Chapter 2: Methodology

2.1 STUDY SELECTION

All studies examining the effectiveness and cost-effectiveness of CRC screening using the tests outlined in the scope of the review were identified. The following selection criteria were applied to the articles identified by the literature search.

Study inclusion criteria

Publication type and date

Publications included primary research (published as full original reports) and secondary research (systematic reviews and meta-analyses). Papers were included for review if they were published between January 1997 and November 2004 inclusive. Earlier papers were accessed where required to provide background material.

Study design

For primary research studies relevant to the effectiveness of screening tests for CRC only randomised controlled trials were included, with the exception of studies examining the accuracy of immunochemical FOBT screening compared with guaiac FOBT screening. Here, where consideration was also made of study designs of a specific level of evidence and above [III-2 and above], to reflect the “best evidence” available on the issue.

The evidence presented in the selected studies was initially considered using the dimensions of evidence defined by the National Health and Medical Research Council (2000). These dimensions (see Table 5) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from appraisal of the literature in the form of a systematic review. The last two domains require expert clinical input as part of their determination. Three sub-domains collectively denote the strength of evidence: level, quality and statistical precision.

Table 5. Evidence dimensions

<table>
<thead>
<tr>
<th>Type of evidence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength of the evidence</td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>The study design used, as an indicator of the degree to which bias has been eliminated by design (see Table 6)</td>
</tr>
<tr>
<td>Quality</td>
<td>The methods used by investigators to minimise bias within a study design</td>
</tr>
<tr>
<td>Statistical precision</td>
<td>The p-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect</td>
</tr>
<tr>
<td>Size of effect</td>
<td>The distance of the study estimate from the “null” value and the inclusion of only clinically important effects in the confidence interval</td>
</tr>
<tr>
<td>Relevance of evidence</td>
<td>The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used</td>
</tr>
</tbody>
</table>

The levels describe groups of research that are broadly associated with particular methodological limitations. However, these levels are only a general guide to quality because each study may be designed and/or conducted with particular strengths and weaknesses. High-level evidence is provided by a well-conducted randomised controlled trial.
Table 6. Designations of levels of evidence*

<table>
<thead>
<tr>
<th>Grade of evidence</th>
<th>Study design</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Evidence obtained from a systematic review of all relevant randomised controlled trials</td>
</tr>
<tr>
<td>II</td>
<td>Evidence obtained from at least one properly-designed randomised controlled trial</td>
</tr>
<tr>
<td>III-1</td>
<td>Evidence obtained from well-designed pseudorandomised controlled trials (alternate allocation or some other method)</td>
</tr>
<tr>
<td>III-2</td>
<td>Evidence obtained from comparative studies [including systematic reviews of such studies] with concurrent controls and allocation not randomised, cohort studies, case-control studies, or interrupted time series with a control group</td>
</tr>
<tr>
<td>III-3</td>
<td>Evidence obtained from comparative studies with historical control, two or more single arm studies, or interrupted time series without a parallel control group</td>
</tr>
<tr>
<td>IV</td>
<td>Evidence obtained from case series, either post-test or pre-test/post-test</td>
</tr>
</tbody>
</table>

* Modified from National Health and Medical Research Council (1999)

For economic studies relevant to the cost-effectiveness of screening tests for CRC, study types considered for inclusion were cost-minimisation analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis (Drummond et al., 1997). The evaluations were classified according to the type of evaluation based on the schema outlined in Table 7. This approach identifies the type of evaluation according to whether it considers both costs and consequences (outcomes) and whether it considers one option or more than one option. Only evaluations deemed to fit into category 4 (full economic evaluation) were included. All foreign currency results have been converted to New Zealand dollars using a current exchange rate to provide a rough estimate of the New Zealand dollar equivalent.

Table 7. Classification of economic evaluations of health care

<table>
<thead>
<tr>
<th>Both costs and consequences examined?</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only one option considered</td>
<td>Partial evaluation</td>
</tr>
<tr>
<td></td>
<td>1A Outcome description</td>
</tr>
<tr>
<td></td>
<td>1B Cost description</td>
</tr>
<tr>
<td>Two or more options considered</td>
<td>Partial evaluation</td>
</tr>
<tr>
<td></td>
<td>3A Efficacy evaluation</td>
</tr>
<tr>
<td></td>
<td>3B Cost analysis</td>
</tr>
<tr>
<td></td>
<td>Full economic evaluation</td>
</tr>
<tr>
<td></td>
<td>4A Cost-minimisation analysis</td>
</tr>
<tr>
<td></td>
<td>4B Cost-effectiveness analysis</td>
</tr>
<tr>
<td></td>
<td>4C Cost-utility analysis</td>
</tr>
<tr>
<td></td>
<td>4D Cost-benefit analysis</td>
</tr>
</tbody>
</table>

Secondary research studies reporting systematic reviews or meta-analyses of RCTs were included if they included a methods section describing how the relevant studies were identified. A methods section should include the search terms used, databases searched, and the dates searching was conducted. However, it is important to note that these papers may not have employed the same inclusion and exclusion criteria as have been used here, or do not consider subsequently published research. Therefore, the results must be interpreted with caution.

Sample size

Studies with samples of at least 15 participants.

Population

People aged 50-75 years who are asymptomatic for CRC.

---

3 This population did not include those already identified as being at greater risk for CRC, as defined by the NZGG (2004)
Index tests and comparators

Studies reporting on:

- FOBT screening compared with no screening
- immunochemical FOBT screening compared with guaiac FOBT screening
- FS screening compared with no screening
- FS and FOBT combined screening compared with FOBT screening, FS screening, or no screening.

Outcomes

Measures of the following key quantitative outcomes relating to diagnostic test performance, health outcomes and outcomes relating to screening presented in the results:

**Diagnostic Test Performance**

- sensitivity (Se)
- specificity (Sp)
- positive predictive value (PPV)
- negative predictive value (NPV).

The methodology used for calculating test sensitivity, specificity, PPV and NPV is presented in Appendix 1.

**Health Outcomes**

- stage of CRC at detection
- CRC morbidity or mortality
- incidence of CRC
- overall mortality.

**Outcomes relating to screening**

- benefits - could be measured by improved health outcomes (or any psychological sequelae) for those allocated to screen group compared to control, as well as if screening is more cost-effective compared to other health interventions for same condition;
- harms (including physical and psychological sequelae) - could be measured by worsened health outcomes (or any psychological sequelae) for those allocated to screen group compared to control, as well as if screening is less cost-effective compared to other health interventions for same condition. Harms include potential risks e.g., risk of being offered a colonoscopy (a more invasive diagnostic test) for those screened by FOBT, (which would be increased by higher test positivity rates) and measured sequelae e.g., adverse effects following FS or colonoscopy tests used as part of screening programme;
- acceptability - this outcome could be measured quantitatively by participation or compliance rates for those allocated to screening (i.e. the proportion of those allocated or invited to screening who actually undertook screening).
Study exclusion criteria

Publication type
The following publication types were excluded:
- non-systematic reviews
- correspondence
- editorials
- expert opinion articles
- comments
- articles published in abstract form
- conference proceedings
- studies that did not clearly describe their methods/results.

Language
Non-English language articles were excluded as none were deemed necessary to the review.

Animal studies
Animal studies were excluded.

2.2 SEARCH STRATEGY

A systematic method of literature searching and selection was employed in the preparation of this review.

An approach using a series of searches over all bibliographic databases was employed. Firstly a core search strategy on colorectal cancer screening was developed for each of the Medline and Embase databases. This was then used as the basis for five separate searches:
- a search for clinical trials using validated filters for trials and systematic reviews/meta-analyses
- a cost-effectiveness search
- a search for immunochemical tests (no date limit)
- a search for guaiac tests and any other references to faecal occult blood testing which might not have been picked up in the other searches or had been published since the first searches were carried out
- a search for flexible sigmoidoscopy used for colorectal screening

searches were not limited by language. All searches except that for the immunochemical tests were limited to the years 1997 and onwards. The first four searches were completed on 29 October 2004. The search on flexible sigmoidoscopy was completed on 17 January 2005.

Principal sources of information
The following databases were searched, using the search strategy outlined in Appendix 2:

Bibliographic databases
- Cinahl
- Current Contents
- Embase
Medline
- Science Citation Index
- Social Science Citation Index
- Cochrane Central Register of Controlled Trials

Review databases
- ACP Journal Club
- Cochrane Database of Systematic Reviews
- Database of Abstracts of Reviews of Effectiveness
- Health Technology Assessment database
- NHS Economic Evaluation database

Main Websites
- Australian National Cancer Initiative
- Clinical Trials.gov
- Current Controlled Trials Meta-Register
- US National Guidelines Clearing House

Other
- One author was contacted for clarification of details in a published paper. Three authors were contacted regarding publication dates for planned analyses.
- Hand searching of journals, contacting of manufacturers, or contacting of authors for unpublished research was not undertaken in this review.
- A complete list of the sources searched for this review is given in Appendix 3.

Search terms used
- index terms from Medline (MeSH terms): exp colorectal neoplasms, mass screening, occult blood, guaiac, immunochemistry, exp immunologic tests, costs and cost analysis, sigmoidoscopy;
- index terms from Embase: exp large intestine cancer, exp rectum cancer, exp colorectal cancer, cancer screening, exp economic evaluation, exp health care cost;
- the above index terms were used as keywords in databases where they were not available and in those databases without controlled vocabulary;
- additional keywords (not standard index terms) were used in all databases: (colorectal or colon$ or bowel or sigmoid or caecal or caecum or rectal or rectum) adj (cancer or neoplas$ or carcino$ or adenocarcino$ or polyp$ or malignant$ or adenoma), screen$, occult blood, (fecal or feces) adj2 blood), fobt, haemoccult, hemoccult, coloscreen$, hemo fec, guaiac, guiac, guajac, fecatest, fecatwin, shionogi-B, immunochemi$, hemeselect, imdia-hem, imdiahem, imdia adj hem, bayerdetect, bayer detect, hemsp, immudia-hem, iatro hemchk, la hemochaser, oc hemodia, flex sure obt, quicktest, dima fob-10, magstream hemsp, insure, bm-test colon albumin, 'inform, monohem, hema fecia, haemafecia, flexible adj5 sigmoid$, cost containment, cost benefit$, cost minim$, cost utili$, cost effectiv$, economic evaluation$, economic analy$, health economi$.

2.3 STUDY SELECTION

Studies were selected for appraisal using a two-stage process. Initially, the titles and abstracts (where available) identified from the search strategy, were scanned and excluded as appropriate. The full text
articles were retrieved for the remaining studies and were appraised if they fulfilled the study selection criteria outlined above.

More than 1986 studies were identified by the search strategy. Ineligible studies were excluded in two phases. First, after applying selection criteria to the search titles and abstracts, articles potentially eligible for review or as useful for background information were retrieved. Excluding the latter, 220 full text articles were considered as potentially eligible for inclusion. Of these, 164 did not fulfil the inclusion criteria and are presented in Appendix 4. These articles were excluded for the following reasons:

- did not meet inclusion criteria (n = 96)
- non-systematic review (n = 30)
- correspondence (n = 7)
- editorial (n = 3)
- expert opinion article (n = 6)
- comments (n = 6)
- abstract only available (n = 11)
- conference proceedings (n = 1)
- unclear methods/results (n = 4)

A total of 64 appraised studies are included in this report. These represent 56 articles, as some review articles were appraised more than once because different results were relevant to different research questions. They are presented in Appendix 5. In addition, publications cited (including ongoing trials and those considered as background material) are presented in the References.

2.4 APPRAISAL OF STUDIES

Data extraction into evidence tables, analysis and presentation

Primary evidence

For primary studies of the effectiveness of screening tests for CRC, data relevant to study quality and statistical precision was extracted using design-relevant checklists (GATE: a Graphic Appraisal Tool for Evidence-based clinical practice. Available from: http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/ Current as of May 2005). Formal checklists use qualitative criteria designed to assess whether the research study was performed in a valid, reliable and rigorous way to minimise bias. Summaries of appraisal results are shown in tabular form as Evidence Tables and include:

- reference (authors, publication date) and country where study was principally conducted
- study design and evidence level (applying NHMRC criteria)
- description of intervention and comparator, and dates of testing
- sample characteristics including number of participants for intervention and comparator groups
- participant inclusion and exclusion criteria
- analyses comparing intervention and comparator groups at baseline
- eligible outcome measures used and outcome verification
- results of analyses comparing intervention and comparator groups on eligible outcomes, including statistically-tested comparisons and reporting relevant statistical data
- adverse effects (and incidence) for the intervention and comparator groups
• comments on the study’s limitation relevant to its internal validity
• authors’ conclusions and reviewers’ conclusions.

In addition to evidence tables, the appraised studies were briefly summarised in the report’s text with reference to study design and specific methodological flaws. Inter-study consistency was considered, results synthesised and overall conclusions drawn based on the study design and the specific problems associated with individual studies.

Secondary evidence
Systematic reviews and meta-analyses were described and critiqued. Data relevant to the methodology and quality of the review was extracted using a design-relevant checklist (GATE: a Graphic Appraisal Tool for Evidence-based clinical practice, available from: http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/ Current as of May 2005). The checklist included assessments of what search strategy was used, whether the review asked a focused question, if the eligibility criteria for included trials were explicit, how the validity of included trials was assessed and whether results of included trials were similar. Summaries of appraisal results are shown in evidence tables, and include:

• reference (authors, publication date) and country where review was principally conducted
• search method employed by the review
• review question, where reported
• review inclusion and exclusion criteria
• results of analyses comparing intervention and comparator groups on eligible outcomes, including statistically-tested comparisons and reporting relevant statistical data
• comments on the review’s limitations relevant to its internal validity
• authors’ conclusions.

In addition, the appraised reviews were briefly summarised in the report’s text, and authors’ conclusions reported.

Economic data
The economic evaluations eligible for review were described and appraised in terms of their design, methods, data sources, key results, sensitivity of the model to value changes in variables, limitations and conclusions. The quality of the economic data included in the review was classified based on Drummond et al.’s (1997) recommendations:

• collection of resource utilisation data
• completeness of costs (including not only direct testing costs but also indirect costs such as treatment costs and the costs of promoting the programme and maintaining participation amongst targeted groups)
• costs presented as per unit of resource used
• an incremental ratio appropriately presented
• a sensitivity analysis performed.

There are four common types of formal economic evaluation as defined by Drummond et al. (1997). Cost minimisation is the simplest evaluation technique, where the new technology is demonstrated to be no worse than existing testing regimens at either the same or lesser cost. Cost-effectiveness analysis is used where the proposed technology is demonstrated to offer an improved outcome, although possibly at an increased cost. The measure of cost-effectiveness is the incremental cost per additional
unit of outcome achieved. Cost-utility analysis requires an additional analysis of benefit to patients through a utility measure (e.g., such as QALYs). Quality adjusted life years (QALYs) are based on both increased numbers of years of life, as well as an assessment of quality through measures such as changes in activities of daily living. Cost-benefit analysis attempts to place monetary or dollar values on all benefits as well as costs. Lives saved are often given monetary value estimates utilising comparisons with other economic sector expenditures, such as averted road deaths or prevention of accidental drowning.

The assessment of the economic literature took into consideration that cost estimates are dependent on both the accuracy of various probabilities (in testing and treatment success rates and patient treatment choice) and also on assumptions, which may vary from one evaluation to another. Treatment effectiveness and best practice regimens vary over time and by localised specialist preference. Every effort was made to use estimates from recent international studies and the results of the respective sensitivity analyses were also included in the literature review. Where possible, the review also extracted information on benefits and indirect costs to patients, such as time costs, recovery costs, and informal care costs.

**LIMITATIONS OF THE REVIEW**

This study has used a structured approach to review the literature. However, there were some inherent limitations with this approach. Namely, systematic reviews are limited by the quality of the studies included in the review and the review’s methodology.

This review has been limited by the inclusion of only English language studies. Restriction by language may result in study bias, but the direction of this bias cannot be determined. However, it is important to note that no important non-English studies were identified by the search strategy. In addition, the review has been limited to the published academic literature, and has not appraised unpublished work. Restriction to the published literature is likely to lead to bias since the unpublished literature tends to consist of studies not identifying a significant result.

Papers published pre-1997 were not considered for appraisal, as these had already been considered by the Working Party on Screening for Colorectal Cancer (1998).

The studies were initially selected by examining the abstracts of these articles. Therefore, it is possible that some studies were inappropriately excluded prior to examination of the full text article. However, where detail was lacking or ambiguous, papers were retrieved as full text to minimise this possibility.

All studies included in this review were conducted outside New Zealand, therefore, their generalisability to the New Zealand population and context may be limited and needs to be considered.

This review examined the effectiveness and cost-effectiveness of the interventions and did not consider ethical or legal considerations associated with these interventions. Interventions were not assessed in terms of their impact on general quality of life.

Although three researchers appraised the articles included in this review they did not cross-validate the data extraction and appraisal process.

This review was conducted over a limited timeframe (September 2004 to March 2005).

This review has greatly benefited from the advice provided by external reviewers. However, it has not been exposed to wider peer review.

For a detailed description of interventions and evaluation methods, and results used in the studies appraised, the reader is referred to the original papers cited.
Chapter 3: Effectiveness of FOBTs

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY

The search identified 11 eligible primary research papers. Below is an overview of study designs and aspects of quality represented by these studies. Full details of the papers appraised, including methods, key results, limitations and conclusions, are provided in evidence Table 8. Studies are presented in chronological order of publication within the table.

Study designs and quality assessments

The search identified four eligible RCTs comparing FOBT screening with no screening. The RCTs were all graded as level II evidence. Ten of the included papers reported further data on outcomes from three of these trials, which compared screening with the guaiac test Haemoccult/Haemoccult II with no screening: the Minnesota RCT (Church et al., 1997; Mandel et al., 1999; Mandel et al., 2000); the Nottingham RCT (Mapp et al., 1999; Moss et al., 1999; Robinson et al., 1999; Parker et al., 2002; Scholfield et al., 2002) and the Funen-1 RCT (Jorgensen et al., 2002; Kronborg et al., 2004). These trials were all of high quality in their design and conduct, with power calculations reported, random allocation to study groups, and adequate blinding of those interpreting key outcome measures. As previously noted by the Working Party on Screening for Colorectal Cancer (1998), which considered previously published data from these RCTs, pertinent analyses in these trials were conducted as intention-to-treat (stated for the Funen-1 trial and assumed for the Nottingham and Minnesota trials). The major differences between these trials are well described in the working party’s report, and will only be briefly outlined here.

The other included paper reported from a randomised controlled trial conducted in Jiashan, China that compared screening with an immunochemical test plus a health questionnaire to no screening (Zheng et al., 2003). This study was of moderate quality in design and conduct, with adequate blinding of those interpreting key outcome measures. However, there was unclear reporting of some aspects of study analysis, most importantly whether analyses were conducted on an intention-to-treat basis, and whether specific analytic techniques to account for the cluster randomisation design were used.

The Minnesota trial

Study setting and sample

Three papers published since 1997 reported on outcomes from the Minnesota trial. This RCT enrolled a sample of more than 45,000 people aged 50-80 years from a volunteer population, and compared both annual and biennial Hemoccult testing with no screening. The investigators recruited the study population by contacting and inviting individuals who were members of civic and fraternal organisations. Potential participants were provided with a brief description of the study, and enrolled by completing a consent form and another form and returning both to the study investigators.

An FOBT in the Minnesota trial required two samples from each of three consecutive stools. After the first three years of this trial all FOBTs were rehydrated prior to development (overall 82.5% of the slides were rehydrated). Screenees were requested to restrict their diet and medications prior to completing FOB test cards. A positive FOBT was defined as being any one from six slides as positive, and FOBT-positive screenees were invited to attend for a diagnostic work-up that included colonoscopy if possible.

This trial stopped actively inviting participants for both annual (after 11 rounds) and biennial (after six rounds) screening in 1992, and reported initial follow-up after an average of 13 years in 1993. Subsequently, the investigators have continued follow-up of the three study groups by postal questionnaires mailed annually, and verified reported lesions and deaths by reviewing medical records and death certificates. None of the papers appraised reported whether or not participants in the
interception group continued with FOBT following the cessation of active invitation, nor whether there
was any uptake of FOBT in the control group.

The participation rate for screening in this trial was 89 percent for biennial and 90 percent for annual
(2000) reported that the average compliances of those invited for screening was 78 percent for biennial
and 75 percent for annual screening.

Outcomes

Test positivity rate

During the two screening intervals for the Minnesota RCT (1976-1982 and 1986-1992) about one
million FOBTs were processed. The positivity rate was 2.4 percent for non-rehydrated tests, increasing
to 9.8 percent with FOBT rehydration. There was a trend towards an increase in the test positivity rate
with increased age at entry into the trial, which was more marked with rehydrated slides and seen in
both men and women. The positivity rates for first compared to subsequent screening rounds were not
reported.

Diagnostic test performance

Regarding diagnostic test performance, in this RCT the sensitivity of the non-rehydrated test was 81
percent and specificity was 98 percent; when the FOBT was rehydrated sensitivity was 92 percent and
specificity was 90 percent (Working Party on Screening for Colorectal Cancer, 1998). Church et al.
(1997) presented a re-analysis of the data for annual screening first reported by Mandel et al. (1993).
They calculated an average sensitivity of around 90 percent for both the screen test and the screening
programme. This information does not differ greatly from that provided in the initial report, where test
sensitivity was calculated at 92 percent (rehydrated).

Stage CRC detected

After the initial 13 years of trial follow-up Mandel et al. (1993) had reported that the cumulative
incidence of CRC, according to study group and Duke’s cancer stage, showed changes in the incidence
of stage A and D cancers that were consistent with earlier detection of CRC by screening.

Subsequently, Mandel et al. (1999) have reported that over the period of 18 years of follow-up a total
of 1292 colorectal adenocarcinomas were diagnosed and validated. A reduced incidence of Duke’s
stage D CRC in both screening groups (47% reduction for annual screening and 32% reduction for
biennial screening) compared to the control groups was still seen at 18-year follow-up (Mandel et al.,
1999). However, it was noted by this review that the change in the incidence of Duke’s stage A CRC
according to study group that was reported by Mandel et al. (1993) was less marked by 1999, seven
years after the end of the second screening interval. The cumulative 18-year incidence* of stage A
cancer in the biennial screening group (approximately 8.4 per 1000) was almost the same as that of the
control group (approximately 8.3 per 1000), and both were only slightly less than the annual screening
group (approximately 9.0 per 1000).

CRC morbidity/mortality

After 18 years of follow-up, Mandel et al. (1999) reported a 21 percent reduction in CRC mortality for
those screened biennially compared to the control group [RR = 0.79 (95%CI 0.62-0.97)]. This
significant risk reduction was better than the data [RR 0.94 (95%CI 0.68-1.31)] reported after 13 years
of follow-up (Mandel et al., 1993). After 18 years of follow-up the reduction in CRC mortality for
annual screening remained essentially the same, at 33 percent [RR = 0.67 (95%CI 0.51-0.83)], as in the
earlier report [RR = 0.67 (95%CI 0.50-0.87)]. However, there was some narrowing of the associated
95%CI around the more recently reported risk reduction.

* These incidence rates are approximations taken from a figure in Mandel et al.’s (1999) paper; the exact rates are not reported
separately in the text.
Incidence of CRC

After 13 years of follow-up Mandel et al. (1993) had reported the incidence of CRC as 12 percent lower in the screened groups than the control groups, but this difference did not reach statistical significance. However, subsequent follow-up of study participants has revealed statistically significant reductions in the incidence of CRC for both the biennial and annual screening groups compared to the control group, with cumulative incidence ratios of 0.83 (95%CI 0.73-0.94) and 0.80 (95%CI 0.70-0.90) respectively (Mandel et al., 2000). Mandel et al. (2000) explained this reduction in incidence as being most likely due to the processes of the screening programme allowing for the identification and removal of the pre-cursor lesions for CRC.

Overall mortality

Consistent with Mandel et al’s (1993) findings, after a further five years of follow-up Mandel et al. (1999) found no significant difference for overall mortality between either the biennial or annual screening group and the control group (Mandel et al., 1999).

Outcomes related to screening

There was no new data reported regarding outcomes relating to screening in the papers appraised.

The Nottingham trial

Study setting and sample

Five papers were appraised reporting outcomes from the Nottingham trial. This RCT enrolled a sample of more than 150,000 people aged 50-74 years (45–74 years in a pilot study) from the general population. It compared biennial Haemoccult testing with no screening and the test cards were not rehydrated. The investigators recruited the study population by identifying men and women aged 50-74 years who lived in the Nottingham area of the United Kingdom according to general practice registration lists. Doctors at each practice were asked to remove any person whom they judged to be ineligible from the list before the investigators randomised the sample. Biennial invitations to participate and FOB testing kits were sent to the screen group only. Initially only those who participated in the first round were re-invited, but from September 1990 previous non-responders were also invited. The control group were not told about the trial and continued to use usual health care facilities.

An FOBT in this RCT required two samples from each of three consecutive stools. Screenees were not requested to restrict their diet and medications prior to completing FOB test cards. A positive FOBT after this process was defined as five or six slides positive. If up to four of the initial six slides were positive, the FOBT was repeated with diet restrictions, and defined as a positive FOBT if one or more slides were positive at retesting. FOBT positive screenees were invited to attend for a colonoscopic examination if possible.

This trial stopped actively inviting participants for biennial screening (after at least three rounds per screenee) in February 1995. Active follow-up of both study groups was continued thereafter, through local hospital records and a central health system register. None of the papers appraised reported whether or not participants in the intervention group continued with FOBT following the cessation of active invitation, nor whether there was any uptake of FOBT in the control group.

The participation rate for screening in the first round of this trial was 53 percent with 38 percent completing all screening rounds. After re-inviting previous non-responders, the overall completion rate for at least one screen increased to 59 percent (Working Party on Screening for Colorectal Cancer, 1998).
Outcomes

Test positivity rate

The test positivity rate was 2.1 percent for first screening test, and 1.2 percent following rescreen within 27 months (Working Party on Screening for Colorectal Cancer, 1998). Four percent of those accepting screening at least once underwent one or more examinations of their rectum and colon (Robinson et al., 1999). Scholefield et al. (2002) reported a cumulative risk (for participants in the intervention group) of having a positive FOBT as 2.6 percent (1977 of 76224). Seventy-three percent of these participants underwent a colonoscopy while the rest had other investigations such as barium enemas, which gave a cumulative colonoscopy rate in the intervention group of 1.9 percent.

Diagnostic test performance

From previous reports, the sensitivity of the screening test as used in this RCT was 53.6 percent, and specificity was estimated at between 96 and 98 percent (Working Party on Screening for Colorectal Cancer, 1998).

Moss et al. (1999) reviewed the sensitivity of the screening test and screening programme sensitivity, using less traditional methodologies to calculate these parameters. They found a crude average test sensitivity for CRC of 54.1 percent using the traditional method, and 54 percent with the proportional incidence method, with test sensitivity being higher in first screening round (Se = 69.7% and 62.7% respectively). Screening programme sensitivity for CRC was 59 percent, and was significantly higher for those aged over 64 years (Se = 64.2%) than those aged 45-64 years (Se = 53.8%). These authors did not report corresponding specificity estimates.

Stage CRC detected

In Hardcastle et al.’s (1996) report from the Nottingham RCT, which reported on data collected up to June 1995, it was noted that there was a significant difference in the proportion of Duke’s stage A CRC lesions detected between the study groups (screen group = 20%, control group = 11%; p<0.001). Furthermore, there was also a significant difference in the proportion of advanced (Duke’s stages C and D) cancers detected (screen group = 46%, control group = 52%; p <0.01). Mapp et al.’s (1999) paper, which focused on survival and included data collected up until December 1996, also presented data on CRC stage at detection. This data was similar to Hardcastle et al.’s (1996) previous report, as there remained a significant difference in the proportion of Duke’s stage A CRC lesions detected between the study groups (screen group = 19%, control group = 12%; p <0.001). Likewise, there was also a significant difference in the proportion of advanced (Duke’s stages C and D) cancers detected (screen group = 45%, control group = 50%; p = 0.02).

CRC morbidity/mortality

The health outcomes from the Nottingham trial of CRC mortality and incidence were the primary outcomes of interest reported in Scholefield et al.’s (2002) paper. After a median of 11.7 years of follow-up there was a 13 percent reduction in CRC mortality for those screened compared to the control group [RR = 0.87 (95%CI 0.78-0.97)]. This significant risk reduction was comparable with the data [RR 0.85 (95%CI 0.74-0.98)] that Hardcastle et al. (1996) had previously reported after 7.8 years of follow-up.

Incidence of CRC

Again consistent with the earlier findings from the Nottingham RCT, Scholefield et al. (2002) reported there was no significant difference found for overall CRC incidence between the screening and control groups (Scholefield at al., 2002).

Overall mortality

Also consistent with previous findings, there was no significant difference found for major causes of death (other than CRC) between the screening and control groups in this trial (Scholefield at al., 2002).
Outcomes relating to screening

Mapp et al. (1999) reported that a benefit for those in the screening group compared to those in the control group for people diagnosed with CRC was an improvement in disease prognosis. This was present for cancers detected at both initial screening and subsequent screening rounds. This benefit derived from screening persisted and was of statistical significance when evaluated according to time since entry into the trial. The authors suggested that by measuring survival from time of entry into the trial, rather than from time of disease detection, the benefit resulting from screening was not simply due to lead-time bias (see Table 8, reporting on Mapp et al., 1999). It is important to note, however, that this outcome is still vulnerable to over-diagnosis and length bias.

Harms resulting as a consequence of FOBT screening in the Nottingham trial were the specific outcome of interest in Robinson et al.’s (1999) paper (which reviewed physical harms). Concerning these harms, there were six deaths in those with early-stage cancer that could have been attributed to ‘over-diagnosis’. The authors concluded that for every 10 CRC deaths prevented by screening in this RCT, one person would be sufficiently harmed by the processes of investigation to require surgical intervention (Robinson et al., 1999). Scholefield et al. (2002) also provided some discussion on physical harms, noting that the most serious complications associated with colonoscopy are cardiovascular. They reported that overall mortality for ischaemic heart disease (IHD) was similar in the intervention and control groups, and also that there was no significant difference between IHD mortality rates between those undergoing colonoscopy following a positive FOBT compared to the control group.

The paper by Parker et al. (2002) evaluated psychological harms, and the limited data derived from the main RCT regarding indicators of psychological harm allowed these authors to conclude that there was no additional risk of suicide in the screen group compared to the control group.

The Funen-1 Trial

Study setting and sample

The Funen-1 RCT enrolled a sample of over 60,000 people aged 45-75 years from the general population. It compared biennial Hemoccult testing with no screening and the test cards were not rehydrated. The investigators recruited the study population by identifying men and women aged 45-74 years who lived in Funen, Denmark. They used hospital and county records and a civil registration system to identify ineligible persons prior to randomisation of the sample. Biennial invitations to participate were sent to the screen group only. Two reminder letters were sent in the initial screening round, and one reminder per round thereafter. Only those who participated in the first round were re-invited for further screening. The control group were not told about the trial and continued to use usual health care facilities.

An FOBT required two samples from each of three consecutive stools. Screenees were requested to restrict their diet and medications prior to completing FOB test cards. A positive FOBT was defined as being any one from six slides as positive, and FOBT positive screenees were invited to attend for a medical review that included colonoscopy if possible.

This trial continued actively re-inviting the intervention group for biennial screening and following up both study groups. Follow-up of both study groups was continued through the Funen patient database, the county public health officer and the Danish National Registry of patients. None of the papers appraised reported whether or not there was any uptake of FOBT in the control group.

The participation rate for screening in this trial was 67 percent in the first round, for rounds thereafter participation of those invited remained high (91-97%) (Kronborg et al., 2004).
Outcomes

Test positivity rate

The FOBT positivity rate in the first round of this RCT was 1 percent. For rescreening rounds the percentages of those testing positive ranged from 0.8 percent to 3.8 percent (round 2 = 0.8%, round 3 = 0.9, round 4 = 1.2%, round 5 = 1.8%, round 6 = 3.8%, round 7 = 1.7%, round 8 = 1.1%, round 9 = 1.4%). Kronborg et al. (2004) explained that the higher positivity rate in round six was in part due to difficulties with reading of the FOBTs and that the problems were rectified in later testing rounds. The cumulative risk of having a positive FOBT requiring more invasive diagnostic testing was 5.7 percent (1766 of 30,762) and at least one colonoscopy was performed in 93.2 percent of these persons. Therefore, from the whole screening group, a cumulative proportion of 5.3 percent underwent colonoscopy one or more times (Kronborg et al., 2004).

Diagnostic test performance

From previous reports, the sensitivity of the screening test as used in this trial was 51 percent, and specificity was estimated at 98 percent (Working Party on Screening for Colorectal Cancer, 1998). There were no further data reported regarding diagnostic test performance reported in the appraised papers.

Stage CRC detected

In Kronborg et al.’s (1996) report from the Funen-1 RCT that considered data collected up until August 1995 it was noted that the proportion of Duke’s stage A CRC lesions detected in the control group (11%) was significantly lower than that in the screen group (22%, p<0.01). Considering data collected up until August 2002, Kronborg et al. (2004) also noted a difference in the proportion of Duke’s stage A disease in the control group (11 %) and the screen group (18 %). This difference was also of statistical significance (p< 0.001).

CRC morbidity/mortality

The health outcomes of CRC mortality and incidence were the primary outcomes of interest in the papers appraised. After 14 years and seven screening rounds, Jorgensen et al., (2002) reported a 14 percent reduction in CRC mortality for those screened compared to the control group [RR = 0.86 (95%CI 0.73-1.00)]. This was lower than the initial 18 percent [RR 0.82 (95%CI 0.68-0.99)] reduction seen after 10 years of screening (Kronborg et al., 1996). Furthermore, after 17 years and nine screening rounds, the reduction in CRC mortality was reduced even further, to 11 percent [RR = 0.89 (95%CI 0.78-1.01)], as recently reported by Kronborg et al., (2004). The authors explained this lessening of risk reduction as likely being due to the decreased numbers of subjects being screened as the screening rounds increased.

Incidence of CRC

In both the subsequent reports from this trial there was no significant difference reported for the incidence of CRC (Jorgensen et al., 2002; Kronborg et al., 2004). This was consistent with the findings after 10 years of screening.

Overall mortality

There was no significant difference reported for overall mortality between the screening and control groups (Jorgensen et al., 2002; Kronborg et al., 2004). This was also consistent with previous findings.

Outcomes relating to screening

Of relevance to the potentially beneficial protective effect of participating in screening, Jorgensen et al. (2002) reported that those who were invited but refused any screening had a significantly increased risk [RR 1.65 (95% CI 1.30-2.08)] of death from CRC compared to those who accepted all screening after
seven rounds. This effect could perhaps be partially explained by different health related behaviours, such as delayed help-seeking between screen-refusers and accepters. However, as this comparison is between those who accepted and rejected screening, it is vulnerable to confounding and selection bias, and cannot therefore be taken as an accurate measure of the protection provided by screening. No information on possible reasons for refusing participation in this trial was reported.

Regarding harms from screening, Kronborg et al. (2004) mentioned the finding that in the screening group there was no mortality resulting from colonoscopy itself. None of the papers appraised had harms from CRC screening as a specific outcome of interest.

The Jiashan trial

Study setting and sample

The Jiashan RCT enrolled a sample of more than 190,000 people aged over 30 years from the general population. Although the median age range of 40-49 years of the sample enrolled in this trial was less than that of interest to this review, this paper was the only other RCT of FOBT screening versus no screening found by the search strategy. It has therefore been included for comparative purposes, given that it compared a programme of once-only immunochemical FOB testing (using a reverse passive haemagglutination test) plus a health questionnaire with no screening.

The investigators used a comprehensive housing registry system to recruit the study population, identifying all the residents of Jiashan County aged 30 years or older. It was not clear how study participants were invited into the trial, although it is stated that all were fully informed prior to the commencement of screening.

An FOBT required eight samples from one stool, with two samples being selected by the laboratory for development. Screenees were requested to restrict specific medications prior to completing FOB test cards. A positive screen was defined as having either a positive FOBT, or a questionnaire score above a pre-determined threshold. Screen-positive subjects were invited to attend for further investigation with a 60cm flexible sigmoidoscope.

From those randomised to screening, just over 80 percent of those invited were recruited into the screening programme. However, the participation rate (from those randomised to screening) for completion of both tests required for this screening programme was less, at 66.4 percent. The proportion of those recruited who completed only the FOBT was not reported.

Outcomes

Test positivity rate

From those screened in Jiashan, 1.2 percent had a positive FOBT plus a ‘positive’ questionnaire result. Three percent had the FOBT as positive, with a ‘negative’ questionnaire. For FOBT alone, the overall positivity rate in this RCT was therefore 4.2 percent. Of the 62,677 screenees who completed both the tests offered by the screening programme 4299 (6.9%) were considered to be ‘screen positive’ and were asked to undergo further testing. Seventy four percent of these screen-positive participants actually went on to undergo flexible sigmoidoscopy.

Diagnostic test performance

There was no data reported from this trial regarding diagnostic test performance.

Stage CRC detected

Zheng et al. (2003) stated that 21 cases of CRC were detected in screen-positive participants who underwent flexible sigmoidoscopy, and that 71.5 percent of these cancers were early stage (10 Duke’s grade A and five Duke’s B). However, no corresponding data were reported regarding CRC stage at detection in the control group and therefore no comparison can be made between the screen and control groups for this outcome.
CRC morbidity/mortality and incidence

The health outcomes of CRC cumulative mortality and incidence were the primary outcomes of interest. The data provided in the report allowed for some calculations to be made of the relative risks for these outcomes. The only statistically significant result from these calculations was a 32 percent [RR = 0.68 (95%CI 0.54-0.87)] reduction in mortality from rectal cancer in screenees compared to controls. There was no significant reduction for either colonic cancer or overall colorectal cancer. As detailed in Table 8, this likely reflected the fact that most screen-positive participants underwent further evaluation of only the distal portion of their bowel, within the reach of the flexible sigmoidoscope. However, it should be noted that the precision of these relative risk estimates may be reduced by the cluster design of the study and unclear reporting of whether the cluster design was taken into account in investigators’ analysis of trial data presented.

It was difficult to evaluate what influence the young age of the study population in this RCT may have had on the efficacy of screening to reduce CRC mortality. No sub-group data that could have enabled a calculation of the effect of screening on CRC mortality in those of different ages was presented in the paper. However, it can be postulated that given the overall young age of the study population (and therefore the probable reduced incidence of CRC in such a young group) the impact of screening would likely have been reduced in this trial. The investigators justified their choice of start age by explaining that CRC occurs at an approximately 10 years younger age in China than in Westerners.

Overall mortality

No data for overall mortality was reported in this paper.

Outcomes relating to screening

There were no data specific to screening outcomes reported in this paper.
Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Sample size</th>
<th>Sample characteristics</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Church et al. (1997)</td>
<td>RCT</td>
<td>Grade II</td>
<td></td>
<td>Intervention</td>
<td>Recruited potential participants aged 50-80 years from Minnesota by informing members of civic and fraternal organisations about the study and asking if they would participate. Excluded those reporting a history of colorectal cancer, familial polyposis, or chronic ulcerative colitis at time of enrollment. Bedridden or otherwise disabled people.</td>
<td>46,561 eligible participants were enrolled. Enrolled subjects were randomly assigned (stratified by age, sex and residence)</td>
<td>Undertaken a re-analysis of 13-year data from Minnesota RCT of annual FOBT screening to refine estimates of: *Screen sensitivity By modifying a model developed by Lang and Ransahoff (which assigns those receiving colonoscopy a reduced CRC mortality rate for the following 5 years, while the remaining subjects experience the same rate as the US population) to include screen sensitivity of the FOBT. By using a truncated product-bimodal model, which assumes that for a cancer to be missed, for each of the 3 stools, either blood is present and both slides fail to react or blood is absent. The probability of detecting the cancer (screen sensitivity) is 1 minus the probability of missing it. Screening programme sensitivity Positive predictivity (PPV) for CRC and polyps according to the number of slides per test that were positive for cancer and polyps. For all colorectal lesions reported in questionnaire, medical records concerning diagnosis were obtained. Cause of death determined by obtaining death certificates for all participants who died and subjecting the certificate to independent committee review, blind to participant screening status. Pathological staging of CRC determined by study pathologist.</td>
<td>*Screen sensitivity Direct crude estimates average = 89% Estimates based on modified Lang-Ransahoff model range = 94-96% Estimates based on truncated product-bimodal model range = 68-99% (median 96%) Screening programme sensitivity Crude estimate = 89% ± 2.2 Screening test PPV For CRC: Crude overall PPV = 2.8% With 6 slides per test, PPV increased as the number of positive slides increased, (Spearman’s ρ = 0.94, p = 0.02). For any kind of polyps (including non-malignant neoplasms), increase in PPV with number of positive slides was non-significant (Spearman’s ρ = 0.77, p = 0.1). *Screen sensitivity was defined as the probability that at least one of the six slides for a case of detectable, pre-clinical CRC will be positive at the time of a single screen. Understood as a function not only of the slide sensitivity, but also of the degree of dependence between the individual slide results.</td>
<td>Authors’ conclusions The methods used herein demonstrate that programme sensitivity of the test that used rehydrated FOB slides annually is in the range of 90% for a programme that is 5 years or more in duration. The answer is important because the diagnostic work-ups associated with the high positivity rate of faecal occult blood test increase costs considerably. The results for screening test PPV are consistent with the FOBT being sensitive for CRC but relatively insensitive for non-malignant neoplasms. Reviewers’ conclusions The study (Minnesota RCT) from which the data for this analysis were taken was well-conducted and reported. Sample size and power calculations had been performed at outset. Allocation of participants to study groups was random and resulted in comparable groups at baseline. Due to the nature of the intervention and controls under study, it was not possible to blind participants to their allocated group, however the trial included blinded assessment of primary outcomes. At least 90% of those allocated to annual screening participated at least once. Follow-up to 13 years was complete for over 99% of study participants.</td>
<td></td>
</tr>
</tbody>
</table>
### Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Mandel et al. (1999) | RCT | Grade II | Intervention  
Guaiac FOBT (Hemoccult®, SmithKline Diagnostics, Palo Alto), Rehydrated (except for initial 3 years of trial), Requested to restrict diet and medications before and during sample collection.  
Annual (11 rounds total) or biennial testing (6 rounds total). Those with positive FOBT (one or more from six slides positive) were invited for diagnostic work-up that included colonoscopy (or DCBE if colonoscopy refused or incomplete), or saw own physician. All consultation results obtained. Detected abnormalities, particularly polyps or CRC, treated and followed up.  
Control  
Usual care.  
All participants sent annual questionnaire to identify any colorectal lesions detected and subject deaths. Repeated efforts made to contact non-responders from both study groups, including test-refusers from screening groups. | Recruited potential participants aged 50-80 years from Minnesota by informing members of civic and fraternal organisations about the study and asking if they would participate.  
Excluded Those reporting a history of colorectal cancer, familial polyposis, or chronic ulcerative colitis at time of enrolment.  
Bedridden or otherwise disabled people.  
46,551 eligible participants were enrolled. Enrolled subjects were randomly assigned (stratified by age, sex and residence) to:  
anual testing group (n=15,570)  
biennial testing group (n=15,587)  
control group – no invitation (n=15,394).  
Randomisation effectively created 3 balanced study groups with respect to age (median age range = 60-69 years) sex (male: female = 0.92) and place of residence.  
18-year follow-up of study participants reported, at which point vital status was known for 88.8% of participants. | After 18-years follow-up  
CRC mortality  
All-cause mortality  
Compliance with screening  
For all colorectal lesions reported in questionnaire, medical records concerning diagnosis were obtained. Cause of death determined by obtaining death certificates for all participants who died and subjecting certificate to independent committee review, blind to participant screening status. Pathological staging of CRC determined by study pathologist.  
Death certificates obtained for 99.8% of known decedants. | CRC mortality  
Control  
Rate* = 14.09 (95% CI 12.01-16.17)  
Annual screening  
Rate* = 9.46 (95% CI 7.75-11.17)  
Biennial screening  
Rate* = 11.19 (95% CI 9.39-12.99) | Authors' conclusions  
Both annual and biennial testing for faecal occult blood are effective methods for statistically significantly reducing colorectal cancer mortality, with the benefit from annual screening appearing to be greater than for biennial screening.  
Reviewers' conclusions  
The study (Minnesota RCT) from which the data for this follow-up report has been derived was well conducted and reported. Sample size and power calculations had been performed at outset. Allocation of participants to study groups was random and resulted in comparable groups at baseline. Due to the nature of the intervention and controls under study, it was not possible to blind participants to their allocated group, but the trial did include blinded assessment of primary outcomes. At least 99% of those allocated to both the annual and biennial screening groups participated at least once. Follow-up for vital status to 18 years was complete for more than 91% of participants in all three groups. After 18 years of follow-up, the CRC mortality rate in the biennial screening group showed a statistically significant reduction of 21% compared to the control group, which was not present at 13-year follow-up. The statistically significant risk reduction of 33% for annual screening remained the same as that found at 13-year follow-up. |
### Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapp et al. (1999)</td>
<td>RCT</td>
<td>Grade II</td>
<td>- Intervention: Invitation to screening programme with guaiac FOBT (Haemoccult®, Rohm Pharma, Wittenstadt). Unhydrated, biennial testing.</td>
<td>Recruited participants from general population of Nottingham aged 45-74 years, identified through general practice registers. Excluded: Those judged by family doctor to have serious illness that would exclude them from participating, including those with a history of colorectal cancer.</td>
<td>Survival of individuals following CRC diagnosis. According to RCT study group and method of detection, study groups were classified as control group or screening group. Cancers in the screening group were further classified by method as screen-detected (first screen and subsequent screen), interval cancers, non-responders and adenoma follow-up.</td>
<td>Survival following CRC diagnosis (Hazard ratio).</td>
<td>Authors’ conclusions: The results of this analysis suggest that, for CRC detected at first screen, about 35% of the observed improvement in prognosis relative to that in cancers diagnosed in the control group is explained by differences in tumour stage and differentiation, and for cancers detected at later screens about 40% of the improvement is explained. The significantly better survival in the study group overall, as measured from date of entry, suggests that this benefit is not simply due to a lead-time effect.</td>
</tr>
<tr>
<td>Nottingham, UK.</td>
<td></td>
<td></td>
<td>- FOBT, then repeated in 3 months. Those with 1 or more slides positive from 2nd FOBT, or 5-6 slides positive from 1st FOBT were offered colonoscopic investigation (or DCBE if colonoscopy refused or incomplete).</td>
<td>Detected adenomas at CRC screening were treated and transferred to endoscopic follow-up programme.</td>
<td>Initially only those who participated in screening (and without CRC or adenomas) were re-invited. From September 1990, previous non-responders were also invited every 2 years.</td>
<td>Survival according to time since entry into the trial.</td>
<td>Multivariate analysis (stage and tumour differentiation included in the model) Control = 1.0 Screen detected First screen = 0.53 (95% CI 0.34-0.81) Subsequent screen = 0.62 (95% CI 0.43-0.90) Interval = 0.98 (95% CI 0.72-1.07) Non-responder = 1.14 (95% CI 0.89-1.12) Adenoma follow-up = 0.14 (95% CI 0.02-0.03)</td>
</tr>
</tbody>
</table>
Table 8.  Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moss et al. (1999)</td>
<td>RCT Grade II</td>
<td>Intervention</td>
<td>Recruitment participants from general population of Nottingham aged 45-74 years; identified through general practice registers.</td>
<td>Test Sensitivity</td>
<td>Test sensitivity</td>
<td>Authors’ conclusions. The findings in this trial suggest that biennial screening using unhydrated Haemoccult will only detect around half of the progressive cancers in an unselected population. The differences between test sensitivity measured by the traditional and proportional methods are of relevance to understanding the natural history of CRC. The difference between estimates of sensitivity using the two methods is greatest for those aged ≥ 65 years at entry to the trial, and in cancers of the distal colon, suggesting there may be a higher proportion of slower growing tumours in these groups.</td>
</tr>
<tr>
<td>Nottingham, UK.</td>
<td></td>
<td>Invitation to screening programme with guaiac FOBT (Haemoccult®, Rohm Pharma, Wellenstaid). Unhydrated. Biennial testing.</td>
<td>Excluded Those judged by family doctor to have serious illness that would exclude them from participating, including those with a history of colorectal cancer.</td>
<td></td>
<td>Proportional incidence method (for follow-up examination):</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average test Se = 54%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Male = 44.3%, Female = 63.0% ((X^2 = 6.4, p = 0.01))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>First screen = 62.7%, Rescreen = 50% ((X^2 = 2.58, p = 0.1))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Age at entry to trial 45-64 years = 54.8%, Age at entry to trial 65 years = 52.6% ((X^2 = 0.05, p = 0.6))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rectum and rectosigmoid = 59.3%, Sigma and descending colon = 55.8%, Transverse and ascending colon = 48.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traditional method:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening round 1 = 69.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening round 2 = 57.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening round 3 = 47.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening round 4 = 48.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening round 5 = 47.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crude average test Se = 54.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average weighted by number of screen-detected cases = 60.1% per round</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Screening programme sensitivity</td>
<td>Using ‘traditional method’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Interval cancers were classified as any cancers diagnosed within 2 years of a negative screen, or after a positive screen where further investigation was negative or refused.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Information on the development of CRC in study groups was obtained from pathology registers of local hospitals, a regional cancer registry, family doctors reports, a national health registry and census survey data.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Test sensitivity</td>
<td>Proportional incidence method (for follow-up examination):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average test Se = 60.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Male = 58.9%, Female = 59.1% ((X^2 = 4.48, p = 0.03))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rectum and rectosigmoid = 58.2%, Sigma and descending colon = 69.1%, Transverse and ascending colon = 49.1%</td>
<td></td>
</tr>
<tr>
<td>Source Country</td>
<td>Study Design Evidence Grading</td>
<td>Comparison interventions and dates of testing</td>
<td>Sample</td>
<td>Outcomes and verification</td>
<td>Results</td>
<td>Comments</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------</td>
<td>---------------------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Robinson et al. (1999)</td>
<td>RCT</td>
<td>Intervention Invitation to screening programme with guaiac FOBT (Haemoccult) Unrehydrated. Biennial testing. Those with 1–4 positive from 6 slides repeated FOBT using 6 stools, with 2-day diet restrictions. If negative 2nd FOBT, then repeated in 3 months. Those with 1 or more slides positive from 2nd FOBT, or 5–6 slides positive from 1st FOBT were offered colonoscopy investigation (or DCBE if colonoscopy refused or incomplete). Detected adenomas or CRC were treated and transferred to endoscopic follow-up programme. Initially only those who participated in screening (and without CRC or adenomas) were re-invited. From September 1990, previous non-responders were also invited every 2 years. Control identified but not invited to screening, usual care. Recruited participants between February 1981 and February 1991. Screening stopped in February 1995, by which time all participants had been offered FOBT at least 3 times. Deaths considered to June 1996.</td>
<td>Recruited participants from general population of Nottingham aged 45–74 years, identified through general practice registers. Excluded: Those judged by family doctor to have serious illness that would exclude them from participating, including those with a history of colorectal cancer. 152,850 eligible participants were recruited and randomly assigned (stratified by age, sex and household) to: screening group (n=76,466) vs control group (n=76,384). Randomisation created 2 well-matched groups at study entry, with ratio (median age range = 55–64 years) and sex (male: female = 0.92). 2599 (1.7%) of participants recruited into the study could not be traced or had emigrated at follow-up, and were excluded from analysis.</td>
<td>Risks from FOBT screening False negative results: • Assessed by comparing the outcomes for patients in the interval cancer group with the control group. (Assumes all cancers presenting in the interval between screens were false negatives.) Harm from colonic investigation: • Assessed by considering both direct injury from procedures and indirect harms such as cardiac effects from bowel preparation or sedation. Overdiagnosis of CRC: • Estimated by counting number of patients dying within 30 days of surgery for any screen-detected CRC or adenoma, or between 30 days to two years of surgery for screen-detected benign adenomas or Duke’s A CRC. May be considered overdiagnosed if certified cause of death not CRC. Cancers in the study group were classified as screen-detected, interval cancers and cancers in those never screened. Information on CRC deaths and harms from pathology registries, regional cancer registry, family doctors’ reports, national health registry and census survey data. All deaths where CRC noted on death certificate scrutinised to determine cause of death, by reviewers blind to study group.</td>
<td>N = 1778 people (4.0% of those accepting screening) underwent examination of the colon and rectum at least once (1474 colonoscopies, 748 DCBEs). Numbers of CRC diagnosed in trial Screen-detected n = 236 Endoscopic follow-up detection n = 8 Interval cancers n = 249 Non-screen participants n = 400 Control group n = 856 False negative results: Proportion of Stage A CRC tumours Interval = 16% (39/249) Control = 11% (85/786), p &lt; 0.05 Proportion of stage C.D and unknown CRC tumours Interval = 54% (134/249) Control = 56% (476/856), p &gt; 0.1 Improved survival for interval cancers compared to control, p &lt; 0.01 Improved survival for interval cancers compared to control remains significant if stage distribution taken into account, p &lt; 0.02 Harm from colonic investigation: Mortality DCBE n = 0 Colonoscopy n = 7 (5 perforations, 1 haemorrhage, 1 snare entrapment). No deaths following colonic investigation Overdiagnosis of CRC: Deaths within 30 days of surgery n = 5 Deaths 30 days−2 years of surgery n = 36 Pre-operative deaths (stage D CRC) n = 2 Total n = 43 From these 43 deaths, 6 had had adenomas or stage A cancer (and died of unrelated cause within 2 years of screen-detected CRC), and therefore could be classified as ‘overdiagnosed’.</td>
<td>Limitations When considering overdiagnosis by screening, it would have been helpful to have had the data from the control group post-operative mortality data too, to make a crude comparison with the usual ‘expected’ number of deaths following this type of surgery for symptom-detected CRC or adenomas. Authors’ conclusions One person will be sufficiently harmed by the process of investigation to require surgical intervention for 10 CRC deaths prevented by screening. Adverse consequences of screening are common to any screening programme, and must be recognised minimised by close attention to quality control and audit. Reviewers’ conclusions The study (Nottingham RCT) from which the data for this follow-up report has been derived was well conducted and reported. Sample size and power calculations were performed at outset. Allocation of participants to study groups was random and resulted in comparable groups at baseline. Due to the nature of the intervention and controls under study, it was not possible to blind participants to their allocated group, but the trial did include blinded assessment of primary outcomes. Follow-up was complete for over 98% of participants.</td>
</tr>
</tbody>
</table>
Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Study interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandel et al. (2000)</td>
<td>Minnesota, USA</td>
<td>RCT</td>
<td>Grade II</td>
<td>Recruited potential participants aged 50-80 years from Minnesota by informing members of civic and fraternal organisations about the study and asking if they would participate. Excluded: Those reporting a history of colorectal cancer, familial polyposis, or chronic ulcerative colitis at time of enrollment. Bedridden or otherwise disabled people. Of 46,551 eligible enrolled and randomised participants, 106 had actually had a diagnosis of CRC prior to randomization, and were omitted from the analysis of cumulative incidence. Randomization (stratified by age, sex and residence) had allocated remaining 44,445 subjects to: annual testing group (n=15,532) biennial testing group (n=15,550) control group – no invitation (n=15,363). Randomisation effectively created 3 balanced study groups with respect to age (median age range = 60-69 years), sex (male: female = 0.92) and place of residence. Follow-up for vital status more than 90%.</td>
<td>Recruit potential participants aged 50-80 years from Minnesota by informing members of civic and fraternal organizations about the study and asking if they would participate. Excluded: Those reporting a history of colorectal cancer, familial polyposis, or chronic ulcerative colitis at time of enrollment. Bedridden or otherwise disabled people. Of 46,551 eligible enrolled and randomised participants, 106 had actually had a diagnosis of CRC prior to randomization, and were omitted from the analysis of cumulative incidence. Randomization (stratified by age, sex and residence) had allocated remaining 44,445 subjects to: annual testing group (n=15,532) biennial testing group (n=15,550) control group – no invitation (n=15,363). Randomisation effectively created 3 balanced study groups with respect to age (median age range = 60-69 years), sex (male: female = 0.92) and place of residence. Follow-up for vital status more than 90%.</td>
<td>Cumulative incidence of CRC after 18 years follow-up. PPV of a positive FOBT test for: CRC Adenomatous polyps For all colorectal lesions reported in questionnaire, medical records concerning diagnosis were obtained. Cause of death determined by obtaining death certificates for all participants who died. Selected certificates were subjected to independent committee review, blind to participant screening status. Pathological staging of CRC determined by study pathologist. Active surveillance continued after screening cessation. Death certificates obtained for 99.9% of subjects known to have died during the 18-year follow-up period.</td>
<td>CRC cumulative incidence/1000: Control: 39 (95% CI 36-43) Annual screening: 32 (95% CI 29-35) Biennial screening: 33 (95% CI 30-36) Cumulative incidence ratio of screen group compared with control group: Annual screening 0.80 (95% CI 0.70-0.90), p&lt; 0.001 Biennial screening 0.83 (95% CI 0.73-0.94), p&lt; 0.002 PPV of positive FOBT for CRC* Annual screening Range = 0.87% to 4.53% (Spearmans rho= 0.94, p = 0.02) Biennial screening Range = 1.12% to 6.13% (Spearmans rho= 0.94, p = 0.02) PPV of positive FOBT for adenomatous polyps at least 1cm diameter Annual screening Range = 5.99% to 7.87% (Spearmans rho= 0.94, p = 0.02) Biennial screening Range = 6.86% to 10.08% (Spearmans rho= 0.83, p = 0.06)</td>
<td>Authors’ conclusions Our findings demonstrate a significant reduction in the incidence of CRC after faecal occult blood testing, which occurred after both annual and biennial screening. The most plausible explanation is the identification and removal of the precursor lesions for CRC. This explanation is supported by the finding that the positive predictive value increased with the number of positive slides in our study. Our study supports the theory of the adenoma-carcinoma sequence and emphasizes the importance of detecting and resecting advanced adenomas. Reviewers’ conclusions The study (Minnesota RCT) from which the data for this follow-up report has been derived was well conducted and reported. Sample size and power calculations had been performed at outset. Allocation of participants to study groups was random and resulted in comparable groups at baseline. Due to the nature of the intervention and controls under study, it was not possible to blind participants to their allocated group, but the trial did include blinded assessment of primary outcomes. At least 90% of those allocated both the annual and biennial screening groups participated at least once. Follow-up for vital status to 18 years was complete for over 91% of participants in all three groups. After 18 years of follow-up, cumulative incidence of CRC in both biennial screening groups has shown a statistically significant reduction compared to control group. This was not present at 13-year follow-up.</td>
</tr>
</tbody>
</table>
Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jørgensen et al. (2002)</td>
<td>RCT</td>
<td>Grade II</td>
<td>Intervention: Guaiac FOBT</td>
<td>Recruited participants from 140,000 general population of Funen, Denmark aged 45-75 years.</td>
<td>CRC mortality</td>
<td>CRC mortality alone: Ratio* = 0.82 (95% CI 0.69-0.97), p = 0.02</td>
<td>Authors’ conclusions</td>
</tr>
<tr>
<td>Funen, Denmark. (Funen-1 trial)</td>
<td></td>
<td></td>
<td>Non-rehydrated, diet-restricted. Biennial testing (7 rounds total). Those with positive FOBT (one or more from six slides positive) were invited for medical review and full colonoscopy (or DCBE if colonoscopy refused or incomplete). Those with CRC or adenomas detected were then invited to surveillance programme. Only screenees who participated in previous rounds (and without CRC or adenomas) were re-invited.</td>
<td>137,485 were eligible. Husbands/wives allocated to same group. Those eligible were randomised by computer (in 14 block groups) to: intervention group - received a mailed invitation (plus 2 reminders) to participate in FOBT screening (n=30,967), control group - no invitation (n=30,964).</td>
<td>CRC adenoma/cancer from Funen patient database, National Registry. Verified cause of death. Investigator review of death certificates was conducted to blind to participant screening status, with independent committee review if unclear whether CRC was cause of death.</td>
<td>CRC total mortality (including treatment complications): Crude ratio* = 0.85 (95% CI 0.73-1.00), p=0.05 Adjusted for age and sex: Ratio* = 0.86 (95% CI 0.73-1.00) [On-treatment analysis Ratio* = 0.70 (95% CI 0.58-0.85)]</td>
<td>Persistent reduction in CRC mortality observed for those in screening group, after 7 biennial screening rounds. However, increasing positivity rate and number of colonoscopies may increase economic cost and participant risk. Significant reduction in risk of death from proximal CRC compared to distal supports the consideration of a screening programme with addition of episodic FS to FOBT screening.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effect of compliance with screening on CRC mortality for those from screening group who accepted specific screening rounds relative to those who refused screening. For screening group: Participation of those invited for each screening round Verified diagnosis of CRC/adenoma from Funen patient database, public health officer and National Registry. Verified cause of death. Investigator review of death certificates was conducted to blind to participant screening status, with independent committee review if unclear whether CRC was cause of death.</td>
<td>Significant difference observed between anatomical sub-site ratios (p = 0.04).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>All-cause mortality Ratio* = 1.00 (95% CI 0.97-1.03), p = 0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Incidence of CRC Ratio* = 1.02 (95% CI 0.91-1.13), p = 0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effect of compliance with screening on risk of death from CRC. Ratio of screening group participants who refused all screening relative to those who accepted all screening rounds = 1.65 (95% CI 1.30-2.08). Participation of those invited for screening: 67% round 1, 92-94% rounds 2-7. *Ratio of screen group compared with control group</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER**
<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Grade</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nottingham, UK.</td>
<td>RCT</td>
<td>Grade II</td>
<td>Intervention: Invitation to screening programme with guaiac FOBT (Haemoccult). Unrehydrated. Biennial testing. Those with 1-4 positive from 6 slides repeated FOBT using 6 stools, with 2-day diet restrictions. If negative 2nd FOBT, then repeated in 3 months. Those with 1 or more slides positive from 2nd FOBT, or 5-6 slides positive from 1st FOBT were offered colonoscopic investigation (or DCBE if colonoscopy refused or incomplete). Detected adenomas or CRC were treated and transferred to endoscopic follow-up programme.</td>
<td>Recruited participants from general population of Nottingham aged 50-74 years. Identified through general practice registers. Excluded: Those judged by family doctor to have serious illness that would exclude them from participating, including those with a history of colorectal cancer.</td>
<td>Psychiatric morbidity: Number of subjects committing suicide. Information on CRC deaths and harms obtained from pathology registers of local hospitals, regional cancer registry, family doctors' reports, national health registry and census survey data. All deaths where CRC noted on death certificate scrutinised to determine cause of death, by reviewers blind to study group.</td>
<td>Number of certified cases of suicide in trial population: Screening group n = 53/75253 (0.62/1000) Control group n = 48/74998 (0.63/1000), p = 0.63. Of the 53 reported suicides in the screening group, 24 were screen-test refusers, 6 had accepted initially but refused subsequent tests, and 23 had accepted on each occasion offered. Only one subject had a positive test, with colonoscopy showing an adenoma.</td>
<td>Limitations: This paper mainly reports on the results of a before-and-after survey undertaken on a small group of screening group subjects who participated in the main RCT. Suicide is the only outcome that could be linked with psychiatric morbidity for which screening and control group data are reported. Authors' conclusions: This study shows that participation in a FOBT screening programme does not cause sustained psychiatric morbidity. Reviewers' conclusions: This study (Nottingham RCT) from which eligible data for outcomes of interest to this report has been derived was well conducted and reported. Sample size and power calculations were performed at outset. Allocation of participants to study groups was random and resulted in comparable groups at baseline. Due to the nature of the intervention and controls under study, it was not possible to blind participants to their allocated group, but the trial did include blinded assessment of primary outcomes. Follow-up was complete for over 98% of participants. There was no significant difference in number of suicides between study groups. However, suicide is likely to provide only a gross representation of psychiatric morbidity between the two study groups. Surveying those from both study groups at outset and during the RCT may have detected more subtle indicators of psychiatric morbidity, such as anxiety.</td>
</tr>
</tbody>
</table>
### Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study Design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nottingham, UK</td>
<td>RCT</td>
<td>Grade II</td>
<td>Intervention to screening programme with guaiac, FOBT (Haemoccult), Unhydrated, Biennal testing.</td>
<td>Recruited participants from general population of Nottingham aged 50-74 years. Identified through general practice registers.</td>
<td>After median 11.7 (range 8.4-18.4) years follow-up</td>
<td>Total person-years observation: Control group = 843.463, Screening group = 844.419</td>
<td>Authors’ conclusions: While the data from this study continue to show a reduction in disease-specific mortality from FOBT screening, despite a compliance rate of only 57%, it is reassuring that there was no significant excess of deaths following the screening process. Further follow-up data are awaited to determine the effect of the screening process on the incidence of this disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Those with 1-4 positive from 6 stools repeated FOBT using 6 stools, with 2-day diet restrictions. If negative 2nd, FOBT, then repeated in 3 months. Those with 1 or more slides positive from 2nd FOBT, or 5-6 slides positive from 1st FOBT were offered colonoscopic investigation (or DCBE if colonoscopy refused or incomplete).</td>
<td>Excluded: Those judged by family doctor to have serious illness that would exclude them from participating, excluding those with a history of colorectal cancer.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Detected adenomas or CRC were treated and transferred to endoscopic follow-up programme.</td>
<td>Initially only those who participated in screening (and without CRC or adenomas) were re-invited. From September 1990, previous non-responders were also invited every 2 years.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zheng et al. (2003)</td>
<td>Jiashan, China.</td>
<td>RCT Grade II</td>
<td>Intervention only screening. If FOBT negative, then questionnaire score &gt; 0.5 was required for diagnostic test (60cm FS). If FS negative, 2nd FOBT performed and colonoscopy. If 2nd FOBT positive, all lesions detected were biopsied and removed if possible at time of diagnostic test. Further treatment and follow-up not reported.</td>
<td>Recruited participants from all residents aged ≥ 30 years from 21 townships of Jiashan county, China.</td>
<td>Screening participation: CRC mortality (Total: Screen group = 208/100,000 (95% CI 196-218); Control group = 244/100,000 (95% CI 233-255); p not reported. RR = 0.85 (95% CI 0.71-1.03); Colon: Screen group = 90/100,000 (95% CI 83-97); Control group = 83/100,000 (95% CI 76-90); p = 0.22; RR = 1.08 (95% CI 0.80-1.46); Rectal: Screen group = 110/100,000 (95% CI 102-118); Control group = 161/100,000 (95% CI 152-170); p = 0.003. RR = 0.68 (95% CI 0.54-0.87).</td>
<td>CRC cumulative mortality rate (8-year): Total CRC: Screen group = 208/100,000 (95% CI 196-218); Control group = 244/100,000 (95% CI 233-255); p not reported. RR = 0.85 (95% CI 0.71-1.03); Colon CRC: Screen group = 90/100,000 (95% CI 83-97); Control group = 83/100,000 (95% CI 76-90); p = 0.22; RR = 1.08 (95% CI 0.80-1.46); Rectal CRC: Screen group = 110/100,000 (95% CI 102-118); Control group = 161/100,000 (95% CI 152-170); p = 0.003. RR = 0.68 (95% CI 0.54-0.87).</td>
<td>Limitations: Unclear whether exclusion of ineligible individuals occurred before or after randomisation. Power calculations not reported. Not reported whether analysis was ‘intention-to-treat’. Not reported whether cluster randomisation taken into account in analysis, thus potential for less precision of effect of screening on CRC outcomes. Because screening tool included questionnaire as well as FOBT, and no separate CRC mortality or incidence data given for those with positive FOBT results, cannot calculate effect of FOBT alone on CRC mortality. This would likely bias trial results in favour of FOBT efficacy because some of those with negative FOBT (but score &gt; 0.5 on questionnaire) also underwent diagnostic tests. Authors’ conclusions: The results from this randomised trial provide strong evidence that mortality from rectal cancer may be significantly reduced by RPHA-FOBT accompanied by an individual risk-assessment questionnaire.</td>
</tr>
<tr>
<td>Source</td>
<td>Country</td>
<td>Study design Evidence Grading</td>
<td>Comparison interventions and dates of testing</td>
<td>Sample</td>
<td>Outcomes and verification</td>
<td>Results</td>
<td>Comments</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>-----------------------------</td>
<td>---------------------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Zheng et al. (2003)</td>
<td>Jiashan, China. (Continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sampling unit was township. Townships matched by population size and age distribution into 10 pairs (smallest 2 townships = 1 unit). From each pair, townships were randomly assigned to: Screen group (n=94423) Control group (n=97838). Randomisation created 2 balanced study groups with respect to gender (male: female = 1.04), age (median age range = 40-49 years) and historical data on incidence, mortality and survival from CRC.</td>
<td></td>
<td></td>
<td>CRC cumulative incidence rate (8-year) Total CRC: Screen group = 395/100,000 (95% CI 381-410) Control group = 401/100,000 (95% CI 386-416) RR = 0.98 (95% CI 0.86-1.13) Colonic CRC: Screen group = 165/100,000 (95% CI 156-174) Control group = 172/100,000 (95% CI 162-182) RR = 0.96 (95% CI 0.78-1.19) Rectal CRC: Screen group = 230/100,000 (95% CI 219-241) Control group = 230/100,000 (95% CI 219-241) RR = 1.0</td>
<td>Proportion of screening group undergoing further testing = 4,299/62,677 (6.9%)</td>
</tr>
</tbody>
</table>
Table 8. Primary research studies appraised investigating the effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Kronborg et al. (2004) | RCT Grade II | Intervention Guaiac FOBT (Hemoccult II) Non-rehydrated, diet-restricted. Biannual testing (9 rounds total). Those with positive FOBT (one or more from six slides positive) were invited for medical review and full colonoscopy (or DCBE if colonoscopy refused or incomplete). Those with CRC or adenomas detected were managed appropriately and then invited to ongoing surveillance programmes. Only screeners who participated in previous rounds (and without CRC or adenomas) were re-invited. Control Not informed about study and used usual healthcare facilities. Each screening round was completed within one year. Study participants from both groups followed from 1965-August 2002. Analysed on an intention-to-treat basis. | Recruited participants from 140000 general population of Funen, Denmark aged 45-75 years. Excluded: Participants from a previous pilot study. Those identified from public registers as having CRC, colorectal adenomas, or distant spread of any malignant disease. 137485 were eligible. Husbands/wives allocated to same group Those eligible were randomised by computer (in 14 block groups) to: Intervention group—receive a mailed invitation (plus 2 reminders) to participate in FOBT screening (n=30967) Control group—no invitation (n=30966) Non-enrollment in study (n=75552) No significant difference between intervention and control groups in age (mean age = 59.3 years) or sex (male:female = 0.92) at enrolment. No significant difference between groups at time of each screening round. Mean age increased from 59.8 to 73.0 years, while male:female ratio decreased from 0.92 to 0.78 during the screening rounds. CRC mortality Anatomical sub-site CRC mortality All-cause mortality Incidence of CRC For screening group: Cumulative risk of having positive FOBT at least once (positivity rate) Cumulative proportion undergoing colonoscopy Participation rate of those invited for each screening round Verified diagnosis of CRC/adenoma from Funen patient database, public health officer and National Registry. Verified cause of mortality as CRC from pre-specified criteria set to assign CRC (including complications attributable to CRC) as cause of death. Investigator review of death certificates was conducted blind to participant screening status, with independent committee review if unclear whether CRC was cause of death. On August 2002: 18762 were alive of the original 30967 in screening group 18716 were alive of the original 30966 in control group. CRC mortality alone: Ratio* = 0.84 (95% CI 0.73-0.96) CRC total mortality (including treatment complications): Crude ratio* = 0.89 (95% CI 0.78-1.01) Adjusted for age and sex: Ratio* = 0.89 (95% CI 0.78-1.02) Effect of screening independent of age (p = 0.77) and sex (p = 0.82) Non-significant tendency for screening effectiveness in preventing death from proximal CRC compared to distal CRC (p = 0.13). All-cause mortality Ratio* = 0.99 (95% CI 0.97-1.02) Incidence of CRC Ratio* = 1.02 (95% CI 0.93-1.12) For screening group: Cumulative risk of having positive FOBT at least once = 5.7% (1.766 of 30.762) Cumulative proportion undergoing colonoscopy at least once = 5.3% (1647 of 30762) Participation of those invited for screening: 67% round 1, 91-94% rounds 2-9. *Ratio of screen group compared with control group. | Authors’ conclusions The present total mortality ratio for death from CRC (including treatment complications) of 0.89 has increased compared with that after seven screening rounds (0.82) as well as that following five screening rounds (0.86). The most probable explanation is the decrease in number of subjects being screened with increasing number of rounds. No more than 43% of the 19644 subjects still alive at the beginning of this 9th round were actually screened. An attempt to avoid this decrease in compliance has been made in other studies by re-invitation, with limited success. Interval cancers still fared better than controls. | Effectiveness And Cost-Effectiveness Of Population Screening For Colorectal Cancer
SECONDARY RESEARCH

The search strategy identified nine relevant reviews (published in 13 articles) that considered the effectiveness of FOBT screening for colorectal cancer. The methods and conclusions are described in Table 9. As discussed in the Methodology chapter above, these papers may not have employed the same inclusion and exclusion criteria as applied in this review and the results must be interpreted with care.

The report by AHTAC (1997) was considered by the Working Party on Screening for Colorectal Cancer (1998). Briefly, this inclusive report conducted for the Minister of Health and Family Services aimed to assess the evidence on the benefits, risks and costs of CRC screening for asymptomatic people (as well as surveillance for those at higher-risk of CRC). The committee considered the evidence up to 1996 for screening with sigmoidoscopy, colonoscopy, barium enema and genetic testing as well as FOBT, using established grading criteria to rank the study types considered for each modality. FOBT screening was the only testing modality for which RCT evidence was available, and only limited data were considered relating to risks resulting from FOBT screening. As a result of its review process AHTAC (1997) recommended that, subject to favourable preliminary testing, Australia develop a programme for the introduction of population screening for CRC using FOBT.

The aim of an Evidence Report/Technology Assessment prepared by the Conseil d’Évaluation des Technologies de la Santé du Québec (CETS) (2000) was to examine in detail evidence for the efficacy and efficiency of different strategies for CRC screening in terms of mortality reduction. The report focussed on CRC in average-risk, asymptomatic people. The report appraised in detail the RCT mortality data available up to 2000 for FOBT testing at that time. CETS (2000) concluded that a CRC screening programme, modelled in principle on the designs of the British and Danish FOBT studies, would significantly reduce mortality from this type of cancer. However, the authors were also clear that such a programme would be based on the existing evidence and that in light of ongoing research it would be advisable to explore other screening strategies, such as the addition of FS to FOBT. The authors then listed pre-conditions that would apply to the implementation of such a programme, including the requirement for feasibility studies to determine the mechanisms and costs of CRC screening.

The journal article compiled by Piedbois and Buyse (2000) provides a review of meta-analyses in colorectal cancer published between 1997-1999, but only reports limited information about the search strategy used. The review included one meta-analysis of CRC screening using FOBT – that conducted by Towler et al. (1998). Piedbois and Buyse (2000) presented only the main results and conclusions derived from the authors of the included meta-analysis, and did not draw any further conclusions about CRC screening.

The review conducted by Craven (2001) focussed on FOBT for asymptomatic CRC population-screening, and included an evaluation of evidence from controlled trials or meta-analyses regarding the influence of FOBT screening on mortality from CRC. Articles published in English between 1980-2000 were considered for inclusion in this review. This review identified the major RCTs that have examined the influence of FOBT screening on CRC mortality, judging the Nottingham and Funen-1 trials as providing the most robust and methodologically sound evidence. The reviewer concluded that despite evidence that CRC mortality can be reduced by approximately 16 percent as a result of FOB screening, mediocre population compliance and questionable Hemoccult performance remain concerns.

A systematic review conducted by McLeod, with the Canadian Task Force on Preventive Health Care (2001a), adapted for a later publication in the Canadian Journal of Gastroenterology (McLeod and members of the Canadian Task Force on Preventive Health Care, 2001b), evaluated the effectiveness of specific screening techniques for CRC in asymptomatic individuals at normal risk. Papers published up to January 2001 were considered for this review, which resulted in the inclusion of data from the four RCTs of FOBT screening versus no screening. The authors summarised the important features of these trials, and included unpublished mortality data from the Göteborg RCT cited from the meta-analysis by Towler et al. (1998). They also pointed out that after 18 years of follow-up, data from the Minnesota RCT had shown a decreased incidence of CRC in intervention compared to control group. McLeod, with the Canadian Task Force on Preventive Health Care (2001a), did not present a meta-
analysis, but concluded that an overall mortality reduction of 15 percent was found, and that in absolute terms, approximately 8.5 deaths from CRC would be prevented if 10,000 people were screened over 10 years.

The comprehensive systematic review conducted by Pignone et al. (2002a) was adapted for journal publication in the Annals of Internal Medicine (Pignone et al., 2002b), for inclusion in an Agency for Healthcare Research and Quality report (Pignone et al., 2002c). It informed the development of the most recent US Preventive Services Task Force recommendations and rationale for CRC screening (US Preventive Services Task Force, 2002). The selection criteria for review were not limited to RCT evidence alone, and other study designs for screening modalities were included if RCTs were not available. Rather than presenting a meta-analysis, this review presented the outcomes of the individual RCTs, related to their different study designs. Potential harms of FOBT screening were discussed by Pignone et al., (2002a), including possible adverse effects of false-positive or false-negative results, but no outcome data for harms specific to FOBT screening were included in this review. The final conclusion that providers and patients should discuss and incorporate patient preferences for CRC screening test is of limited relevance to the current health system in New Zealand, where opportunistic screening for CRC is not currently recommended.

A review prepared by the University of California (Walsh and Terdiman, 2003) and published in JAMA considered the status of the evidence for colorectal cancer screening, including methods of screening that may become available in future. For FOBT screening, it included the reports of clinical outcomes published to 2002 from the three large, long-term RCTs of FOBT screening – Minnesota (Mandel et al., 1993), Nottingham (Hardcastle et al., 1996) and Funen-1 (Kronborg et al., 1996). The review reported that serial FOBT screening decreased CRC mortality from between 15-30 percent depending on the frequency of screening, reported the absolute risk reduction, and noted that the Minnesota RCT had demonstrated that both annual and biennial FOBT screening reduced CRC incidence. Much of the review section on FOBT screening discussed issues relating to how the test is performed; safety issues or potential or real harms arising from FOBT screening were not assessed. In their final comments on CRC screening, the authors concluded that direct and indirect evidence indicated that the available tests for CRC screening were effective, but differed in their sensitivity, specificity, cost and safety, and that the available evidence did not support choosing one test over another.

One of the objectives of the Quebec Association of Gastroenterology Task Force was to evaluate the testing methods available (Barkun et al., 2004). FOB testing was one of the modalities included in the committee’s evaluation, which used a process of study selection by committee members followed by full committee review of selected studies to determine the most relevant data applicable to screening for CRC. Although the committee found the most evidence for FOBT screening, in its eventual recommendations it stated that there would be limitations to a screening programme for Quebec based on the use of this modality.

Towler et al.’s (2004) thorough systematic review and meta-analysis of CRC screening trials using the Hemoccult FOBT for the Cochrane Library, which was adapted for publication in the BMJ (Towler et al., 1998), was most recently updated on 1 September 2003. However, the most recent substantive update was undertaken on 6 September 1998, and therefore the search for this review included only data available up to January 1997. The primary result of this meta-analysis was that FOBT screening with the Hemoccult test reduces CRC mortality by 16 percent (RR 0.84, 95% CI 0.77-0.93). It is difficult to see why an uncontrolled trial conducted in Burgundy, France was included in this review, given the lack of relevant outcome data that was available from that trial at the time. The findings of Towler et al.’s (2004) meta-analysis could be criticised on the basis that there were several methodological differences between the included trials, and that the number of trials included in the meta-analysis was small. However, the methodological conduct and reporting of the overall review are difficult to fault otherwise, lending weight to the conclusions of this review. It is the only secondary research article included in this chapter for which the reviewers personally sourced and included unpublished mortality data (from the Göteborg RCT). These reviewers also highlighted in their discussion the general paucity of reported data regarding harms from FOBT screening (Towler et al., 2004).
CONCLUSIONS

The Working Party on Screening for Colorectal Cancer (1998) mentioned a fourth RCT comparing Haemoccult screening to no screening, conducted in Göteborg (Kewenter et al., 1994, in Working Party on Screening for Colorectal Cancer, 1998). Since that time, no mortality results have been published from this trial (although they were expected in 1999) and no further outcome data from this RCT was available for appraisal. Towler et al. (2004) included unpublished mortality results from the Göteborg trial (by contacting the trial investigators as part of their search strategy, see Table 9) that were available in 1998 in their meta-analysis. The main result of their meta-analysis was that for biennial FOBT screening with the Haemoccult test compared to no screening, the overall reduction in risk of death from CRC was 16 percent [RR = 0.84 (95% CI 0.77-0.92)].

Fletcher (1999, 2002) has commented on the topic of meta-analyses that combine CRC mortality data from the strong studies of screening efficacy. Fletcher made the point that pooling the results from such trials was misleading, as the questions addressed by each RCT were different, and further noted that the trial results suggested that biennial screening for > 10 years reduced relative risk by 18-21%, whether or not the test kit was rehydrated. Despite Fletcher’s (1999, 2000) comments, a comparison of the efficacy results for haemoccult screening derived from the trials included in Towler et al.’s (2004) meta-analysis, and the results of the statistical tests that were used to evaluate the meta-analysis did not suggest heterogeneity between trial results. Therefore, in the context of this review, a meta-analysis of the subsequent CRC mortality results reported from the most recent data published from the population-based RCTs of biennial haemoccult screening (Scholefield, 2002; Kronborg, 2004) was conducted.

The present meta-analysis was done using Stata version 7.0 (StataCorp, 2001). A fixed-effects model was used, as there was no significant heterogeneity identified between the results from the RCTs included. The data used for the analysis were the rate of death from CRC in screen compared to control groups. This did not include complications from treatment of CRC, which was reported as a separate additional result by Kronborg et al. (2004). The result of this meta-analysis was a rate ratio of death from CRC in screen compared to control groups of 0.86 (95% CI 0.79-0.93). In other words, screening with the haemoccult FOBT reduced the risk of death from CRC by 14%. It should be noted that adding the results for biennial haemoccult screening from the Minnesota RCT (Mandel, 1999) to the meta-analysis did not make a consequential difference to the meta-analysis result.

As the Working Party on Screening for Colorectal Cancer (1998) found, there is high-quality, statistically significant evidence that FOBT screening with the guaiac test Haemoccult FOBT reduces mortality from CRC. In addition, the evidence that biennial Haemoccult screening can reduce CRC mortality is further supported by the more recent data reported from the Minnesota trial, which as discussed above had not found a significant risk reduction after 13 years of follow-up, but did after 18 years of follow-up. Evidence available from ongoing follow-up of the three Haemoccult RCTs suggests that the initial mortality reductions found have been sustained for the populations in which screening has stopped, but decreased for the Funen-1 trial population, to whom screening has continued to be offered. However, some caution has to be taken when applying results from static RCT study populations to actual screening programmes, as not only are no new screenees invited, but the study population ages as it progresses, and the chance of screenees having a colorectal lesion decreases with each round of screening.

As detailed above and in Table 8, the direct evidence available from the Jiashan RCT concerning the efficacy of an immunochemical test to reduce CRC mortality is less robust. Although the data from this RCT cannot be assumed as completely applicable to screenees aged 50-75 years, it does suggest that a reduction in rectal cancer may be achievable with the use of this test. It can be further postulated that if all screenees with a positive FOBT in this trial had been assessed by colonoscopy instead of FS, then a reduction in risk for colonic cancer may also have been found.

Some new evidence, as detailed above, regarding the benefits and harms from CRC screening has become available from the Nottingham RCT. However, little new data regarding possible psychological harms became available as a result of this review.
Acceptability of FOBT screening would be one of the most crucial factors affecting success in reducing CRC mortality for a population-based screening programme. To replicate the mortality reductions found in the RCTs, participation in a screening programme would need to be equivalent or higher. The working party found that, for the Nottingham and Funen-1 trials, between 60-67 percent of those invited to screening from the general population participated in at least one round. Although the screening process was not strictly comparable, a similar participation rate (66.4%) was found by the Chinese investigators, who were also inviting members of the general population. Achieving adequate rates of participation for other screening programmes has not been easy for many countries, including New Zealand, and may represent one of the largest challenges for FOBT screening.

Similar conclusions can be drawn from the results of this review as those made by other reviewers, including the Working Party on Screening for Colorectal Cancer (1998). FOBT screening can reduce mortality from CRC. The evidence available since this topic was last considered in New Zealand has not changed substantially. The question of whether immunochemical tests provide more optimal performance than guaiac will be evaluated in Chapter 4.
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Health Technology Advisory Committee (AHTAC) [1997]</td>
<td>Search: 1990-1996. Supplemented by identification of key references for earlier articles. Databases searched: Medline, EMBASE, CancerLit. Also identified articles from the Internet and reference lists of retrieved articles. Current Contents searches were continued up to the date of publication (1997). Accessed “grey” literature by contacting health departments, academic institutions and reviewing submissions to the AHTAC Working Group. Key Words: colorectal cancer, screening, faecal occult blood test, barium enema, economics, costs, epidemiology, prevention, survival, acceptability, compliance.</td>
<td>All study types evaluating FOBT screening were considered, but greater weight was given to research material scoring highly on NHMRC (1995) guidelines. For RCTs of FOBT screening “quality appraisal criteria were applied to identified studies, but none were excluded”</td>
<td>4 RCTS of FOBT screening versus no screening: (Minnesota, Nottingham, Funen-1, Göteborg). Summary data is presented from the 3 RCTs that had reported mortality (Göteborg had not reported mortality data at time of review).</td>
<td>Review did not specify clear PICO questions. Authors’ conclusions (for FOBT screening) Because not all cancers bleed and not all bleeding is due to cancer, the simple FOBT will not always detect cancer when it is present and will detect some bleeding when it does not come from cancer. Nevertheless, on the basis of the published evidence showing that reductions in CRC mortality can be achieved through a population-screening programme using FOBT, it is recommended that Australia develop a programme for the introduction of population screening for CRC for the average risk population aged over 50 years. (Also see Chapter 4 for immunochemical versus guaiac FOBT screening).</td>
</tr>
</tbody>
</table>

Effectiveness and Cost-effectiveness of Population Screening for Colorectal Cancer
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conseil d’Évaluation des Technologies de la Santé du Québec (2000) Canada.</td>
<td>Search: 1966-1997. Databases searched: Medline, EMBASE, PASCAL, CancerLit. Key Words: Initially limited to screening and colorectal neoplasia for years 1966-1997. Colonic neoplasms then included for 1966-1987, from 1988-“present” (likely to be late 1999 or early 2000) query modified to include the terms colorectal or colonic or rectal combined with adenoma(s), adenocarcinoma(s), carcinoma(s), polyp(s), neoplasm(s) or cancer(s).</td>
<td>Review focused on asymptomatic people at average risk for CRC. The included studies were chosen on the basis of the level of evidence they provided. Priority given to RCTs, and other research designs were only mentioned if there was no RCT evidence nature, or as supplementary information. For FOBT screening versus no screening: 4 RCTS (Minnesota, Nottingham, Funen-1, Göteborg), 1 non-controlled trial (Burgundy), 5 case-control studies. [Göteborg RCT and Burgundy trial had not reported mortality data at time of this review]. Main Results from RCTs: PPV Non-rehydrated 1st Test 10-17% Hydrated 2.2% Mortality Presented the results of a meta-analysis using data from biennial screening participants from Minnesota (13-year follow-up), Nottingham (8 years) and Funen-1 (10 years) RCTS: Reduction in CRC mortality = 14%, OR= 0.86 [95% CI 0.78-0.95]. Using data from 18-year follow-up [Minnesota]: Reduction in CRC mortality = 16%, OR= 0.84 [95% CI 0.76-0.93]. Test positivity rate Non-rehydrated 1st test ranged from 1.0% - 2.1% Hydrated 9.8% Compliance For biennial screening, percentage of participants completing at least one FOBT ranged from 60%- 67% in general population samples, to 78% in volunteer sample.</td>
<td>Authors’ conclusions The authors discussed the results from meta-analyses conducted by other authors that include only the data from Nottingham and Funen-1 studies, which give comparable mortality reduction and odds ratio to their own results. The authors concluded that FOBT screening led to a reduction in CRC mortality, despite the lack of evidence for a decrease in CRC incidence or overall mortality. (Also see Chapter 4 immunochemical versus guaiac FOBT screening)</td>
<td></td>
</tr>
</tbody>
</table>
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piedbois and Buyse (2000) France and Belgium</td>
<td>Search: 1997-1999. Databases searched: CancerLit and Current Contents. Also reviewed reference lists of relevant papers, hand searched contents tables of selected journals and scanned websites of selected journals. Key Words: Not reported.</td>
<td>Inclusion criteria: Quantitative meta-analyses of comparative studies in management of CRC. Exclusion criteria: Reviews without a global and quantitative estimation of the difference between two procedures were excluded.</td>
<td>One meta-analysis of FOBT screening versus no screening. The meta-analysis, conducted by Towler et al., (1998) included 4 RCTs of FOBT screening (330,000 subjects). Mortality ‘ITT’ analysis Subjects allocated to screening by FOBT had a statistically significant reduction in mortality from CRC of 16%, RR = 0.84 (95% CI 0.77-0.93). ‘On-treatment’ analysis Subjects actually screened had a greater mortality reduction RR = 0.77 (95% CI 0.57-0.89).</td>
<td>Lack of information on the search strategy. Authors’ conclusions The authors praise the meta-analysis included in their article but comment that adding the two non-randomised trials (113,000 subjects) to the meta-analysis could have been avoided, as it did not bring any new evidence.</td>
</tr>
</tbody>
</table>
### Table 9. Secondary research appraised relevant to the effectiveness of FOBT screening on outcomes from CRC compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
- written in English  
- written/published between January 1980-January 2000  
- directly related to occult blood screening of asymptomatic population  
- that were a controlled trial or meta-analysis.  
Excluded articles:  
- relating solely to other methods of screening (e.g., FS, colonoscopy)  
- relating to other aspects of establishing a national screening programme (e.g., economic analysis)  
- consisting of ‘grey literature’  
- involved multiple or serial publications of the same data  
Methodological quality of selected articles was assessed using a specific appraisal format and hierarchy of evidence to grade articles according to research design. | 5 controlled trials and 1 meta-analysis reporting evidence that FOBT screening influences CRC mortality.  
5 controlled studies evaluating the effectiveness of Hemoccult test had mortality as primary outcome measure.  
Overall reduction in CRC mortality of between 6-18% with 2-yearly screening. Considers the Nottingham and Funen-1 RCTs as providing the most robust and methodologically sound evidence, with their respective reductions in mortality of 15% and 18% respectively.  
One meta-analysis (Towler et al., 1998), combined the data from the separate though related haemoccult trials, and demonstrated that those allocated to screening had an overall reduction in CRC mortality of 16%. | Methodologically sound review process.  
Authors’ conclusions  
The results of the controlled trials using Hemoccult screening reflect the advantage of repeated screening over a once-only test, followed by diagnostic evaluation using colonoscopy. It is possible that with multiple opportunities to detect cancer, even a low-sensitivity test such as Hemoccult may achieve a higher sensitivity. However, despite the recent evidence that CRC mortality can be reduced by approximately 16% as a result of FOBT screening, mediocre population compliance and questionable Haemoccult performance have meant that in effect the chance of detecting cancer is, on balance, only slightly greater than the chance of missing it.  
(Also see Chapter 4 for immunochemical versus guaiac FOBT screening). |
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>McLeod, with the Canadian Task Force on Preventive Health Care (2001a); McLeod and members of the Canadian Task Force on Preventive Health Care (2001b) Canada</td>
<td>Search: January 1966 -January 2001. Databases searched: Medline. Also identified articles from reference sections of review articles published before January 2001 and expert consultation. Key Words: screening, colorectal neoplasia.</td>
<td>Included articles concerning Hemoccult testing or FS as the first step in a multiphase CRC screening strategy, or colonoscopy as a single-phase screening strategy in asymptomatic people at average risk and high risk for CRC. Excluded articles concerning screening with digital rectal examination and DCBE. Outcomes: Effectiveness of CRC screening. Papers graded according to predetermined hierarchy of evidence. Discussed the critical appraisal of the evidence at expert meetings, to reach unanimous conclusions.</td>
<td>For Hemoccult screening 4 RCTs (Minnesota, Nottingham, Funen-1, Göteborg), comparing data for CRC screening versus no screening. 1 meta-analysis [Towler et al., 1998] Sensitivity For screening programmes with biennial testing (included hydrated and non-rehydrated kits), sensitivity ranged from 38-79%. Overall, the sensitivity of FOBT for detecting CRC was approximately 50% in three of the RCTs. PPV Ranged from 2.2-4.6% in the trials that included some subjects with rehydrated testing. Ranged from 10-17.0% in the trials using only non-rehydrated testing. Mortality From the 4 RCTs, in summary the relative risk reduction for CRC mortality for screened to control groups was around 15%. In absolute terms approximately 6.5 deaths from CRC would be averted if 10,000 people were screened over 10 years. Incidence After 18 years of follow-up, Minnesota RCT showed cumulative incidence ratio for CRC in screened to control groups of 0.80 (95% CI 0.70-0.90) for annual screening and 0.83 95% CI 0.73-0.94) for biennial screening. Compliance Ranged from 59-78%</td>
<td>Discrepancies noted between some data in text and tables in this report. Authors’ conclusions For individuals at normal risk, there is evidence to support the use of annual or biennial faecal occult blood testing over age 50 years. There remain concerns about the sensitivity of Haemoccult testing and its value as a screening test.</td>
</tr>
</tbody>
</table>
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pignone et al. (2002a);</td>
<td>Search: January 1966-September 2000. Updated by general and focused searches from January 1995-December 2001. Databases searched: Medline and British National Health Service Economic Evaluation database. Also identified articles from the second edition of the Guide to Clinical Preventive Services, existing systematic reviews, hand searches of reference lists of key articles and expert consultation. Key Words: colorectal neoplasms, occult blood, sigmoidoscopy, colonoscopy combined with mass screening, then colorectal neoplasms combined with barium sulfate, enema.</td>
<td>Population: average-risk adults over the age of 50 years</td>
<td>4 RCTs (Minnesota, Nottingham, Funen-1, Göteborg), 1 systematic review (Towler et al., 1998).</td>
<td>Authors’ conclusions For FOBT, 3 high-quality randomised controlled trials have shown disease-specific mortality reductions of 15-33% over 8 to 13 years. Although CRC screening is supported by strong direct and indirect evidence, current data are insufficient to define which strategy is most effective or cost-effective. Providers and patients may benefit from discussing the pros and cons of the different methods and incorporating patients’ preferences in the decision about how to screen.</td>
</tr>
<tr>
<td>(2002b); Pignone et al.</td>
<td></td>
<td>Intervention: ‘home’ FOBT screening</td>
<td>Sensitivity</td>
<td></td>
</tr>
<tr>
<td>(2002c) USA.</td>
<td></td>
<td>Outcomes: benefits and adverse effects. Adverse effects included complications of the screening test, false positives, and economic costs. Inclusion criteria developed to guide decisions about inclusion of articles, generally seeking to identify and include the highest quality evidence available. For ‘home’ FOBT screening included randomised controlled trials (for other screening modalities, observational and diagnostic accuracy studies were evaluated when RCTs were not available). For adverse effects (of any test) included RCTs, observational studies and case series. Two reviewers reached consensus decision on which abstracts to retrieve and then which full text articles to include.</td>
<td>Minnesota screening programme, 13-year follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*differentiated from ‘office’ FOBT screening, which reviewers defined as that undertaken by a physician during digital rectal examination.</td>
<td>• annual testing Se=49%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial testing Se=39%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nottingham and Funen-1 screening programmes (unrehydrated FOBT), 8-10 years of follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial testing Se=27% (49% of cancers occurring in participants)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mortality</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Percentage reduction* for intervention compared with control group</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minnesota (rehydrated FOBT), 18 year follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• annual 33%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial 21%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nottingham (unrehydrated FOBT) - years follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial 15%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Funen-1 (unrehydrated FOBT) 10 years follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial 18%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Göteborg (rehydrated) after 2 rounds of testing found a non-significant percentage reduction of 12% (Cited from Towler et al., 1998).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Incidence</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Percentage reduction* for intervention compared with control group</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minnesota (rehydrated FOBT), 18-year follow-up:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• annual 20%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• biennial 17%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Acceptability</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(measured by adherence to FOBT testing in RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Approximately 50% of participants completed all tests in the series</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Approximately 80% of initial acceptors completed the second test in the series. *statistically significant</td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Search method</td>
<td>Selection criteria</td>
<td>Results</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Walsh and Terdiman (2003)</td>
<td>Search: January 1966-August 2002. Databases searched: PubMed. Also identified articles from published meta-analyses, reference lists of key articles and expert consultation. Supplemented search with hand searching and proceedings from national professional organisation meetings. Key words: colorectal neoplasia, occult blood, sigmoidoscopy, barium enema, colonoscopy all combined with screening.</td>
<td>Population: asymptomatic subjects who were at average risk for CRC. Intervention: FOBT screening. Outcomes: clinical outcomes considered included mortality, cancer incidence, and identification of adenomas. Inclusion criteria: English language For FOBT screening included randomised controlled trials (for other screening modalities, observational and diagnostic accuracy studies were evaluated when RCTs were not available). Exclusion criteria: Not reported.</td>
<td>For FOBT screening: 3 RCTs (Minnesota, Nottingham, Funen-1) Mortality Results involving more than 250,000 subjects followed for up to 18 years, consistently demonstrated that serial FOBT screening: • reduces colorectal cancer mortality by 15-33% • results in an absolute risk reduction for CRC death from between 0.8 per 1000 person-years (with biennial screening) up to 4.6 per 1000 person-years (with annual screening) Incidence One RCT (Minnesota) demonstrated that FOBT screening reduces CRC incidence by 17-20%, after 18 years of follow-up.</td>
<td>Not restricted to published literature. Several internal review validity processes were not reported, including how study selection was undertaken, how data was extracted, or how assessment of internal validity of included studies was undertaken. Authors’ conclusions FOBT performed on a single occasion for the detection of CRC and adenomas shows poor sensitivity, but the key to the success of FOBT lies in serial testing. The recommendation that all men and women aged 50 years or older undergo screening for colorectal cancer is supported by a large body of direct and indirect evidence. At present, the available evidence does not currently support choosing one test over another.</td>
</tr>
<tr>
<td>Source</td>
<td>Search method</td>
<td>Selection criteria</td>
<td>Results</td>
<td>Comments</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Barkun et al. (2004) Canada</td>
<td>Search: 1966-2002. Databases searched: Medline. Key Words: Randomized clinical trials, clinical trials, colorectal cancer, screening, patient outcomes, reviews, meta-analyses, cost-effectiveness, systematic reviews.</td>
<td>Each committee member was assigned a sub-group of topics and selected all the RCTs and case-control studies relevant to CRC screening. Selection of additional studies for that topic was left up to that member. All data then presented to the whole committee and final decision to include or exclude article was made by committee. Included articles were graded from Levels I to III according what level of evidence they contained.</td>
<td>4 RCTs assessing the Hemoccult test (Minnesota, Funen-1, Nottingham and Göteborg) 1 systematic review (McLeod et al., 2001b) Found Level I evidence (from properly conducted RCTs) that screening with FOBT reduces mortality in asymptomatic patients who are over age 50 years. The RR of CRC death is 0.84 (95% CI 0.77-0.93) overall, and 0.77 (95% CI 0.57-0.89) in compliant individuals.</td>
<td>Only searched one database. Authors’ conclusions Although the most clinical efficacy data is available for FOBT, this strategy is limited by the high-false positive rate, low compliance and low detection yield. There remain concerns about the feasibility and small clinical benefits of such screening.</td>
</tr>
</tbody>
</table>
Table 9. Secondary research appraised relevant to effectiveness of FOBT screening on outcomes from CRC, compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Towler et al. (2004)</td>
<td>Search: up to January 1997</td>
<td>Population: Adults participating in Hemoccult screening trials worldwide (volunteer participants or those identified from a population register)</td>
<td>6 controlled trials</td>
<td>Not restricted to published literature. Meta-analysis combined data from RCTs that differed in several ways: different screening intervals, used rehydrated and non-rehydrated FOBTs, duration of follow-up, which may effect the analysis results.</td>
</tr>
<tr>
<td>Towler et al. (1998)</td>
<td>Databases searched: Medline, Current Contents, Cochrane Controlled Trials Register, Supplemented search by retrieving references from a previous review by Towler et al. (1995) referenced as a background paper and by writing to the principal investigators from the six known screening trials requesting unpublished data and clarification of trial methods and results.</td>
<td>Intervention: Hemoccult FOBT screening (annual or biennial; rehydrated or non-rehydrated; Hemoccult as additional to FS screening or alone).</td>
<td>4 RCTs (Minnesota, Nottingham, Funen-1, Göteborg)</td>
<td>Authors’ conclusions The estimate of mortality reduction from the Haemoccult trials is now well quantified and the confidence intervals are narrow enough to allow the conclusion that colorectal cancer screening is likely to be of net benefit for some population groups. Based on the overall trial results, if a biennial screening programme were offered to 10,000 people, of whom approximately two-thirds attended for at least one Hemoccult test, 8.5 colorectal cancer deaths would be prevented over approximately 10 years. There are some important areas needing research, including that there is insufficient information about the harmful physical and psycho-social effects of screening, whether the necessary quality of screening and follow-up be achieved outside the trials, and whether screening expenditure would constitute a sound use of resources given local priorities in health care.</td>
</tr>
<tr>
<td>Australia</td>
<td>Key Words: colorectal neoplasms, screening, randomised controlled trial (RCT), controlled clinical trial (CCT).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Towler et al.</td>
<td>Search: up to January 1997</td>
<td>Population: Adults participating in Hemoccult screening trials worldwide (volunteer participants or those identified from a population register)</td>
<td>2 non-randomised CTs (Burgundy, New York). Burgundy had not reported data for intermediate outcomes or mortality.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Databases searched: Medline, Current Contents, Cochrane Controlled Trials Register, Supplemented search by retrieving references from a previous review by Towler et al. (1995) referenced as a background paper and by writing to the principal investigators from the six known screening trials requesting unpublished data and clarification of trial methods and results.</td>
<td>Intervention: Hemoccult FOBT screening (annual or biennial; rehydrated or non-rehydrated; Hemoccult as additional to FS screening or alone).</td>
<td>2 non-randomised CTs (Burgundy, New York). Burgundy had not reported data for intermediate outcomes or mortality.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Main outcome: Effectiveness of screening for each trial overall was evaluated by comparing CRC mortality in the screen and control groups.</td>
<td>Effectiveness of screening</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate outcomes:</td>
<td>Annual Screening</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Hemoccult sensitivity for CRC</td>
<td>Minnesota RR = 0.67 (95% CI 0.51-0.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• CRC stage distribution in screen and control groups</td>
<td>New York regular attenders RR=1.18 (95% CI 0.48-2.88), first time attenders RR = 0.63 (95% CI 0.31-1.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• CRC incidence in screen and control groups</td>
<td>Biennial Screening</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Physical harms (complications) of follow-up colonoscopy or FS</td>
<td>Minnesota RR = 0.95 (95% CI 0.74-1.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Proportion of those allocated to screening who actually attended screening.</td>
<td>Nottingham RR= 0.86 (95% CI 0.74-0.99)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inclusion criteria:</td>
<td>Funen-1 RR = 0.82 (95% CI 0.68-0.99)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>All randomised and non-randomised controlled trials with above population (randomisation may be of individuals or groups) that have available mortality results (published or unpublished).</td>
<td>Göteborg RR = 0.88 (95% CI 0.69-1.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exclusion criteria:</td>
<td>Meta-analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No trials of Hemoccult screening were excluded.</td>
<td>Overall (intention to screen) RR = 0.84 (95% CI 0.77-0.92). Chi² for heterogeneity 0.33, df4, p &gt; 0.5.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Three authors contributed to assembling a set of articles of trials that met the inclusion criteria. Trials independently assessed for quality by two authors Data from trials independently extracted by two authors.</td>
<td>Exclude New York data RR= 0.84 (95% CI 0.77-0.93)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adjust for those who attended for screening RR = 0.77 (95% CI 0.57-0.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intermediate outcomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Se: range = 46% (unrehydrated) to 92% (hydrated). PPV: range 2.2-4.2% (rehydrated), range 5.6-17.7% (unrehydrated)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRC stage distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>All 5 trials found more early stage (Duke’s A) and less late stage (Duke’s C/ C&amp;D) CRC in screen compared to control groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRC incidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Similar cumulative rates for screen and control groups in all trials after 8-13 years of follow-up.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Physical harms (perforation/haemorrhage per 10,000 procedures)</td>
<td>Minnesota = 12 per 10,000 colonoscopies Göteborg = 30 per 10,000 endoscopies (colonoscopy +FS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proportion who attended at least one screening round</td>
<td>Range = 60-90% (higher in Minnesota RCT, with volunteer population, than European trials).</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Accuracy of FOBTs

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY

The search identified seven eligible primary research studies. Below is an overview of study designs and aspects of quality represented by these studies. Full details of the papers appraised, including methods, key results, limitations and conclusions, are provided in evidence Table 10. Studies are presented in chronological order of publication within the table.

Study designs and quality assessments

Of the seven eligible studies comparing immunochemical FOBT to guaiac FOBT screening, six were graded evidence level III-2. Of these six studies, three were of cross-sectional design (Rozen et al., 1997; Rozen et al, 2000; Cheng et al., 2002), one was a case-control study (Saito et al., 2000), one was a cohort study that appeared to have been conducted retrospectively (Zappa et al., 2001) and one was a non-randomised clinical trial (Ko et al., 2003). The other eligible study was a randomised controlled trial (Cole et al., 2003), and therefore was graded evidence level II.

The cross-sectional studies conducted in Israel (Rozen et al., 1997 and 2000) appear to have employed the same study design, the earlier study comparing four FOBTs, and the later study comparing two FOBTs. These consecutive studies were of reasonable quality in their design and conduct, with adequate blinding of those interpreting FOBT results. However, the internal validity of both studies was threatened by the potential for review bias in assessing outcomes, as it was not clear whether those who conducted and interpreted the follow-up tests used as reference standards for the FOBTs were blind to the subjects’ FOBT results. Furthermore, results from lower bowel endoscopies that had been conducted prior to FOB testing, whether FS or colonoscopy, were taken as a reference standard for a substantial proportion (approximately 41%) of the study subjects with a negative FOBT. The reasons for these inconsistencies in both method of large bowel assessment and timing of assessment were not well explained in the reports, and undermined the quality of both of these studies. Although the cross-sectional study undertaken in Taiwan by Cheng et al. (2002) that compared two FOBTs was of reasonable design, the study cannot be graded as being of good quality. This is because the report did not specify whether either FOBT or reference standard interpretation were conducted blind.

The study conducted in Japan that employed a case-control design compared the performance of two FOBTs by calculating the odds ratios of dying from CRC for those screened by each test (Saito et al., 2000). Small numbers of cases and controls limit the strength of this study’s findings. Important differences between the FOBT screening processes for each test, in particular significantly different compliance rates for diagnostic work-up following a positive FOBT, may have further distorted the results of this study. The intended design of Zappa et al.’s (2001) study was not reported in their paper, but appears to have been of retrospective cohort design. By establishing the screening histories of participants involved in a screening programme, and then linking this information to new diagnoses of CRC, comparisons were made between the detection rates for CRC of a guaiac FOBT screening test and two immunochemical tests, compared to the number of cases of CRC that would have been expected in the absence of screening. Although the results of this study are highly significant, potential biases between the cohorts studied cannot be reliably excluded, reducing the quality of the evidence from this study.

The inherent potential for bias present in the non-randomised trial design utilised by Ko et al. (2003) means that the results of this study provided only mediocre evidence with regards to comparative FOBT performance. The reliance of the investigators on verification of outcome results from medical notes, and apparent low numbers of participants undergoing further work-up after positive FOBT results also weakened the reliability of their findings.

The high-quality study design, methods and reporting of the RCT conducted by Cole et al. (2003) mean that the evidence attained regarding the FOBT comparisons that they made was highly valid. The applicability of their findings to other populations is however weakened by the low participation rates.
The quality of evidence from many studies of diagnostic techniques has been questioned by authors such as Knottnerus et al. (2002), who commented that methodological flaws are common in diagnostic studies. To compare validity and reliability between diagnostic tests, the ideal study design is a cross-sectional survey in which both the tests being compared, as well as a reference standard, are performed (Greenhalgh, 1997). The cross-sectional studies appraised for this review were of reasonable or low quality, therefore limiting the conclusions that can be drawn from their results. To compare other aspects of diagnostic tests’ usefulness for population screening, such as acceptability, ideally requires a prospective comparison between using each test, for which the randomised controlled trial is the most robust method (Knottnerus et al., 2002). This review identified only a small amount of evidence, from one randomised controlled trial which compared FOBTs in this manner.

**Study settings and samples**

Three of the studies recruited participants from those attending clinical services that undertook CRC screening as well as providing other clinical services (Rozen et al., 1997, Rozen et al., 2000; Ko et al., 2003). Although the ages of participants in all of these studies were broadly similar, at approximately 60-65 years, to that of interest to this review, the samples for all three of these studies included participants who had symptoms that could have been attributable to CRC. Despite the fact that both study populations recruited by Rozen et al. (1997 and 2000) were mainly (97%) asymptomatic, only 22 percent of so-called asymptomatic screenees were classified as being of average risk. Given that a higher prevalence of detectable disease in a study population will be expected in a population that includes those who have a higher likelihood of having the disease of interest, the predictive values of the tests compared in these studies will have little relevance to their use in screening asymptomatic, average-risk general populations. This limits the applicability of the study findings from all three of these studies to the question of interest to this review.

Three other studies evaluated FOBTs as used in population-based screening programmes. The screening programme that provided the setting for the cross-sectional study conducted by Cheng et al. (2002) required voluntary attendance (Cheng et al., 2002). The sample in this study included only asymptomatic screenees and excluded those with personal history of CRC, but did not exclude other groups who may have been at higher risk for CRC such as those with past history of polyps or family history of CRC. Combined with the inclusion of a young population age, the diversity of the study sample makes the test predictive values calculated rather meaningless and the applicability of the results hard to clarify. There were no data presented to allow calculations for different study sub-populations.

The screening programmes that provided the setting for the other two studies (Saito et al., 2000 and Zappa et al., 2001) specifically invited members of certain ages (40 years and over, or between 50-70 years respectively) from the general population to attend for CRC screening tests. Their results regarding comparative FOBT performance are thus most applicable to the CRC screening of asymptomatic average-risk individuals aged 50-74 years of age. Cole et al. (2003) recruited participants aged 50-60 years from the general population of Adelaide but, as noted in Table 10, the low overall response rate for this RCT limits the applicability of the trial results.

In summary, four of the appraised studies provided evidence from population-based settings, and of these only the populations of three studies could be considered as broadly representative of asymptomatic adults at average risk for CRC.

**Interventions, comparators and study outcomes**

All of the studies compared immunochemical FOBT to guaiac FOBT as screening tests. The guaiac test for which robust evidence exists from screening RCTs concerning reduction in CRC mortality is the Haemoccult/Haemoccult II.

**Test positivity rates**

For the three papers appraised that studied populations that could be considered as broadly representative of asymptomatic adults at average risk for CRC (Saito et al., 2000; Zappa et al., 2001; Cole et al., 2003), test positivity rates are detailed in Table 10. For the Haemoccult II (HOII) test Saito et al. (2000) reported that 5.7 percent of tests were positive compared to 2.7 percent for the
immunochemical test (HemSp) they studied, but did not detail positivity rates for first compared to subsequent screening rounds. Zappa et al. (2001) reported an overall positivity rate for HOII of 4.5 percent (first screen round 4.7%, subsequent rounds 4.3%), compared to the overall positivity rate of 4.0 percent (first round 4.5%, subsequent rounds 3.7%) for the immunochemical tests they used. Cole et al. (2003) used the guaiac test HaemoccultSENSA (HOS), reporting that 5.7 percent were positive in the one-off screen offered in their RCT. This did not differ significantly from the percentages of positive tests reported for the two immunochemical tests they studied (Insure = 6.8%, FlexSure = 4.6%).

### Haemoccult/ Haemoccult II

Three of the studies used the Haemoccult II (HOII) test. None of these studies reported rehydrating the HOII test cards, and two were explicit that they required two samples from each of three consecutive stools. However, different pre-HOII dietary requirements and testing/re-testing strategies restrict the comparability of these studies. The comparisons made between immunochemical tests and HOII were also diverse. One comparison was made of HOII with HemeSelect alone, one comparison with Flexsure alone (these comparisons were both made in the study conducted by Rozen et al. (1997)), one comparison with HemSp alone and one comparison of HOII where the results from both the HemeSelect and HemSp tests were grouped together.

Only one study compared HOII’s single-use test performance with immunochemical testing (Rozen et al., 1997). In this study HOII was less sensitive for any colorectal neoplasms (Se = 27%) and for neoplasms larger than 1cm and CRC combined (Se = 63%) than HemeSelect, which had sensitivities for the same lesions of 34 percent and 86 percent respectively. Conversely HOII was of comparable sensitivity to the other immunochemical test FxS (Se = 63%) for larger adenomas and CRCs. As outlined in Table 10, these tests were all of similar specificity (Sp= 95-99%) for the colorectal lesions considered.

Zappa et al. (2001) used the proportional incidence method to calculate two-year test sensitivity. Although (as the authors pointed out) this method can under-estimate programme sensitivity, it did allow for the main goal of their study, which was to compare HOII with the immunochemical tests. Compared to the use of either HemeSelect or HemSp (which are both immunochemical tests that use a reverse passive hemagglutination process to detect the presence of blood in the stool) in their screening programme, HOII (Se = 50%, 95% CI 34-63) performed significantly worse than the immunochemical tests (Se = 82%, 95%CI 67-92) with respect to overall programme sensitivity for CRC (Zappa et al., 2001). The authors did not explain how, or whether, these immunochemical tests differed from one another. Nevertheless, when the results of both of these studies are considered together, it can be postulated that HOII is less sensitive for serious colorectal lesions that include CRC than the HemeSelect test for CRC, despite the comparable specificity of these tests.

The case-control study conducted by Saito et al. (2000) offered the third comparison between HOII and an immunochemical test, suggesting that HemSp may offer reduced likelihood of dying from CRC than HOII. However, the confidence intervals that surrounded the odds ratios for both the tests were wide, even for those screened within one year prior to death from CRC. The study by Saito et al. (2000) therefore added little additional robust evidence about the comparative performance of HOII.

### HaemoccultSENSA

Four studies evaluated HaemoccultSENSA (HOS) compared to immunochemical tests. Three of these reported outcomes of diagnostic test performance comparisons (Rozen et al., 1997 and 2000; Ko et al., 2003) with the fourth evaluating participation as its primary outcome (Cole et al., 2003).

HOS was compared to Flexsure OBT (FxS) in all four of these studies. Both studies by Rozen et al. (1997, 2000) found the single-use test sensitivity of FxS (Se= 18-23%) for any colorectal neoplasms to be lower than that of HOS (Se = 31-35%), with the 2000 study finding a significant difference between these proportions. For neoplasms larger than 1cm and CRC combined, there was no significant difference, although HOS (Se= 50-63%) did perform slightly better than FxS (Se= 35-63%). FxS consistently had a higher test specificity (Sp = 97-99%) than HOS (Sp = 92-95%) for both categories of lesions considered in these studies, and this difference was of a statistically significant difference in the later study (Rozen et al., 2000). Although Ko et al. (2003) also compared the test performances
between HOS and FxS, finding no significant differences between the tests’ positive predictive values, as detailed in Table 10, the robustness and applicability of their findings is questionable.

The earlier cross-sectional study conducted in Israel (Rozen et al., 1997) also included a comparison of HOS to the HemeSelect. As with HOII, HemeSelect demonstrated substantially better sensitivity for any colorectal neoplasms (Se = 35%) and for neoplasms larger than 1 cm and CRC combined (Se = 86%) than HOS (Se = 31% and Se = 63% respectively). HemeSelect was also more specific (Sp = 98-99%) than HOS (Sp = 92-93%).

Previous authors have commented that newer guaiac tests, such as Haemoccult SENSA, appear to have better sensitivity than Haemoccult II (Scholefield and Moss, 2002; Young, St John et al., 2002). One study that was eligible for the question under consideration in this review also evaluated the test performance of both HOII and HOS on the same population (Rozen et al., 1997). This study reported similar sensitivities of both tests for any colorectal neoplasia (Se = 27-31%) as well similar sensitivities for neoplasms larger than 1 cm and CRC combined (Se = 63%). Likewise, the specificities of both these guaiac tests, as compared in this particular study, were not significantly different. Specificity for any colorectal neoplasia was high (Sp = 93-95%) as well as for neoplasms larger than 1 cm and CRC combined (Sp = 92-95%). As discussed above, the positive and negative predictive values (reported in Table 10) for the tests compared in this study are of limited applicability to the purpose of this review.

One of the main drawbacks of HOS is a requirement for screenee adherence to strict dietary restrictions, to avoid false positive test results. Rozen et al. (2000) state that not developing the HOS test cards until three days after the last card is prepared allows for breakdown of dietary vegetable peroxidases, and reduces false positives, while making FOB testing easier for the patient. The results of the RCT by Cole et al. (2003) confirm that, for their study population, using the FxS test, which had similar stool sampling procedures but did not require the same dietary restrictions as HOS, significantly increased the likelihood of FOBT completion by screenees.

CFOBB test

Cheng et al. (2002) compared a test described as being a chemical test involving the same reaction (with peroxidase) as Haemoccult, to an immunochemical test called OC-Hemodia. Limited useful data could be extracted from this study report, and so although the results suggested superior immunochemical test performance, the supporting evidence for this association was weak.
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rozen et al. (1997)</td>
<td>Cross-sectional survey design</td>
<td>Grade III-2</td>
<td>Immunochemical Hemocell OBT (HOS)</td>
<td>N= 403</td>
<td>Mean age 60.2 (±10.8 s.d.) years, female = 54%.</td>
<td>Positive for ≥ 1 FOBT n=46.</td>
<td>Limitations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Immunochemical HemoSelect (HS)</td>
<td>Recruited from consecutive attendees at a CRC screening service and symptomatic patients coming for evaluation of abdominal complaints.</td>
<td>Asymptomatic = 97% (average risk= 22%, family history CRC = 45%, cured breast cancer = 5%, adenoma follow-up = 17%, cured CRC follow-up = 7%), Symptomatic = 3% (abdominal symptoms)</td>
<td>Any colorectal neoplasia detected</td>
<td>Sensitivity: FxS=23% (95% CI 7-39) HS= 35% (95% CI 16-57) HOII= 27% (95% CI 10.44) HOS= 31% (95% CI 13-49)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Guaiac Hemoccult II (HOII)</td>
<td>Excluded if:</td>
<td></td>
<td>Specificity: FxS= 97% (95% CI 95-99) HS= 99% (95% CI 97-100) HOII= 95% (95% CI 93-98) HOS= 93% (95% CI 90-95), FxS&gt; HO (p&lt; 0.05), PPV: FxS = 35% (95% CI 13-58) HS= 67% (95% CI 35-90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Guaiac HemoccultSENSA (HOS)</td>
<td>Included participants had:</td>
<td>Outcome verification by endoscopic examination. N=237 (58.8%) had after FOBT because of positive FOBT, or other screening protocol indication [n= 90 had FS, n= 147 had other indications for endoscopic examination</td>
<td>FxS = 29% (95% CI 11-47) HOII= 23% (95% CI 9-37), PPV: FxS= 95% (95% CI 93-98) HOII= 95% (95% CI 93-97) HOS= 95% (95% CI 93-97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All tests SmithKline Diagnostics, California, US</td>
<td>N=166 (41.2%) with negative FOBT and no other indications for endoscopy at time of FOBT testing had had a normal endoscopic examination before completing FOBTs [n=97 had FS, mean 2.9 (s.d. ± 1.5) years before, n=69 had colonoscopy, mean 2.3 (s.d. ± 1.5) years before]</td>
<td>Adenomas ≥ 1 cm or CRC detected Sensitivity: FxS= 63% % (95% CI 24-91) HOII= 86% (95% CI 42-100)</td>
<td></td>
<td>Specificity: FxS= 97% (95% CI 95-98) HS= 98% (95% CI 96-99) HOII=95 % (95% CI 93-97) HOII= 92% (95% CI 89-95), PPV: FxS = 29% (95% CI 10-56) HS= 50% (95% CI 21-79)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All participants received instructions about dietary and medication restrictions for 3 days prior to and during 3 days of sample collection. Compliance with instructions assessed by questionnaire. Participants took faecal samples from 2 parts of 3 consecutive stools, and smeared them on sample cards of FxS, HOII, and HOS, and returned completed cards to study investigators. The HS FOBT was performed in a laboratory on samples from the smeared FxS test cards. All tests completed by one technician blinded to clinical diagnosis. If any one test was positive, then FOBT testing was regarded as positive and full colonoscopy performed after tests. Study received ethical approval in 1996, but actual dates of testing not specified., N=166 (41.2%) with negative FOBT and no other indications for endoscopy at time of FOBT testing had had a normal endoscopic examination before completing FOBTs [n=97 had FS, mean 2.9 (s.d. ± 1.5) years before, n=69 had colonoscopy, mean 2.3 (s.d. ± 1.5) years before]</td>
<td>Adenomas ≥ 1 cm or CRC detected Sensitivity: FxS= 63% % (95% CI 24-91) HOII= 86% (95% CI 42-100)</td>
<td></td>
<td>Ranking these four tests, HS had the best overall Se, PPV and NPV for clinically significant neoplasia, followed by FxS, HOII and then HS. Because there were few neoplastic findings in this small-sized, mainly asymptomatic and repeatedly screened population, the results, especially relating to Se, should be regarded as a pilot, and evaluated in a larger screening study.</td>
<td></td>
</tr>
<tr>
<td>Raizen et al. (1997)</td>
<td>Cross-sectional survey design</td>
<td>Grade III-2</td>
<td>Immunochemical FlexSure OBT (FxS)</td>
<td>Recruited from consecutive attendees at a CRC screening service and symptomatic patients coming for evaluation of abdominal complaints.</td>
<td>Asymptomatic = 97% (average risk= 22%, family history CRC = 45%, cured breast cancer = 5%, adenoma follow-up = 17%, cured CRC follow-up = 7%), Symptomatic = 3% (abdominal symptoms)</td>
<td>Results not reproduced.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All tests SmithKline Diagnostics, California, US</td>
<td>Excluded if:</td>
<td></td>
<td>The outcomes of this study support the immunochemical test HS as more specific, for all neoplastic lesions as well as those that were, or had a greater chance of being, malignant. However, the applicability of the predictive values and the study findings in general is limited, given the diverse study population.</td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 10.** Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening.
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rozen et al., (2000) Israel</td>
<td>Cross-sectional survey design Grade III-2</td>
<td>Immunochemical FlexSure OBT (FxS) Guaiac HemoccultSENSA (HOS) Both tests SmithKline Diagnostics, California, USA. Certain medications were withdrawn 1 week before test card preparation. All participants initially received instructions about dietary restrictions for 3 days prior to and during the days of sample collection. After first 403 participants, no dietary limitations were used. Participants took faecal samples from 2 parts of 3 consecutive stools, smeared them on sample cards of FxS and HOS, and returned the completed cards to study investigators. Development took place &gt;3days ≤14 days of last test card completion, to allow for breakdown of unstable dietary peroxidases. All tests completed by one technician blinded to clinical diagnosis. If either test was positive, then FOBT testing was regarded as positive and it was recommended that total colonoscopy be done after FOBTs. Study received ethical approval in 1996, but actual dates of testing not specified.</td>
<td>N=1410 Mean age 60.9 (±11 s.d.) years, female = 53%. Recruited from consecutive attendees at a CRC screening service Asymptomatic n=1367 (97%). This group consisted of participants who: were of average risk (21%); had family history of CRC (47%); had a cured breast cancer (6%), were undergoing colorectal adenoma follow-up (19%); had a history of CRC (47%). Symptomatic n = 43 (3%) All came for evaluation of abdominal symptoms. Included participants had: • to agree to and be capable of keeping to medication limitations • to be able to prepare the FOBTs • to have a recent history of a normal endoscopic examination of large bowel, if not, it was performed after FOBT testing Excluded if: • non-co-operation • active rectal bleeding • known ulcerative colitis.</td>
<td>For each FOBT, the outcomes of Se, Sp, PPV, NPV were analysed for: • subjects with any colorectal neoplasia detected • subjects having adenomas ≥1 cm or CRC detected • subjects having only adenomas detected (Specificity was estimated from results of total colonoscopy or a negative FS in the absence of any positive FOBTs if subject had not had colonoscopy). Outcome verification by endoscopic examination. N=681 (46%) had this after FOBT because of positive FOBT, or other screening programme. Adenomas ≥1 cm or CRC detected Sensitivity: FxS=35% (95% CI 14-56) HOS= 50% (95% CI 28-72), ns. Adenomas ≥ 1 cm or CRC detected Sensitivity: FxS= 9% (95% CI 9-19) HOS= 12% (95% CI 7-26), ns. All endoscopic lesions were measured, removed and sent for examination by one pathologist. Positive endoscopic finding of &quot;neoplasia&quot; = presence of an adenomatous polyp (any size) or CRC. Subjects classified according to the largest adenoma detected or presence of CRC.</td>
<td>Either FOBT positive n=89 (HOS only positive n=64, FxS only positive n=14, both FOBTs positive n=11) Any colorectal neoplasia detected Sensitivity: FxS=18% (95% CI 8-28) HOS= 35% (95% CI 22-47), Difference in proportions p&lt; 0.05 Specificity: FxS= 99% (95% CI 98-99.5) HOS= 96% (95% CI 95-97), p&lt; 0.05. PPV: FxS = 42% (95% CI 22-61) HOS = 25% (95% CI 15-35). NPV: FxS = 95% (95% CI 93-97) HOS= 97% (95% CI 96-98), not significant. Both FOBTs positive: Sensitivity: FxS= 9% (95% CI 9-19) HOS= 11% (95% CI 7-16), all p&lt;0.05, NPV= 77% (95% CI 96-98), ns. Adenomas ≥ 1 cm or CRC detected Sensitivity: FxS= 35% (95% CI 14-56) HOS= 50% (95% CI 28-72), ns. Specificity: FxS= 99% (95% CI 98-99) HOS= 95% (95% CI 94-96), p&lt;0.05. PPV: FxS = 29% (95% CI 11-47) HOS= 13% (95% CI 6-21). NPV: FxS = 99% (95% CI 99-100) HOS= 99% (95% CI 99-100). Both FOBTs positive: Sensitivity: FxS= 25% (95% CI 6-44), for HOS p&lt;0.05, Sp= 100% (95% CI 99-100) for all p&lt;0.05, PPV= 46% (95% CI 43-48) for all p&lt;0.05, NPV= 99% (95% CI 98-100), ns.</td>
<td>Limitations Some industry support Possible disease progression bias, due to different timings of reference standard tests among study population. Possible review bias as not clear whether pathologist blind to FOBT and clinical history. Possible verification bias, as only those with positive FOBT had whole bowel examination by colonoscopy. Some of those with negative FOBT only had distal bowel examined (by FS). Authors' conclusions Within an endemic screening programme, we have shown that the sensitive guaiac HS test, without dietary restrictions, had a better sensitivity for all colorectal neoplasias than the specific immunochemical F5 test. The specificity of HS although lower than F5, was still acceptable for our population screening programme. Reviewers' conclusions The outcomes of this study support the guaiac HS test as significantly more sensitive than F5x for all neoplastic lesions. However, this is not the case for lesions that are malignant or have a greater chance of being malignant. The applicability of the predictive values and the study findings in general is limited, given the diverse study population, and variable past screening histories of participants.</td>
</tr>
</tbody>
</table>
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source/Autor (Year)</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Saito et al. (2000) | Case-control | Grade III-2      | Immunochemical HemSp (HmS), Fujirebio Inc., Tokyo, Japan, (hemaggulitination) Guaiac Hemoccult II (HII), Smithkline Diagnostics, Sunnyvale, USA. 2-step method used for HII, testing. For step I subjects had no diet restrictions and took faecal samples from 2 parts of 3 consecutive stools, smeared them on sample cards and returned completed cards to study investigators. If this was positive, the process was repeated with diet restrictions. Only those with second test positive were classified as positive. The HmS FOBT was performed in laboratory on a one-day sample smeared onto filter paper cards. Subjects with either the HII or HmS test positive recommended to have a work-up examination by FS + DCBE and/or total colonoscopy. Recommendation made up to 3 times for those who refused initially. FOB testing undertaken as part of a CRC screening programme offered annually to all Kizukuri town inhabitants aged 40 years. Guaiac testing from 1980-1985, immunochemical testing from 1986-1993. Sample selected from all inhabitants of Kizukuri town (population 23,000) aged ≥40 years. Identified from list of residents in study area. Cases n=51 Those who died of CRC during 1982-1993, who had been diagnosed as having CRC, aged over 39 years and who had been living in the screening programme area after 1980. Excluded those for whom histological or other substantial evidence of CRC diagnosis was not available. Controls n=152 Three controls per case were randomly selected from list of residents in 1980, matched by sex and year of birth to cases. Required to have been alive at the time the case was diagnosed, and free of CRC before 1980 (when the programme began). Specific demographic data for cases and controls not reported. Participation rates for each test Odds ratio (OR) of dying from CRC: • if participated in FOBT screening programme • if participated in FOBT screening programme and screened with HII • if participated in FOBT screening programme and screened with HmS Screening history verification Screening histories of cases and controls were investigated by reviewing the records of the screenee. Screening histories of cases and controls compared during the same year-segment in the 4 years prior to case diagnosis. The FOBT that led to the diagnosis of a case was included in a screening history (but not whether diagnostic work-ups were completed after positive FOBT). Outcome verification Cases identified by death certificates, Information on CRC diagnoses (e.g., histological type, anatomical location) verified by reviewing hospital medical records. If histological information not available subjects were excluded unless there was other substantial information 1980-1985: HII test • offered to 54,350 persons • accepted by 21,862 (40%) • positive tests = 1,291 (5.7%) • follow-up examinations = 799 (64%) • cancers detected = 19 1985-1993: HmS test • offered to 98,727 persons • accepted by 41,994 (43%) • positive tests = 1,139 (2.7%) • follow-up examinations = 526 (46%) • cancers detected = 26 Compliance rate for diagnostic work-up 46% vs 64% (p<0.001, X² test) OR of dying from CRC if participated in FOBT programme (by years since most recent screening): 0-1 years OR = 0.20 (95% CI 0.08-0.49) 1-2 years OR = 0.17 (95% CI 0.04-0.75) 2-3 years OR = 0.77 (95% CI 0.13-4.56) 3-4 years OR = 1.14 (95% CI 0.10-12.66) OR of dying from CRC if participated in FOBT programme (by years since most recent screening with HII test): 0-1 years OR = 0.36 (95% CI 0.11-1.17) 1-2 years OR = 0.88 (95% CI 0.08-10.26) 3-4 years OR = 1.0 (1 case, 1 control) OR of dying from CRC if participated in FOBT programme (by years since most recent screening with HmS test, adjusted for previous screening history with HII): 0-1 years OR = 0.19 (95% CI 0.05-0.70) 1-2 years OR = 0.56 (95% CI 0.06-5.54) 2-3 years OR = 0.63 (95% CI 0.04-9.73) 3-4 years OR = 2.89 (95% CI 0.15-53.91) Limitations: No power calculations reported Possible self-selection bias as those who chose to undergo CRC screening may systematically differ from those who did not participate. Non-simultaneous testing with the different FOBTs. Low rates of follow-up examination after positive test for both FOBTs. Significantly higher rate of follow-up examinations in HII group favours reduced mortality from HII screening. Authors’ conclusions: Results suggest a reduced risk of dying of CRC after being screened with either the HmS and HII tests sequentially, or the HmS test alone. In combination with the result of a previous case-control study, the efficacy of the screening programme using the immunochemical test has been strongly suggested. However, influence of self-selection bias must be considered when interpreting the results. Reviewers’ conclusions: The ORs of dying from CRC were lower for screening with the HmS test than with HII, but only reached statistical significance for those screened with HmS during the preceding one year. Small numbers of cases/controls for FOBT screening may have accounted for the wide confidence intervals seen as time since most recent screening increased. Effectiveness And Cost-effectiveness Of Population Screening For Colorectal Cancer
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zappa et al. (2001)</td>
<td>Cohort (not specified, but appears to have been conducted retrospectively)</td>
<td>*Immunochemical&lt;br&gt;*HemSp (HmS) - also called 'Immudia', Fujirebio Inc., Tokyo, Japan. Or&lt;br&gt;*HemeSelect (HS), SmithKline Diagnostics, Palo Alto, USA. Guaiac&lt;br&gt;Hemoccult II (HOII), SmithKline Diagnostics, Palo Alto, USA.&lt;br&gt;Method of testing and definition of positive FOBT for above tests not reported.&lt;br&gt;Subjects with a positive FOBT were offered diagnostic work-up examination by colonoscopy (colonoscopy + DCBE if colonoscopy not possible to complete)&lt;br&gt;FOBT testing undertaken as part of a population-based CRC screening programme offered biennially (actual interval 2.5 years) to residents aged 40-70 years in 24 municipalities of Florence from January 1992-December 1997 (from 1995 onwards invitation restricted to those age 50-70 years). From 1992-1996 some municipalities used guaiac test and others used immunochemical testing. From 1997 onwards only immunochemical testing used.&lt;br&gt;*Both of these are reverse passive hemagglutination tests (RPHA)</td>
<td>N= 41,774 subjects aged 50-70 years underwent 54,308 FOBTs after a screening invitation during period considered in this study. Male n= 24,957 Female n= 29,351 Excluded&lt;br&gt;• Age &lt; 50 years.&lt;br&gt;• 8,008 subjects who had participated in a previous FOBT study at beginning of this study period. No difference evident between the guaiac and immunochemical tests in terms of age or sex distribution of subjects.</td>
<td>FOBT programme sensitivity&lt;br&gt;CALculated using proportional incidence method (compares incidence of interval cancers to underlying incidence of cancers expected in the absence of screening to give the proportion of expected cancers in which diagnostic anticipation was obtained by screening).&lt;br&gt;Relative Risk (RR) of developing an interval cancer 2 years after a negative FOBT&lt;br&gt;Outcome verification&lt;br&gt;Incident cases of CRC for district of Florence were provided by regional Cancer Registry. Data from Cancer Registry also used to calculate CRC incidence expected in absence of screening. Cancers classified according to Duke’s classification (only Duke’s A considered ‘early’ cancer for this study). All CRCs diagnosed within 2 years after a negative screen were classified as interval cancers (as were 3 cancers that occurred in FOBT positive screeners who refused further work-up).&lt;br&gt;Screening history verification&lt;br&gt;Screening histories of incident cases of CRC obtained through linkage with a computerised database to obtain CRC screening histories of matched subjects. Linked records further checked manually (to eliminate mismatch).</td>
<td>Proportion of tests performed as first screening test in programme significantly higher for HOII than RPHA (50% versus 44%, p&lt; 0.01)&lt;br&gt;HmS and HS tests (RPHA)&lt;br&gt;• 20,893 tests done (16,670 subjects)&lt;br&gt;• 35,886 person-years observation after 2 years following RPHA&lt;br&gt;• overall positive tests = 839 (4%)&lt;br&gt;• screen-detected cancers = 73&lt;br&gt;• overall detection rate (DR) for CRC = 3.5% (95% CI 2.8-4.4)&lt;br&gt;• interval cancers within 2 years = 9&lt;br&gt;• expected CRC = 51.2&lt;br&gt;HOII test&lt;br&gt;• 33,415 tests done (25,104 subjects)&lt;br&gt;• 65,723 person-years observation after 2 years following HOII&lt;br&gt;• overall positive tests = 1,561 (4.5%)&lt;br&gt;• screen-detected cancers = 66&lt;br&gt;• overall DR for CRC = 2.0% (95% CI 1.6-2.4)&lt;br&gt;• interval cancers within 2 years = 47&lt;br&gt;• expected CRC = 93.5&lt;br&gt;DR for CRC at first screen for RPHA= 4.5% (95% CI 3.2-6.1) versus HOII = 2.7% (95% CI 2.0-3.7), p= 0.05&lt;br&gt;DR for CRC at subsequent screens for RPHA = 2.7% (95% CI 1.9-3.8) versus HOII = 1.2% (95% CI 0.7-1.8), p&lt; 0.01&lt;br&gt;Overall FOBT programme sensitivity:&lt;br&gt;RPHA = 82% (95% CI 67-92) vs. HOII = 50% (95% CI 34-63), p&lt; 0.01&lt;br&gt;RR of developing interval cancer after negative screen with HOII compared to RPHA:&lt;br&gt;RR = 2.64 (95% CI 1.3-6.7)</td>
<td>Limitations&lt;br&gt;Not clear where FOBTs from different regions were analysed or what constituted a positive result. This introduces possibility for work-up bias between tests, given the different testing strategies employed across the municipalities studied.&lt;br&gt;No concurrent testing with the different FOBTs.&lt;br&gt;No data provided to enable assessment of performance of RPHA tests individually.&lt;br&gt;Authors’ conclusions&lt;br&gt;Our study confirms that RPHA is more sensitive compared with the guaiac test. Overall, the risk of developing an interval cancer after a negative guaiac test is almost 3 times higher compared with RPHA.&lt;br&gt;Reviewers’ conclusions&lt;br&gt;The results of this study provide highly significant data that the sensitivity of a screening programme using the RPHA test is superior to that using HOII, in a population setting. However, there are potential sources of bias in the conduct of this study that cannot be excluded from the information reported.</td>
</tr>
</tbody>
</table>
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheng et al. (2002) Taiwan</td>
<td>Cross-sectional Grade III-2</td>
<td>Immunochemical OC-Hemodia (OCH), Eiken Chemical Co., Tokyo, Japan. Guaiac CFOBB test, Shih-Yung Instruments Co., Taipei, Taiwan. This test involves a chemical reaction of toluidine with peroxidase. Subjects asked to follow diet instructions for 3 days prior to stool collection. Stool specimens were sent for FOBTs on the day of the colonoscopic examination. FOB testing undertaken as part of a voluntary health-screening programme at a Cancer Centre in Taipei, Taiwan. Study dates January 1997-December 2000.</td>
<td>Selected from source population of n = 7617 asymptomatic adults who underwent colonoscopic examinations in the health-screening programme during study period. Study sample n = 7411 examinations met criteria for inclusion in analysis. Mean age = 46.8 ± 9.9 years. Age range ≤ 20 to ≥ 80 years 69% of sample between 31-50 years of age. Male:female ratio = 1.23:1 n = 206 (3%) excluded for reasons including: • incomplete colon examination • presence of symptoms (e.g. bloody stool) • history of CRC, colitis, inflammatory bowel disease • previous positive FOBT • colorectal lesion found at colonoscopy of undefined site or size. Subjects with a personal history of colonic polyps or family history of CRC were not excluded.</td>
<td>Percentages of positive FOBTs found for different colonoscopy findings: Se, Sp, PPV and NPV for both FOBTs could be calculated from study data (if colonoscopy result is used as a reference standard). Complications of colonoscopy Outcome verification Screening colonoscopies all performed by trained specialists, under conscious sedation. Reports recorded on standardised form that included size and location of any lesions detected. Examinations classified on basis of most advanced lesion identified. Lesion categories were: • polyps • advanced neoplasm (defined as a polyp larger than 1cm, polyps with severe villous or dysplastic features • cancer.</td>
<td>From 7411 colonoscopic examinations, 719 (9.7%) abnormal findings. Abnormal findings included 93 (1.3%) advanced neoplasms and 16 (0.2%) cancers. OCH test Positive tests in screenees with colonoscopy findings of: • normal = 8.4% • polyps = 16.8% • advanced neoplasm = 48.3% • cancer = 87.3% Significant trend for increased positivity with more serious lesions, p&lt; 0.0001 For polyps: Se= 17% (95% CI 14-20) Sp= 92% (95% CI 91-92) PPV= 18% (95% CI 15-21) NPV= 91 (95% CI 90-92) For advanced neoplasms: Se= 48% (95% CI 38-59) Sp= 92% (95% CI 91-92) PPV= 8% (95% CI 6-10) NPV = 99 (95% CI 99-99) For cancers: Se= 88% (95% CI 62-98) Sp= 92% (95% CI 91-92) PPV= 3 (95% CI 1-5) NPV= 100 (95% CI 100-100) CFOBB test Positive tests in screenees with colonoscopy findings of: • normal = 11.2% • polyps = 12.8% • advanced neoplasm = 20.2% • cancer = 37.6% Significant trend for increased positivity with more serious lesions, p&lt; 0.001 For polyps: Se= 13% (95% CI 10-15) Sp=89% (95% CI 88-90) PPV= 11% (95% CI 9-13) NPV=90% (95% CI 90-91) For advanced neoplasms: Se= 20% (95% CI 12-29) Sp= 89% (95% CI 88-90) PPV= 3% (95% CI 1-4) NPV = 99 % (95% CI 98-99) For cancers: Se=38% (95% CI 14-61) Sp=89 % (95% CI 88-90) PPV=1% (95% CI 0.1-2) NPV=99% (95% CI 99-100) Complications from colonoscopy: Perforation n = 2, Bleeding n = 5, Mortality n = 0</td>
<td>Limitations Sample included those at higher risk for CRC. Limited details given regarding FOBT testing methods used. Possible review bias, for both FOBTs and colonoscopy findings, as not clear whether FOBT reporting was blind to patient colonoscopy result, nor if colonoscopists were blind to FOBT results. Authors’ conclusions For both FOBTs, the FOBT-positivity rate increased with pathologic severity, and this trend was statistically significant. This study confirmed that immunochemical rather than chemical (guaiac) FOBT should be used for CRC screening. Reviewers’ conclusions The main aim of this study was to compare the performance characteristics of the FOBTs with FS and colonoscopy. Using the limited useful data that could be extracted from this paper, the results of this study also suggest that the OCH test had superior performance over the CFOBB test for both Se and Sp, especially for more serious neoplastic lesions.</td>
</tr>
<tr>
<td>Source Country</td>
<td>Study design</td>
<td>Evidence Grading</td>
<td>Comparison interventions and dates of testing</td>
<td>Sample</td>
<td>Outcomes and verification</td>
<td>Results</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>------------------</td>
<td>-------------------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Cole et al. (2003)</td>
<td>RCT</td>
<td>Grade II</td>
<td>Immunochemical InSure (InS) Enterix Inc., Portland ME, US</td>
<td>Immunochemical FlexSure OBT (FxS) Beckman Coulter Inc., Palo Alto, US, Gudaz, HemoccultSENSA (HOS), Beckman Coulter Inc., Palo Alto, US.</td>
<td>Participation rates for an FOBT that does not require dietary restrictions (FxS) compared to one that does (HOS).</td>
<td>Participation rates at 12 weeks: InS test 240/606 (39.6%), FxS test 185/606 (30.5%). Participation rates for an FOBT with simplified sampling (InS) compared to one requiring traditional spatula sampling of faecal material (FxS). Participation for InS compared to FxS: 39.6% compared to 30.5%, ( \chi^2 = 10.6, p = 0.002 ). Participation for InS compared to HOS: 39.6% compared to 30.5%, ( \chi^2 = 7.39, p = 0.007 ). Participation for FxS compared to HOS: 30.5% compared to 23.4%, ( \chi^2 = 3.71, p = 0.055 ).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Identical invitations to screening, with relevant test instructions, brief questionnaire and consent request sent to all invitees. Reminders at 6 weeks.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For InS, invitees requested to use manufacturer’s brush to take samples from toilet water of 2 consecutive stools and smear them on sample cards.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For FxS, invitees requested to use manufacturer’s spatula to take samples from 1 part of 3 consecutive stools and smear them on sample cards. For HOS, instructions requested diet and medication restrictions and use of manufacturer’s spatula to take samples from 2 parts of 3 consecutive stools and to smear on sample cards.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All tests developed by 2 investigators, according to manufacturer’s instructions. Screening offered from April-August 2001.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Effectiveness and Cost-effectiveness of Population Screening for Colorectal Cancer**
Table 10. Primary research studies appraised investigating the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ko et al. [2003] US</td>
<td>Non-randomised clinical trial Grade III-2</td>
<td>Immunochromual FlexSure OBT (FxS) SmithKline Diagnostics, Palo Alto, US, Guaiac HemoccultSENSA (HOS), Beckman Coulter Inc., Palo Alto, US.</td>
<td>Sample n= 5929 Male = 5810 (98%) Mean age (± sd) = 65.4 (±10.5) years FxS group mean age = 64.8 (±10.3) years HOS group mean age = 65.8 (±10.6) years Sample selected from attendees of a general medical clinic in Virginia, clinic operated as 2 firms, with separate patient lists. Patient characteristics of age (mean 62.7 years), race, income and self-reported cardio-respiratory and diabetic status were similar between the 2 firms.</td>
<td>Participation rates for each test Positive predictive value for • any adenoma or malignancy • adenoma &gt;1cm or malignancy • malignancy</td>
<td>FxS test: • ordered for n= 2965 attendees • 1410 (48%) of tests ordered were returned by attendees. HOS test: • ordered for n= 2964 attendees • 1396 (47%) of tests ordered were returned by attendees</td>
<td>Limitsations Sample characteristics not well-reported. Unclear if cluster sampling approach taken into account in analysis. This may have led to reduced precision of effect estimates between study groups. Possible selection bias due to systematic differences between patient groups and FOBT referral practices of clinical staff at the two firms. Likely verification bias. Authors' conclusions We found few differences between the guaiac-based and immunochemical-based tests in terms of completion rate, positivity rate or PPVs for adenoma or polyps CRC. Immunochromual tests were associated with higher rates of inappropriate specimens for processing. This suggests that the theoretical benefits of immunochromual tests are not borne out in practice. Reviewers' conclusions Although similar participation rates for FOBTs suggests equal acceptability, the likely sources of bias and under-reporting of methods compromise the validity of the study results. Furthermore, the general applicability of the study findings is indeterminate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For both tests, participants were given 3 test cards and test collection paper and spatulas. Asked to take faecal samples from 2 parts of 3 consecutive stools, smear them on sample cards and return completed cards to study laboratory. Tests processed in laboratory according to manufacturer's instructions. Subject's primary care practitioner notified of test results, and was responsible for referring subject for further diagnostic evaluation (such as colonoscopy or DCBE) or for repeat FOBT if test inadequate). FOBT testing undertaken in line with clinic guidelines that recommended annual CRC screening with FOBT. Information for study collected on FOBTs ordered by clinic staff between 1st August 2000 and 30th September 2001.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average return time: FxS = 26 ± 32 days, HOS = 30 ± 38 days, p= 0.001 Tests cancelled by laboratory: FxS = 136 (10%), HOS = 77 (5%), p = 0.001 Positive tests: FxS = 128 (9%), HOS = 122 (9%), p= 0.72 Referred for total colonic examination: FxS = 87 (68%), HOS = 78 (64%), p= 0.50 Completed total colonic examination: FxS = 69 (54%), HOS = 64 (52%), p= 0.73 Positive predictive value (PPV) Any adenoma or malignancy: FxS = 58%, HOS = 59%, p= 0.87 Adenoma &gt;1cm or malignancy: FxS = 17%, HOS = 30%, p= 0.09 Malignancy: FxS = 7%, HOS = 14%, p= 0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SECONDARY RESEARCH

The search strategy identified five relevant reviews that considered the accuracy of immunochemical FOBT compared to guaiac FOBT screening for colorectal cancer. The methods and conclusions are described in Table 11. As discussed in the Methodology chapter above, these papers may not have employed the same inclusion and exclusion criteria as applied in this review and the results must be interpreted with care.

The report by AHTAC (1997) was considered by the Working Party on Screening for Colorectal Cancer (1998). Briefly, this report, conducted for the Minister of Health and Family Services, aimed to assess the evidence on the benefits, risks and costs of CRC screening for asymptomatic people (as well as surveillance for those at higher-risk of CRC). For FOBT, the committee considered the evidence up to 1996, and included a review of the performance characteristics of immunochemical and guaiac FOBTs. However, the immunochemical FOBT accuracy studies reported in this review included diverse study populations, and therefore limit the generalisability of findings to screening populations; the authors themselves advised caution when interpreting comparisons in test performance, given differences in study design and conduct. AHTAC’s (1997) overall recommendation as a result of the review was that, subject to favourable preliminary testing, Australia should develop a programme for the introduction of population screening for CRC using FOBT. However, a caveat to this recommendation was that uncertainties surrounded which FOBT method is the most appropriate, and that pilot and feasibility studies should investigate which method provides the best balance between sensitivity, specificity and costs.

The aim of an Evidence Report/Technology Assessment prepared by the Conseil d’Évaluation des Technologies de la Santé du Quebec (CETS) (2000) was to examine in detail evidence for the efficacy and efficiency of different strategies for CRC screening in average-risk, asymptomatic people in terms of mortality reduction. However, they also provided a small section discussing the performance characteristics of different methods of FOBT. CETS (2000) concluded this report with the belief that a CRC screening programme, using biennial unrehydrated guaiac FOBT, would significantly reduce mortality from this type of cancer. However the authors stated that such a programme would be based on the existing evidence, and that in light of ongoing research it would be advisable to explore other screening strategies. The authors then listed pre-conditions that would apply to the implementation of such a programme, including the requirement for feasibility studies to determine the mechanisms and costs of CRC screening, and validate the performance of FOBTs considered.

The review conducted by Craven (2001) focussed on FOBT for asymptomatic CRC population-screening and included an evaluation of evidence regarding the performance characteristics of immunochemical and guaiac tests. Although this review identified several controlled studies of robust design, all of which had been published in or prior to 1996, it was unable to make a conclusive decision regarding which type of test had superior performance. A suggestion was made by the reviewer, based on comments from some of the studies reviewed, that a dual-test approach using guaiac followed by immunochemical tests may maximise both sensitivity and specificity (Craven, 2001).

Young et al.’s (2002) review for the World Health Organization and World Organization for Digestive Endoscopy aimed to evaluate the published evidence concerning the various FOBT methods available, with the goal of providing guidance on which FOBT to use in a given population setting. This report provided an informative introductory section outlining the practical advantages and disadvantages associated with each of guaiac and immunochemical FOBTs, as well as a short appendix listing the names and manufacturers of some of the main tests for faecal haemoglobin that were available at the time of their report. These reviewers did not recommend the rehydration of the Hemoccult test, and considered the newer and more sensitive HemoccultSENSA test to be the guaiac test of choice, although only for populations that complied with the required dietary restrictions. HemeSelect (also previously called Immudia HemSp) was the immunochemical FOBT that Young et al. (2002) found to compare most favourably with the guaiac tests, as well as being the most widely studied, and was considered by this review to be the most suitable immunochemical test (ICT) for CRC screening. The study selection requirements employed by this review aimed to include only studies of a certain methodological level. Given this prerequisite, the main conclusion reached that there was no extensively studied FOBT that fulfilled the needs for all target populations worldwide was likely to be
based on valid and reliable evidence. The authors finally recommended that the choice of FOBT for screening in any given population setting should take into consideration other health system and population factors.

The review, conducted by Piper (2004) for the Blue Cross Blue Shield Association Technology Evaluation Center, ascertain to evaluate whether there was sufficient evidence to evaluate the performance of ICTs in general, or of specific ICTs, and to compare performance to standard guaiac FOBTs. It also examined the evidence on patient compliance with various FOBT formats to determine if compliance was more likely with any or with a specific ICT versus guaiac FOBTs. This review addressed clearly focussed, well-defined questions. However, most of the eligible studies that assessed the performance of immunochemical FOBTs compared to guaiac tests were conducted with tests that were no longer available in the US at the time of the review. Given this fact, the Technology Evaluation Center finally concluded was that there was insufficient evidence to permit conclusions regarding the use of immunochemical FOBTs for CRC screening and health outcomes (Piper, 2004).

CONCLUSIONS

Several authors (Young et al., 2002; Fletcher, 1996, in Allison 2003) have commented that, since there is direct evidence from RCTs that screening with haemoccult FOBT reduces mortality from CRC, then evidence of comparable or superior test performance of ICTs to haemoccult may be taken as proof of their effectiveness in reducing mortality. The opinion of Young (1996, in Working Party on Screening for Colorectal Cancer, 1998, Chapter 6) considered previously by the working party was that only HemeSelect (previously called Immudia-HemSp) met criteria of improved sensitivity and acceptable specificity compared with Haemoccult/Haemoccult II.

The studies identified by this review evaluated a diverse range of both types of test. Further design and quality limitations, as discussed above, limited the conclusions that could be drawn from the published literature about which test offers the optimal performance. The limited evidence available since the working party considered FOBT screening regarding immunochemical test performance compared to guaiac, also suggested that only HemeSelect performs as well as, or better than the guaiac tests HOII and HOS when compared head-to-head in cross-sectional studies. There may be other immunochemical tests available that are more sensitive than those evaluated in the papers considered eligible for inclusion in this review, but their applicability in screening populations would need to be confirmed.

There was good evidence that the simplified testing process and sampling kit of the immunochemical test Insure encouraged a greater proportion of people from the general population to participate when invited to complete FOB screening tests.

The conclusions of this chapter are similar to those made in other recent reviews. There is limited definitive evidence regarding superior immunochemical FOBT screening performance over the guaiac tests.

HemeSelect is the immunochemical test that compares most favourably with the guaiac tests. It has now progressed into the commercial tests Magstream HemSp and Bayer Detect™ (Young, 2004). However, Insure (a brush-sampling immunochemical FOBT) may be more acceptable. The conclusions on this topic should be revisited if further reliable evidence on the comparative performance of screening FOBTs becomes available.
Table 11. Secondary research appraised relevant to accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Health Technology Advisory Committee (AHTAC) [1997]</td>
<td>Search: 1990-1996. Supplemented by identification of key references for earlier articles. Databases searched: Medline, EMBASE, CancerLit. Also identified articles from the Internet and reference lists of retrieved articles. Current Contents searches were continued up to the date of publication (1997). Accessed &quot;grey&quot; literature by contacting health departments, academic institutions and reviewing submissions to the AHTAC Working Group. Key Words: colorectal cancer, screening, faecal occult blood test, barium enema, economics, costs, epidemiology, prevention, survival, acceptability, compliance.</td>
<td>All study types evaluating FOBT screening were considered, but greater weight was given to research material scoring highly on NHMRC (1995) guidelines. For studies examining test performance characteristics, those that offered a reasonable prospect of providing valid measures, as determined by two independent reviewers, were included. All identified studies examining issues relating to compliance in which study groups numbered at least 100 were included.</td>
<td>Narrative summaries of 9 studies evaluating performance characteristics of immunochemical FOBTs are presented. 6 studies evaluated the performance characteristics of immunochemical tests alone, and 3 included comparisons of the performance of ICTs to Hemoccult or HemoccultSENSA tests. Narrative summaries of 7 studies evaluating performance characteristics of guaiac FOBTs are presented. All evaluated the performance characteristics of guaiac tests alone, with at least 3 studies of long-term screening RCTs. For immunochemical FOBTs: HemeSelect PPV = 7.5% Se (range) = 70-95% Sp (range) = 88-96% Detectacol (Fecal Human Hemoglobin Test) PPV = not reported Se = 83% Sp = 96% Unspecified brands PPV = 1.8% Se = 25-87% Sp = 97% For guaiac FOBTs: Hemoccult/Hemoccult II PPV = 5-10%(CRC), 35-50% (large adenoma) Se = 50-80% Sp (range) = 95-98% HemoccultSENSA PPV = not specified Se = 77% Sp = not specified</td>
<td>Review did not specify clear PICO questions. The populations of the studies evaluating the performance of ICTs discussed in this review were diverse and included populations with known or symptomatic CRC, high-risk populations and other asymptomatic individuals. Authors’ conclusions The FOBT is the only test for which there is level I evidence [based on RCTs] demonstrating a reduction in mortality, using the test Hemoccult II. It is important not to assume that all other types of FOBTs will confer the same or greater benefit. It can be assumed, however, that if mortality benefits are demonstrated for Hemoccult, which is of relatively low sensitivity, then the benefits would be greater using tests of greater Se and Sp. Many other FOBTs are available, but their reliability needs further investigation. (Also see Chapter 3 for FOBT screening versus no screening)</td>
</tr>
</tbody>
</table>
### Table 11. Secondary research appraised relevant to the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conseil d’Évaluation des Technologies de la Santé du Québec (2000) Canada.</td>
<td>Search: 1966 – 1997. Databases searched: Medline, EMBASE, PASCAL, CancerLit. Key Words: Initially limited to screening and colorectal neoplasia for years 1966-1997. Colonic neoplasms then included for 1946-1987. From 1988–'present' (likely to be late 1999 or early 2000) query modified to include the terms colorectal or colonic or rectal combined with adenoma(s), adenocarcinoma(s), carcinoma(s), polyp(s), neoplasm(s) or cancer(s).</td>
<td>Review focused on asymptomatic people at average risk for CRC. The included studies were chosen on the basis of the level of evidence they provided. Priority given to RCTs, and other research designs were only mentioned if there was no RCT evidence, or as supplementary information.</td>
<td>For studies evaluating comparative accuracy of immunochemical and guaiac FOBTs, number not specified. However, some comparative performance characteristic analyses results are presented in narrative form within the section on FOBTs for screening. Immunochemical tests: HemeSelect PPV = 21% (95% CI 17-25) Se (for carcinomas and adenomas) = 97% Guaiac tests: Hemoccult/Hemoccult II PPV = 23% (95% CI 18-30) Se (for carcinomas and adenomas) = 76% HemoccultSENSA PPV = 9.2% (95% CI 7.6-11.2) Se (for carcinomas and adenomas) = 83% For low or average-risk individuals, found the published values for guaiac tests to be Se = 16-98%, and Sp=86-98%. Combined tests: HemoccultSENSA/HemeSelect combined PPV = 31% (95% CI 25-37)</td>
<td>Few details reported about studies evaluating performance of ICTs and guaiac FOBTs. Lack of specificity data for immunochemical tests. Authors’ conclusions: Based on the existing evidence, a CRC screening programme based on biennial guaiac FOBT screening would reduce mortality from CRC. In light of ongoing research it would be advisable to explore other screening strategies. Feasibility studies and pilot trials should be conducted, and one of the purposes of such work would be to validate the performance of FOBTs. (Also see Chapter 3 for FOBT screening versus no screening)</td>
</tr>
</tbody>
</table>
Table 11. Secondary research appraised relevant to the accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
- written in English  
- written/published between January 1980-January 2000  
- directly related to occult blood screening of asymptomatic population  
- primary research or a review.  
Excluded articles:  
- relating solely to other methods of screening (e.g., F5, colonoscopy)  
- relating to other aspects of establishing a national screening programme (e.g., economic analysis)  
- consisting of ‘grey literature’  
- involved multiple or serial publications of the same data  
Methodological quality of selected articles was assessed using a specific appraisal format and hierarchy of evidence to grade articles according to research design. | 6 studies evaluating which type of FOBT, immunochemical or guaiac, is most effective in terms of sensitivity and specificity for CRC asymptomatic population screening.  
4 controlled studies used a repeated measures design to directly compare both types of test.  
2 studies explored guaiac test performance following manipulation of the standard test procedure.  
Found that even with the more rigorous studies, problems of inadequate follow-up, imperfect sampling and poor compliance (which averaged 43%) existed.  
Evidence for FOB test performance varies considerably, and this may be due to several methodological factors:  
- the different types of FOBT used  
- whether or not the guaiac slides are rehydrated  
- the number of FOB samples taken from each subject  
- profiles of the subjects tested  
- compliance with testing. | Methodologically sound review process.  
Authors’ conclusions  
The evidence drawn from this review is that no single currently available FOBT provides optimal screening performance. Despite the apparent superiorly of the immunological test over the guaiac test, in terms of absolute cancer yield, the fruitless investigation of so many false positive results would generate enormous follow-on costs. Conversely, guaiac FOBT insensitivity also casts doubt on its value in population screening.  
(Also see Chapter 3 for FOBT screening versus no screening) |
Table 11. Secondary research appraised relevant to accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young et al. (2002)</td>
<td>Search: 1990 – 30th September 2001 (from 1993 for CancerLit). Databases searched: Medline, CancerLit. Limited to articles with English language abstracts. Key Words: colorectal cancer, colon cancer, bowel cancer, colorectal neoplasia, adenomas, adenomatous polyps, colonic polyps, FOBT, fecal/faecal occult blood test, screening, mass screening, guaiac fecal/faecal occult blood test, immunochemical fecal/faecal occult blood test, comparison, test comparison.</td>
<td>Evidence from articles was included in the review if the study: • was judged to be well-designed (controlled comparisons of FOBTs, with colonoscopic follow-up) • had a significant cohort size (≥ 500 subjects in main cohort) • reported acceptable measurements of performance including test positivity, sensitivity (Se) and specificity (Sp).</td>
<td>The search strategy yielded 975 articles. Performance characteristics (ranges) of different types of FOBTs from screening populations of ≥ 500 subjects</td>
<td>Final number of articles included in review is not summarised. Not clear how data extraction and review of internal validity of included studies was undertaken. No precision of test effect estimates presented. Authors’ conclusions Of the guaiac tests, Hemoccult and HemoccultSENSA require the most serious consideration as they are the most studied. The latter is more sensitive for cancer but at a higher TPR. Of the immunochemical tests (ICTs), HemeSelect and Immudia HemSp are the best studied. No other laboratory-developed ICT has yet been evaluated in large population screening studies, as Flexsure OBT is no longer available commercially. At present there is no extensively studied FOBT that fulfils the needs for all target populations worldwide. The choice of FOBT should take into consideration population dietary compliance and colonoscopy resources, in addition to quality control of the screening method and follow-up.</td>
</tr>
<tr>
<td>World Health Organization and World Organization for Digestive Endoscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Guaiac tests
Rehydrated Hemoccult: Test positivity rate (TPR*) = 8-15%, Se#= 92%, Sp** = 90%
Hemoccult II:
TPR= 1-5%, Se= 37-81%, Sp= 94-99%
HemoccultSENSA:
TPR= 5-17%, Se= 79%, Sp= 88-92%

Immunochmmical tests
Heme Select [previously called Immudia HemSp]:
TPR= 3-6 %, Se= 69%, Sp= 95%
Flexsure OBT:
TPR= 1.7-2.9% Se and Sp not reported

Combined guaiac/immunochemical
HemoccultSENSA/ HemeSelect (“two-tier”):
TPR= 3%, Se= 66%, Sp= 98%

*TPR = Percentage of people tested returning a positive FOBT result.
#Se = sensitivity for cancer (studies required long-term follow-up to allow detection of interval cancers to be included in review).
**Sp = specificity for ‘neoplasia’ (includes cancer and adenomas)
**Table 11. Secondary research appraised relevant to accuracy of immunochemical FOBT screening for CRC compared with guaiac FOBT screening (continued)**

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piper (2004) Blue Cross Blue Shield Association Technology Evaluation Center, US</td>
<td>Search: 1985-June 2004. Restricted to articles with English language abstracts. Databases searched: Reported that an electronic search was supplemented with hand searches of prior reviews in this area, recent articles, the Cochrane Collaboration Library, and articles identified through expert contacts. Key Words: fecal, faecal, immuno, colon cancer.</td>
<td>Population: average-risk men and women ≥ 50 years. Comparators: ICTs compared to guaiac FOBTs Outcomes: CRC incidence, stage and mortality (or, given that Hemoccult II is only test with direct evidence for these outcomes, indirect evidence of these outcomes in studies comparing the performance of ICTs and guaiac) To review FOBT performance Included studies that met minimum criteria of: • published as full-length article in the English language • enrolled a well-described population for CRC screening (preferably population or community-based) excluded subjects with known causes of bleeding • FOBT completion by 85% enrollees • endoscopy for all enrollees • compared 2 or more FOBTs, including at least one ICT (preferably included Hemoccult II as one of the FOBTs compared) • no rehydration for Hemoccult tests • reported FOBT results for CRC, adenomas&gt;1cm or both combined. To review compliance with FOBTs Included studies that were published in peer-reviewed journals and compared compliance for at least 2 FOBTs or different sets of performance conditions for the same test.</td>
<td>FOBT performance 7 studies met minimum criteria. 4 compared ICTs to Hemoccult II, 2 compared ICTs to HemoccultSENSA, 1 compared 2 different ICTs. All studies were assigned a quality rating of ‘fair’. None enrolled an average-risk CRC screening population. No studies were designed to assess programmatic screening performance (repeated screening over years). Basic conclusions from evidence for comparative performance of FOBTs: • PPV for FOBT studied was almost always the best value for FOBT with best Sp • ICTs have better Se than Hemoccult II, but not necessarily better than HemoccultSENSA • IFTs have better Sp than HemoccultSENSA, but Sp is not clearly as good or better than Hemoccult II. Only 2 studies evaluated FOBTs that were commercially available in the US at that time. FOBT compliance 3 studies + 1 meta-analysis. No quality rating reported for included studies. 1 study assessed effect of faecal sampling method on patient compliance and found that requirements for less faecal samples and less contact with faecal material significantly increased compliance. Several tested the effect of dietary restrictions versus none, but study settings, populations and specific dietary restrictions differed across studies such that it was not possible to draw conclusions regarding the effect on patient compliance.</td>
<td>Did not specify which electronic databases were searched. Not clear how data extraction and review of internal validity of included studies was undertaken. Authors’ conclusions The evidence on clinical performance of immunochemical FOBTs currently available in the US compared to guaiac FOBTs is insufficient for drawing conclusions. There is insufficient evidence to permit conclusions regarding the use of immunochemical faecal occult blood testing for CRC screening and health outcomes. The use of immunochemical FOBT testing does not meet the technology evaluation centre criteria.</td>
</tr>
</tbody>
</table>
Chapter 5: Effectiveness of Flexible sigmoidoscopy

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY

The search identified two eligible primary research papers that presented data on incidence of colorectal cancer or overall mortality for the Norwegian Telemark Polyp Study, a pseudo-randomised controlled trial graded as Level III-1 evidence. Mortality data at 10-year follow-up was reported in the Working Party on Screening for Colorectal Cancer’s (1998) report for the National Health Committee of New Zealand. Thirteen-year follow-up data is now available for this trial. Full details of the two papers appraised relevant to the Telemark Polyp Study, including methods, key results, limitations and conclusions, are provided in evidence Table 12. Studies are presented in chronological order of publication within the table.

In addition to the Norwegian trial, three large RCTs investigating flexible sigmoidoscopy are underway. As mortality and incidence data are not yet available from these trials, they have not been formally appraised. However an overview of these ongoing trials is provided as their long-term health outcome data will be of critical interest to the National Screening Unit once available.

The Telemark Polyp Study1 (TPS-I)

Study setting and sample

The Telemark Polyp Study was the first prospective, controlled study that aimed to investigate the impact of FS on colorectal cancer incidence and mortality. In 1983, 400 men and women aged 50-59 years (born in January or February) were randomly drawn from the population registry of Telemark, Norway (from 9957 people in this age group) and offered a one-off flexible sigmoidoscopy. The participation rate was 81 percent (n=324). The trial was pseudo-randomised. A control group of 399 people were drawn (without month of birth restrictions) from the same registry but were not contacted upon enrollment. Those in whom polyps were detected had a full colonoscopy with polypectomy in 1985 and 1989.

In 1996, survivors from both the screening and control groups were invited to have a colonoscopic investigation for detection and removal of polyps (Thiis-Evensen et al., 1999; Thiis-Evensen et al., 2001). Those allocated to the screening group who had refused screening at baseline in 1983 (n=76) were not invited for follow-up. Of those invited, participation for follow-up in 1996 was 71 percent, which is 58 percent of the original population of 400 people who were selected for the trial.

Outcomes

Stage CRC detected

In 2001, Thiis-Evensen et al. presented data on the prevalence of adenomas at the 13-year follow-up in 1996, including high-risk adenomas (severe dysplasia, adenomas 10mm or greater in diameter, villous components). There was no difference in the prevalence of adenomas found in the screening group (37%) compared with the control group (43%) [RR=0.9; 95% CI 0.7-1.1; p=0.3]. The relative risk of having adenomas 5mm or greater in diameter in the screening group compared with the control group was 0.7; 95% CI 0.5-0.95; p=0.03. There was a non-significant trend toward fewer high-risk adenomas being found at follow-up in those in the screening group (8%) compared with the control group (13%) [RR=0.6; 95% CI 0.3-1.0; p=0.07]. The authors argue that the one-time FS screening may have a limited long-term impact on adenoma prevalence in an average-risk population, but may have some protective effect against high-risk adenomas consistent with CRC prevention.
Incidence of CRC

Thiis-Evensen et al. (1999) described CRC incidence and overall mortality outcomes at 13-year follow-up. Medical records and Norwegian Cancer Registry files were searched to register any cases of CRC in the period 1983-1996. Using intention-to-treat analyses, the registered incidence of CRC between 1983 and 1996 (including those identified at screening and follow-up) was 2/400 enrolled in the screening group and 10/399 enrolled in the control group [RR=0.20; 95% CI 0.03-0.95; p=0.02]. None of the attendants of the baseline FS (in the screening group) developed CRC, compared with 11 cases in the Control Group (CG) and non-attendants of the Screening Group (SG) group (p=0.006). Thiis-Evensen et al. (1999) concluded that the endoscopic screening with polypectomy reduced the incidence of CRC in the Norwegian average-risk population.

Overall mortality

Total mortality rates and causes of death were obtained from searching Norwegian Bureau of Statistics data. Higher overall mortality was observed in the screening group (n=55/400 enrolled, 14%) than the control group (n=35/399 enrolled, 9%) [RR=1.57; 95% CI 1.03-2.4; p=0.02]. These rates compare to 12 percent for the population from which the samples were drawn (Thiis-Evensen et al, 1997). The possibility was raised by Thiis-Evensen et al. (1999) that participation in screening may have made screenees less motivated to improve their lifestyle which may have increased all-cause mortality rates in the screened group. However the small size of this trial means that the difference in mortality may be a chance variation. Another limitation of the study is that allocation to screening or control group was not truly randomised; the screening group was drawn from those born in January or February whereas month of birth restrictions were not applied to the control group. As there was some month-of-birth-related variation in total mortality across the cohort, the month-of-birth variation between the screening and control groups may have introduced selection bias into the results, hypothesised as related to seasonally-related hormonal or dietary changes in pregnancy (Hoff et al, 2001). Moreover, those enrolled in the control group were unaware of enrolment and therefore were not subjected to the same exclusion criteria as the screening group, which may have compromised comparability of the intervention and control groups.

Conclusion

The Telemark trial was primarily designed as a small-size feasibility study (Hoff et al, 2001). Larger, adequately controlled trials of comparable screening and control groups are necessary to demonstrate reliably any impact of screening in CRC incidence and mortality at follow-up.
Table 12. Primary research studies appraised investigating the effectiveness of FS screening on outcomes from CRC compared with no screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>Evidence</th>
<th>Grade</th>
<th>Country</th>
<th>Study design</th>
<th>Evidence</th>
<th>Grade</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>This-Evensen et al. (1999)</td>
<td>RCT</td>
<td>Grade III-1</td>
<td></td>
<td>Telemark, Norway</td>
<td>Intervention</td>
<td></td>
<td></td>
<td>In 1983, 400 men and women aged 50-59 years were randomly drawn from the population registry of Telemark (from 19997 people in this age group). Those in the screening group were drawn from those born in January or February whereas those in the control group were drawn irrespective of month of birth. 324 (81%) of the screening group (SG) accepted FS screening (mean age: 54.4 years). A control group (CG) of 399 people was drawn from the same registry but not contacted upon enrollment.</td>
<td>Medical records (for all but five people) and the files of The Norwegian Cancer Registry (with reportedly 99% completeness) were searched to register any cases of CRC in the period 1983-1996. Total mortality rates and causes of death were investigated through the Norwegian Bureau of Statistics.</td>
<td>The registered incidence of CRC between 1983 and 1996 (including those identified at screening and follow-up) was 2/400 enrolled in the SG and 10/399 enrolled in the CG [RR=0.20; 95% CI 0.03-0.95; p=0.02]. None of the attendants of the baseline FS developed CRC compared with 11 cases in the CG and non-attendants of the SG (p=0.006). Higher overall mortality was observed in the screening group (n=55/400 enrolled, 14%) than the control group (n=35/399 enrolled, 9%) [RR=1.57; 95% CI 1.03-2.4; p=0.02]. These compare to 12% for the population from which the samples were drawn.</td>
<td>Limitations: The screening group at 13-year follow-up excluded those who had refused the initial FS, which may have biased incidence data at follow-up. Allocation to group was not strictly random; the screening group was selected based on birth dates of January and February whereas no such restriction was made for the control group members which were age- and sex-matched. This may have introduced a selection bias into the mortality statistics that may vary as a function of time of year of birth (e.g., seasonally related hormonal or dietary changes in pregnancy). Moreover, those enrolled in the control group were unaware of enrolment and therefore were not subjected to the same exclusion criteria as the screening group [which were not detailed in this paper]. Behaviour patterns relating to risk factors (e.g., diet, smoking) may have changed in the screening group where participants were aware of the study. These factors may have compromised the comparability of the screening and control groups. Patients were unable to be blinded to group allocation. It is difficult to blind the endoscopist at follow-up to whether the patient belonged to the SG or CG as unsedated patients tend to disclose their previous experience with colonoscopy.</td>
</tr>
</tbody>
</table>
Table 12. Primary research studies appraised investigating the effectiveness of FS screening on outcomes from CRC compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>This-Evensen et al. (1999) Telemark, Norway (Continued)</td>
<td>Of attendees, there were no significant differences between SG and CG in age, sex distribution, family history of CRC, smoking status, prevalence of atherosclerotic disease, diabetes mellitus, inflammatory bowel disease, abdominal complaints, extracolonic cancer, or various bowel symptoms.</td>
<td>Authors’ conclusions The authors concluded that the endoscopic screening with polypectomy reduced the incidence of CRC in the Norwegian normal population. With respect to mortality, due to the nature of the intervention it was not possible to blind participants invited for a baseline FS to their allocated group, whereas those in the control group were unaware of their enrollment. The authors argue that, whilst unlikely, one cannot rule out the possibility that participation in screening may have made screenees less motivated to improve their lifestyle as they may have felt “relieved of the risk of malignancy”. Reviewers’ conclusions The intervention group difference in all-cause mortality may be a chance variation given the relatively small sample size of the trial. The limitations of the study also may lead to confounding and bias, which cast doubt concerning the robustness of the study findings. Larger trials are necessary to demonstrate screening effects on CRC incidence and mortality.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 12. Primary research studies appraised investigating the effectiveness of FS screening on outcomes from CRC compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thiis-Evensen et al. (2001)</td>
<td>RCT Grade III-1</td>
<td>Intervention Offered a one-off flexible sigmoidoscopy in 1983. Those in whom polyps were detected had a full colonoscopy with polypectomy in 1985 and 1989. Control Were not informed of their status as enrolled controls usual care and received no contact until 13-year follow-up.</td>
<td>In 1983, 400 men and women aged 50-59 years were randomly drawn from the population registry of Telemark (from 9957 people in this age group). Those in the screening group were drawn from those born in January or February whereas those in the control group were drawn irrespective of month of birth. 324 (81%) of the screening group (SG) accepted FS screening (mean age: 54.4 years). A control group (CG) of 399 people were drawn from the same registry but not contacted upon enrollment. Exclusion criteria: none mentioned for enrollment. At 13-year follow-up, people who had emigrated (n=2), were deceased (n=47 in screening group, n=41 in control), and those who had refused FS screening in the screening group (n=76), were not invited for colonoscopy. Participation of those invited for 13-year follow-up in 1996: 210/277 (76%) of the SG and 356/399 (68%) of the CG. Overall 71% of those invited to participate at 13-year follow up did so, mean age was 67 years (range: 63 – 72 years), 48% were women. Of those invited for follow-up in the SG, 189 (68%) chose colonoscopy, and 12 (4%) chose sigmoidoscopy. Of those invited for follow-up in the CG, 202 (60%) chose colonoscopy, and 19 (5%) chose sigmoidoscopy. Overall, 18 (3%) of those invited for follow-up (9 in each group) had had a colonoscopy for gastrointestinal symptoms within the previous six months.</td>
<td>Prevalence of adenomas including high-risk adenomas (severe dysplasia, adenomas 10 mm or greater in diameter, villous components) for the SG and CG at 13-year follow-up.</td>
<td>In 1996 (at 13-year follow-up), 153 (73%) of individuals in the SG and 180 (75%) in the CG had a total of 1605 polyps detected and removed. Adenomas were found in 78 (37%) of those in the screening group and 103 (43%) in the control group (RR=0.9; 95% CI 0.7-1.1; p=0.3). In the SG, 61 (29%) had one or two adenomas and 17 (8%) had three or more, compared with the CG where 78 (32%) had one or two adenomas and 25 (10%) had three or more adenomas [Relative Risk for having 3 or more adenomas in the SG c.f. CG was 0.8; 95% CI 0.4-1.5; p=0.5]. The relative risk of having adenomas 5mm or greater in diameter in the SG c.f. the CG was 0.7; 95% CI 0.5-0.95; p=0.03. High-risk adenomas were found in 16 (8%) of those in the screening group and 32 (13%) of those in the control group [RR=0.6; 95% CI 0.3-1.0; p=0.07].</td>
<td>Limitations The screening group at 13-year follow-up excluded those who had refused the initial FS, which may have biased data at follow-up. Those enrolled in the control group were unaware of enrollment and therefore were not subjected to the same exclusion criteria as the screening group (which were not detailed in this paper). Behaviour patterns relating to risk factors (e.g., diet, smoking) may have changed in the screening group where participants were aware of the study. These factors may have compromised the comparability of the screening and control groups. Patients were unable to be blinded to group allocation. It is difficult to blind the endoscopist at follow-up to whether the patient belonged to the SG or CG as unsted patients tend to disclose their previous experience with colonoscopy. Authors’ conclusions After 13 years, there was no significant difference in adenoma prevalence between those offered an on-off FS and those in the control group provided with usual care. There was a trend toward more high-risk adenomas in the control group compared with the screening group. This may suggest a protective effect of FS screening against the development of high-risk adenomas consistent with the CRC preventive effect of polypectomy.</td>
</tr>
</tbody>
</table>
Table 12. Primary research studies appraised investigating the effectiveness of FS screening on outcomes from CRC compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tid-Eversen et al. (2001) Telemark, Norway (Continued)</td>
<td></td>
<td>Of attendees, there were no significant differences between SG and CG in demographic or clinical variables.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reviewers’ conclusions
The limitations of the study may lead to confounding and bias, which cast doubt concerning the robustness of the study findings. Larger trials with adequate randomisation and strictly comparable groups are necessary to demonstrate reliably the effect of screening on the more significant health outcomes of CRC incidence and mortality.
Overview of eligible trials currently underway with baseline data

Three large RCTs investigating flexible sigmoidoscopy are underway. Two parallel multi-centre trials employing the same protocol in the UK and Italy are investigating once-only screens for people aged 55-64 years, and a US trial is investigating five-yearly screening for people enrolling aged between 55 and 74 years (as part of a broader cancer screening trial). Only limited baseline data on acceptability, safety, feasibility and yield is currently available (following the comprehensive search strategy completed on 17 January 2005). An overview of these trials is provided below.

The UK Flexible Sigmoidoscopy Screening Trial

The UK Flexible Sigmoidoscopy Screening Trial is a multi-centre, randomised controlled trial investigating whether single (once only) flexible sigmoidoscopy screening offered at around the age of 60 years can lower the incidence and mortality of colorectal cancer. The control was usual care, receiving no further contact from investigators. The study design was described by Atkin et al. (2001) and baseline data following completion of the recruitment and screening phases is now available (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002). The UK trial aims to identify the optimal interval (55-59 or 60-64 years) for single screening by including people aged within the range 55-64 years. General practitioners with each participating practice from 14 geographical centres across the UK scanned their lists of patients to remove those ineligible. Remaining men and women aged 55-64 years (n=375,744) were approached by postal questionnaire (with reminder letter sent to non-respondents) and asked whether they would attend for FS screening if invited as well as questions relating to exclusion criteria (strong family history of CRC, health problem precluding benefit from screening, worrying bowel symptom requiring investigation). Fifty-five per cent were interested and those eligible were then randomised (ratio of one to two, where the unit of randomisation was the household) to FS screening or control. Attendance was relatively high with 71 percent (n=40,674) of those invited attending for screening (following a single enema). This approach of enrolling only those eligible who had previously expressed interest in FS screening was chosen to increase compliance rates and statistical power in the study. Population coverage of 39 percent was obtained using the two-stage recruitment method.

Following a single enema (usually self administered at home), screenees were examined using a 60 cm video-endoscope (using carbon dioxide for insufflating the bowel). During screening, small polyps were removed and those with no polyps or low risk polyps detected were discharged (based on the argument that the risks of colonoscopy may outweigh the benefits in individuals with only low-risk adenomas). A repeat FS (usually due to inadequate bowel preparation) was required for 5 percent of screenees (60% of these were performed on the same day). The 5 percent (n=2131) of screened patients found to have high-risk polyps (three or more small adenomas, size 1cm of greater, villous, severely dysplastic, or malignant) were referred for colonoscopy (plus surveillance). Prevalence of neoplasia was high, with distal adenomas detected in 12.1 percent (n =4931) of those screened by FS and distal cancer in 0.3 percent (n=131). Proximal adenomas were detected in 18.8 percent (n=386) of those undergoing colonoscopy and proximal cancer in 0.4 percent (9), 62 percent of which were Duke’s stage A or locally excised.

The screening regimen was considered to be safe; there was one perforation in more than 40,000 flexible sigmoidoscopies (including over 19,000 polypectomies) and four perforations among 2,377 people (0.2%) having a colonoscopy. Twelve people were admitted to hospital for bleeding after FS. Eighty percent reported on the day of their FS (98% response rate) that they experienced no or mild pain, 3 percent experiencing severe pain. In a three month follow-up questionnaire (91% response rate), only 5 percent reported finding the test more than mildly embarrassing, 98 percent were glad that they had the test and 97 percent said they would encourage a friend to have a test if asked.

Pilot studies began in September 1994 with recruitment for the main study conducted between October 1996 and March 1999. The last baseline FS occurred in July 1999. The cohort will be followed up for 15 years using records held at the Office of National Statistics. Results from the pilot phase and baseline findings described above suggest that the screening regimen is likely to be acceptable and safe in terms of attendance rates, adverse events, and psychological morbidity, and feasible in terms of workload (UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002; Atkin et al, 1998; Wardle et al, 2000; Sutton et al, 2000; Taylor et al, 2000). Ultimately, these potential costs and harms will need to be weighed against the health benefits of screening evident from long-term follow-up.
high yield of advanced lesions and early colorectal cancer is consistent with substantial reductions in colorectal cancer incidence and mortality. In view of the large number of prevalent cancers detected at entry, the investigators are of the view that it might be necessary to wait several years before seeing any effect on incidence reduction. Therefore they intend to wait until 2008 before conducting the first analysis of incidence and mortality data (personal correspondence, Professor Wendy Atkin, principal investigator, Flexible Sigmoidoscopy Screening Trial, 15 February 2005).

The Italian SCORE Trial

A parallel multi-centre trial using the same protocol as the UK trial (with some variations described below) is underway across six geographical centres across Northern Italy (known as the SCORE Trial). The methods are as described for the UK trial and sample, attendance, yield, safety and acceptability outcomes are briefly outlined here as reported by Segnan et al. (2002).

Men and women aged 55-64 years (n=236,568) were approached by postal questionnaire (with no reminders sent). Twenty-four per cent (n=56,532) responded with interest in FS screening and those eligible (n=34,393, 16% of those originally approached) were then randomised to FS screening or control (in a 1:1 ratio where the unit of randomisation varied between centres). Attendance was moderately high with 58 percent (n=9999) of those invited attending for screening (n=9911 actually examined). Population coverage of only 9 percent was obtained using the two-stage recruitment method, significantly less than the 39 percent achieved in the UK trial, which may relate to design differences (the lack of reminder letters sent after the first approach letter) as well as those related to the characteristics and cultures of the source populations.

Screenees were examined using a 140cm colonoscope in five centres and a sigmoidoscope in one centre (Genova). The sigmoidoscopy could be fully completed on a single occasion in 79.8 percent of those remaining, the examination was terminated due to pain or bowel adhesions (37%), or was terminated due to inadequate visualisation of the colonic mucosa following unsatisfactory bowel preparation (63%). In the latter situation, only where no visualisation was possible were new tests offered (with 86% attending). The researchers discussed the need to repeat sigmoidoscopy in the event of unsatisfactory bowel preparation as a crucial problem, particularly given the apparent reluctance of physicians to request a second sigmoidoscopy examination when only partial visualisation was possible.

Eight percent (832) of screened patients were referred for colonoscopy, 93 percent (775) of whom attended. Prevalence of neoplasia was high, with distal adenomas detected in 1070 (10.8%) of those screened by sigmoidoscopy. Proximal adenomas were detected in 116 (15.5%) of 747 screenees without cancer at sigmoidoscopy who then underwent colonoscopy. In total, 54 people had screen-detected colorectal cancer, a rate of 5.4 per 1000 (54% of which were Duke’s stage A), including 47 with distal cancer detected during sigmoidoscopy, and seven colorectal cancers detected at colonoscopy. The high yield of advanced adenomas is consistent with the projected impact of sigmoidoscopy screening for colorectal cancer incidence. The wide variability in detection rates across trial centres led the SCORE trial investigators to conclude that quality-control procedures are needed when planning a population-based colorectal cancer-screening programme.

The screening regimen was considered to be safe; there was one perforation in more than 9911 flexible sigmoidoscopies and one perforation and one haemorrhage requiring hospitalisation following polypectomy among 775 people having a colonoscopy. Minor self-limited complications occurred in 60 (0.6%) of those who had a sigmoidoscopy and in 30 (4%) of those receiving colonoscopy. Sixty percent reported on the day of their FS (95% response rate) that they experienced mild discomfort; 5 percent reported finding the test more than mildly embarrassing.

The main differences between the UK and Italian trial designs related to randomisation procedures and criteria for colonoscopy referral. Cluster randomisation (using the physician as the unit of randomisation instead of the household) was undertaken in the SCORE trial in three Italian centres to reduce the probability of contamination in the context of open-access endoscopy (i.e., through the physician). Regarding the criteria for referral, in Italy all people with distal polyps larger than 5mm were referred whereas in the UK trial, only those with distal polyps larger than 10mm were referred for colonoscopy. Further, only in the Italian trial was there an option to refer people with incomplete sigmoidoscopies and with any distal polyp detected for colonoscopy. In a separate paper using data
from this trial, Senore et al. (2004) investigated systematically the difference in the yield of proximal lesions of these two different referral policies. Of patients with polyps greater than 5mm, 29 (6.9%) were detected with an advanced proximal neoplasm (including four colorectal cancers). The prevalence of proximal advanced neoplasia was 9.4 percent among patients with high-risk distal polyps and 2.5 percent among those with low-risk lesions (adjusted odds ratio, 3.19; 95% CI, 1.06-9.59). This latter rate among people with low-risk 6 mm-9 mm distal polyps is similar to the prevalence among people without distal polyp. Senore et al. (2004) concluded that restricting colonoscopy referral to patients with high-risk distal polyps might represent a cost-effective strategy in a screening context.

Recruitment for the main study was conducted between October 1995 and April 1999. Colorectal cancer incidence and mortality will be followed up through regional registries of hospital discharge records, death certificates, and local cancer registries. The investigators estimate that they will commence analysing incidence data after 7-8 years of follow-up (personal correspondence, Dr Carlo Senore, SCORE Trial, 16 March 2005).

The USA Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO)

The Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) is a randomised, controlled community-based, multi-centre study evaluating the effectiveness of cancer screening tests on site-specific cancer mortality. Sponsored by the National Cancer Institute, the trial is being conducted in 10 screening centres in the United States. Several papers document the history, design, recruitment strategies, quality control procedures and adherence to follow-up of this important trial (Gohagan et al, 2000; Prorok et al, 2000; Simpson et al, 2000; Weissfeld et al, 2000; Weissfeld et al, 2002). Randomisation began in November 1993 and was completed in July 2001, enrolling more than 154,000 asymptomatic men and women aged 55 to 74 years. Follow-up is anticipated for at least 13 years from randomisation. Participants in the intervention arm of the trial undergo several periodic cancer screening tests including flexible sigmoidoscopy, chest radiograph, digital rectal examination and prostate-specific antigen screening (for men), and cancer antigen (CA) 125 screening and vaginal ultrasound (for women). Relating to colorectal cancer, the trial is evaluating the effect of a 60cm flexible sigmoidoscopy (performed by trained nurses, certified internists, or specialists) on colorectal cancer mortality. Initially the trial design dictated that people enrolled in the study would receive an initial FS at inception and again at three years. To be consistent with prevailing practice in the United States, the screening interval was subsequently extended to five years for those enrolled during or later than December 1995.

Participants completed a baseline questionnaire recording sociodemographic characteristics, medical history, cancer-screening history over the previous three years, and family history. Recruitment occurred through mass mailings. Eligibility criteria included: aged between 55 and 74 years, no current cancer treatment except for basal cell or squamous cell skin cancer, no known prior cancer of the colon, rectum, prostate, lung, or ovaries, no surgical removal in these sites, no participation in another cancer screening or prevention trial, no finasteride use (in men) or tamoxifen use (in women) in the past six months, provision of informed consent, no more than one prostate-specific antigen test in the past three years (for men randomised after April 1995), and no colonoscopy, sigmoidoscopy, or barium enema in the past three years (for individuals randomised after April 1995). Cancer cases are identified through annual follow-up questionnaires, and all deaths are identified through vital status tracing mechanisms (Hayes et al, 2000).

At each positive FS examination (where a polypoid lesion or mass was observed), examiners recorded location, shape and estimated size (of the four largest); however, lesions were usually not biopsied or removed. Screenees were referred to their personal physician for diagnostic workup and follow-up. Data from repeat FS or colonoscopy relating to pathology, size and location of each lesion were recorded. Following the first round of FS screening, some preliminary data on indicative baseline yield has been released, with yield of advanced distal adenomas reported as 2.5 percent, and cancer in the distal colon as 27 per 10,000 (Schoen et al, 2003). Comprehensive, finalised baseline data is not yet available, though has been provisionally accepted for publication and will be published in 3-4 months (personal correspondence, Dr Schoen, investigator, Flexible Sigmoidoscopy Screening Trial, 21 February 2005).

Yield data is available from participants drawn from the PLCO trial who received a three-year repeat flexible sigmoidoscopy following negative examination on their initial FS (Schoen et al, 2003). Of
11,583 screenees eligible for follow-up after an initial negative examination, 9317 (80.4%) returned for their three-year repeat sigmoidoscopy. A total of 13.9 percent (n=1292) had a polyp or mass detected of whom 73.6 percent (n=952) underwent follow-up diagnostic testing: 104 with repeat FS, and 847 with colonoscopy. Yield results for the distal colon included 214 non-advanced adenomas (2.3%), 72 advanced adenomas (0.77%) and six cancers (0.06%). It is important to note that this sample included individuals with inadequate examinations at initial FS due to inadequate preparation or depth of insertion of less than 50 cm. These results therefore may include false negatives from the initial screen, however 80.6% (n=58/72) of advanced distal adenomas detected at follow-up were found in a portion of the colon that had been reportedly adequately examined at the baseline screen. While admitting that the detection of abnormalities is modest, the trial investigators argue that they are nevertheless of concern with respect to the impact of longer screening intervals. Ultimately, decisions about screening interval need to consider evidence of reduced incidence and mortality (from long-term follow-up) that is significant when compared with the potential costs and harms of screening. Cancer incidence and mortality data as well as deaths from all causes for the PLCO trial is unlikely to be available for 5-7 years (personal correspondence, Dr Schoen, investigator, Flexible Sigmoidoscopy Screening Trial, 23 February 2005). Interpretation of mortality data will need to consider the impact of a multiple range of site-specific cancer screening methods being employed, as well as opportunistic screening that occurs outside of the study protocol for people enrolled in control as well as screening groups.

SECONDARY RESEARCH

The search strategy identified only two relevant systematic reviews that reported on RCTs relating to the effectiveness of FS screening for colorectal cancer. The methods and conclusions of these reviews are described in Table 13. These papers may not have employed the same inclusion and exclusion criteria as applied in this review and the results must be interpreted with care.

The report by Australian AHTAC review (1997), which was considered by the Working Party on Screening for Colorectal Cancer (1998), aimed to assess the evidence on the benefits, risks and costs of CRC screening for asymptomatic people (as well as surveillance for those at higher-risk of CRC). A thorough search strategy was employed, including reference lists of retrieved papers and grey literature. Narrative summaries of case-control trials were presented, but only one RCT was identified, the Norwegian Telemark Polyp Study, which was a pseudo-randomised trial graded as Level III-1 evidence. The 13-year follow-up data reported in the current review was identified through personal correspondence with the trial’s investigators as indicating significantly reduced incidence of colorectal cancer in those enrolled in the screening group compared with those in the control group. Limitations of the study, including its small sample size and the possibility that comparability of the screened and control groups was compromised by the study design, were discussed. The AHTAC (1997) review’s authors concluded that sigmoidoscopy showed promise as a screening instrument, but definitive evidence from randomised controlled trials currently underway is required before widespread screening can be given serious consideration.

The other review identified was by Walsh and Terdiman (2003) from the University of California and considered the evidence for colorectal cancer screening including flexible sigmoidoscopy (this review was also appraised in Chapter 3 in relation to FOBT, and Chapter 6 in relation to combined screening). The search considered English language articles published to August 2003 identified using the PubMed database, reference lists of key articles, consultation with experts, and hand searching of reference lists of key articles, journals, and proceedings of professional meetings. Studies were identified that evaluated colorectal cancer screening in healthy individuals and assessed clinical outcomes. The review reported briefly on one small pseudo-randomised controlled trial, the Telemark Polyp Study in Norway. The trial was said to have demonstrated that one-time FS screening could reduce colorectal cancer incidence, however no reduction in colorectal cancer mortality was observed in the screened group. The review mentioned large trials in progress, including the PLCO trial in the US and the UK trial (appraised in this chapter). In addition to RCT evidence, mortality outcomes from several retrospective case-control studies were reported. The review section also included data on performance characteristics, and a description of how the FS test is performed and results interpreted. A limitation of the systematic review was that it lacked detailed reporting of the methodological quality and design of the studies appraised. The reviewers concluded that while there was direct and indirect evidence that the available tests for CRC screening for men and women aged 50 years or older were effective, the evidence appraised did not support choosing one test over another.
CONCLUSIONS

The current review only identified two papers reporting on one completed RCT relating to the impact of FS on colorectal cancer incidence and mortality, the pseudo-randomised Telemark Polyp Study (Thiis-Evensen et al, 1999, Thiis-Evensen et al, 2001). This trial was originally planned as a feasibility study and its small size (400 allocated to the screening group) and variations in the criteria for allocation to intervention and control groups raise concerns about their comparability. Three large multi-centre trials are currently underway which are designed to investigate the impact of FS screening on CRC incidence and mortality. These include the UK Flexible Sigmoidoscopy Screening Trial and the Italian Score Trial (using similar protocols) which offered one-time FS screening for those aged between 55 and 64 years. These trials are also exploring the optimal age for the one-time FS screen within this age range. The Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) being conducted in the US is evaluating the effectiveness of repeated cancer screening tests, including FS screening, on site-specific cancer mortality for 55-74 year olds. Incidence and mortality data will not become available until at least 2008 for the UK trial (and probably for the Italian trial which used the similar protocol and recruitment timeframe), and 2010-2012 for the US trial.

Given the lack of incidence and mortality data available for these ongoing trials, the impact on health outcomes of flexible sigmoidoscopy screening is currently unknown. In the absence of this information, one cannot recommend the introduction of a national screening programme using flexible sigmoidoscopy. However once long-term follow-up data on these major trials is available, this possibility will need to be revisited. Preliminary data from the trials do suggest that population-based screening may be feasible and acceptable, with good participation rates. This holds promise for FS as a screening tool if the health benefits anticipated by these investigators are realised. Cost-effectiveness also needs to be established.
Table 13. Secondary research appraised relevant to the effectiveness of FS screening for CRC compared with no screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Health Technology Advisory Committee (AHTAC) (1997) Australia</td>
<td>Search: 1990-1996. Supplemented by identification of key references for earlier articles. Databases searched: Medline, EMBASE, CancerLit. Also identified articles from the Internet and reference lists of retrieved articles. Current Contents searches were continued up to the date of publication (1997). Accessed “grey” literature by contacting health departments, academic institutions and reviewing submissions to the AHTAC Working Group. Key Words: colorectal cancer, screening, faecal occult blood test, barium enema, economics, costs, epidemiology, prevention, survival, acceptability, compliance.</td>
<td>All study types evaluating FS screening were considered, but greater weight was given to research material scoring highly on NHMRC (1995) guidelines. For studies examining test performance characteristics, those that offered a reasonable prospect of providing valid measures, as determined by two independent reviewers, were included. All identified studies examining issues relating to compliance in which study groups numbered at least 100 were included.</td>
<td>For FS screening: Provided narrative summaries of non-controlled studies. Identified only one RCT underway, the pseudo-randomised controlled Norwegian trial (Telemark Polyp Study), and reported 10-year follow-up incidence data of 1/400 CRC in the screening group (who did not attend for screening) compared with 4/399 in the control group. Following personal correspondence with the author, G Hoff, the AHTAC review also reported 13-year follow-up data of two cancers in the screening group and 10 in the control group.</td>
<td>Review did not specify clear PICO questions. Authors’ conclusions Sigmoidoscopy shows promise as a screening instrument, but definitive evidence from randomised controlled trials (currently underway) is required before widespread screening can be given serious consideration.</td>
</tr>
<tr>
<td>Walsh and Terdiman (2003) University of California, San Francisco, US.</td>
<td>Search: January 1966-August 2002. Databases searched: PubMed. Also identified articles from published meta-analyses, reference lists of key articles and expert consultation. Supplemented search with hand searching and proceedings from national professional organisation meetings. Key Words: colorectal neoplasia, occult blood, sigmoidoscopy, barium enema, colonoscopy all combined with screening.</td>
<td>Population: Asymptomatic subjects who were at average risk for CRC. Intervention: FS screening (other screening modalities reviewed but not reported on here) Outcomes: Clinical outcomes considered included mortality, cancer incidence, and identification of adenomas. Inclusion criteria: English language Observational and diagnostic accuracy studies were evaluated when RCTs were not available (for FS). Exclusion criteria: Not reported.</td>
<td>For FS screening: Identified one small RCT, the pseudo-randomised Telemark Polyp Study in Norway, which demonstrated that one-time FS screening could reduce colorectal cancer incidence. However no reduction in colorectal cancer mortality was observed in the screened group. No large RCTs were identified reporting on mortality. Described two large RCTs in progress, including the Prostate, Lung, Colorectal and Ovarian screening trial in the US, and the UK trial. No outcome data was available to report. The review reported on mortality outcomes from several well-designed, retrospective case-control studies. The review section also included data on performance characteristics, and a description of how the FS test is performed and results interpreted.</td>
<td>Search not restricted to published literature, although only one database considered. There was minimal reporting of design or sample characteristics of reviewed studies, or analysis of their limitations and strengths. Authors’ conclusions The recommendation that all men and women aged 50 years or older undergo screening for colorectal cancer is supported by a large body of direct and indirect evidence. At present, the available evidence does not support choosing one test over another.</td>
</tr>
</tbody>
</table>
Chapter 6: Effectiveness of screening with FOBT and FS combined

PRIMARY RESEARCH: STUDY DESIGNS AND QUALITY

The search identified two eligible primary research studies reporting data on CRC incidence and/or mortality outcomes for FOBT and FS screening combined. Below is an overview of study designs and aspects of quality represented by these studies. Full details of the two studies appraised, including methods, key results, limitations and conclusions, are provided in evidence Table 14 comparing FOBT and FS combined with FOBT alone. No eligible studies were identified comparing FOBT and FS combined with no screening.

Three other RCTs were also identified related to the effectiveness of FOBT and FS combined but did not report data on CRC incidence and/or mortality. The outcomes focus of these studies was on screening compliance, diagnostic yield and colonoscopic findings from once-only screening. Full appraisals of these three studies of the accuracy of combined FS and FOBT screening are presented in Appendix 6.

An on-going trial, the Norwegian NORCCAP study, compared FOBT and FS combined with FS alone. Gondal et al. (2003) presents baseline screening compliance and diagnostic yield findings for this study and these are summarised in this chapter.

Studies are presented in chronological order of publication in Table 14.

The search identified two eligible RCTs that compared FOBT and FS combined with FOBT alone and reported on incidence and/or mortality data. One RCT (Rasmussen et al, 1999), known as the Funen-2 trial was graded as level II evidence, with adequate design and conduct. The other RCT (Rasmussen et al, 2003), known as the Funen-3 study was graded as level III-1 evidence due to limitations in study design and conduct, including pseudo-randomisation, possible subject selection bias, and limitations in the comparative analysis.

**FOBT and FS combined screening compared with FOBT screening alone**

The Funen-2 trial

Study setting and sample

In the Funen-2 RCT a sample of 10,978 asymptomatic persons aged 50-75 years was invited for CRC screening between September 1992 and November 1995. Persons already taking part in the Funen-1 screening trial (see Chapter 3) with FOBT (Haemoccult II) testing alone (Kronborg et al, 1996) were excluded. The trial compared two groups randomised before invitation, with one group of 5495 persons invited to undergo one FOBT (Haemoccult II) and one flexible sigmoidoscopy (FS) and the other group of 5483 persons invited to undergo one FOBT (Haemoccult II) alone (Rasmussen et al, 1999). The Funen-2 RCT study by Rasmussen et al. (1999) reported on the additional diagnostic yield of screening with combined once only FOBT and FS compared to once only FOBT alone. A later study known as the Funen-3 study (Rasmussen et al, 2003) compared the additional diagnostic yield of screening with biennial FOBT (Haemoccult-II) over 16 years (eight rounds) from participants in the Funen-1 RCT (Kronborg et al, 1996; Jorgenson et al, 2002; Kronborg et al, 2004 – see Chapter 3) compared to participants from the trial arm of the Funen-2 RCT of persons receiving once-only FOBT and FS (refer to Funen-3 trial).
Those invited to undergo a combined FOBT and FS were required to send in their test samples prior to undergoing the FS, with the physician performing the FS unaware of the FOBT results. Bowel preparation for FS consisted of two days of oral administration of laxatives and one suppository two hours before FS examination. An invitation to undergo full colonoscopy with polypectomy was offered to persons detected with possible neoplasia from FS examination (all polyps > 3mm diameter, and/or mucosal ulcerations, and/or stricturing carcinoma). Similar to the Funen-1 trial, for the FOB testing the Haemoccult II test cards were not rehydrated. A FOBT required two samples from each of three consecutive stools. Screenees were requested to restrict their diet and medications prior to completing Haemoccult II test cards. A positive FOBT was defined as being any one from six slides as positive, and FOBT positive screenees were invited to attend for a medical review that included colonoscopy if possible.

**Outcomes**

Lower compliance in the Funen-2 RCT was reported for the combined FOBT + FS procedure compared to FOBT alone (40% versus 56%, \( p<0.0001 \)). Compliance decreased with increasing age in both groups and the response rate for men and women was similar in the combined FOBT + FS group and lower for men than women having FOBT alone. Compliance rates in non-responders (those enrolled after not responding to first invitation and reminded after 3-weeks) was 32 percent and 31 percent respectively. Complete FS (60cm) was performed in 85 percent of persons examined and partial FS (at least 40 cm) in 96 percent of persons examined. Poor bowel preparation and pain were the main reasons reported for incomplete examination.

**Test positivity rate**

The criteria for a positive test with FS resulted in 18.6 percent of persons undergoing a full examination of the colon with colonoscopy compared to only 2.3 percent of persons with a positive FOBT only. The predictive value of a positive test for CRC was 2.8 percent after combined FOBT + FS and 5.4 percent after FOBT alone. The PPV for a positive FS and FOBT for CRC was 0.27, advanced adenomas 0.42. For negative FOBT and positive FS this was 0.01 and 0.14 and for positive FOBT alone PPV was 0.05 and 0.19 respectively.

**Diagnostic test performance**

No sensitivity/specificity of tests included.

**Stage CRC detected**

The CRC stage at detection was more favourable in the FOBT + FS (six Duke’s A, three B, one C, and two with distant spread) compared with the FOBT alone group (one Duke’s A, one B, and two C). The diagnostic yield of colorectal neoplasia was 12 CRCs for combined FOBT and FS (eight detected by FS and a further four confirmed by colonoscopy) versus four CRCs for FOBT alone and 72 large adenomas versus 14 respectively.

**CRC morbidity/mortality**

Overall, 11 persons died of CRC in the FOBT + FS group and 14 in the FOBT alone group during the study follow-up period (as at December 1997, 24-62 months after inclusion in the screening programme).

**Incidence of CRC**

The CRC clinical status of persons at December 1997 (24-62 months after inclusion in the screening programme) was the same in the FOBT + FS group as in the FOBT only group, with 38 persons in each group with CRCs detected. The interval cancer rate in both groups did not differ significantly among all those who had been screened (\( p=0.24 \)). No CRCs were detected in the follow-up period in persons under surveillance after the detection of adenomas in either screening group. CRCs detected after a negative screening test (interval cancer incidence) were significantly higher among those screened with
FOBT alone (18/22, 82%) compared with those screened with combined FOBT + FS (8/20, 40%), (p=0.01).

Outcomes related to screening

The Funen-2 trial was not designed as a mortality study given the short follow-up period and non-repeat screening. Although mainly baseline findings in terms of compliance and diagnostic yield were presented limited CRC incidence and mortality data were provided from the CRC clinical status of persons assessed at December 1997 (24-62 months after inclusion in the screening programme). Compliance with FS was poor and this may be an issue related to acceptability. The invasiveness of FS, bowel preparation requirements and type of screening invitation are known to influence compliance. Although the diagnostic yield was greater in the combined screening group despite low compliance, the addition of FOBT to FS added nothing to the predictive value of FS. There were 28 persons with negative FS and positive FOBT, but no CRCs and only one advanced adenoma was detected with follow-up colonoscopy, indicating little benefit in diagnostic yield with the addition of once-only FOBT to once-only FS. The greater resource use through the high proportion of follow-up colonoscopies after a positive test by FS could have been reduced by changing the criteria for a positive test to adenomas of at least 10mm diameter and/or multiple polyps which the authors argue would not have increased the risk of CRC.

The Funen-3 study

Study setting and sample

The Funen-3 study compared the diagnostic yield from CRC screening in two age-matched groups from two different randomised trials (Rasmussen et al, 2003). One trial was the Funen-1 RCT, a screening trial with biennial FOBT (Haemoccult II). Results, including CRC incidence and mortality data from this trial after the seventh and ninth rounds of screening over 14-18 years, have been reported (Jorgenson et al, 2002; Kronborg et al, 2004 – see Chapter 3). This study included data from the eighth round of biennial FOB testing. These results were compared to data from the once-only combined FOBT + FS study arm from the Funen-2 RCT (see above).

After excluding persons under 50 years of age and through the random sampling age-matching process, 25,151 persons remained from the Funen-1 FOBT trial and 4460 persons from the Funen-2 FOBT and FS trial. There were no significant age-sex distribution differences between the two trials. There were small pre-invitation but post-randomisation adjustments made so as to exclude deaths, persons with CRC diagnoses, and those with other advanced malignancies.

Outcomes

Compliance to screening with biennial FOBT testing in the Funen-1 trial in the first round was 65.5 percent (95% CI 64.9-66.1) and with once only combined FOBT + FS in the Funen-2 trial was 39.8 percent (95% CI 38.3-41.2), significantly different (p<0.0005). In subsequent rounds compliance in the Funen-1 trial ranged from 90 percent to 94 percent of invited participants.

Test positivity rate

After eight FOBT screening rounds, 1.6 percent (1,425/90,881 tests) were positive, and 8.2 percent (95% CI 7.8-8.6) of persons had one or more positive FOB tests and 87 percent of these had a complete colonoscopy. In the Funen-2 trial, 20.3 percent (95% CI 18.5-22.2) of persons had a positive FOBT and/or FS test (19.4% positive FS, 2.8% FOBT, 1.4% had both tests positive) and 91 percent of these had a complete colonoscopy.

Diagnostic test performance

No sensitivity/specificity of tests included.
Stage CRC detected

The number of CRCs detected in those invited rather than those screened resulted in a higher detection rate in the biennial FOBT programme of 6.5/1000 versus 2.7/1000. Advanced adenomas (≥10 mm) and/or villous structure and/or severe dysplasia were found in 2.3 percent (95% CI 2.1-2.6) of persons screened in the biennial FOBT program over 8-rounds and in 3.3 percent (95% CI 2.5-4.2) of persons in the combined FOBT + FS trial, significantly different proportions (p=0.011). These proportions among invited persons were similar and not significantly different (p=0.32). There was no significant difference in the distribution of Duke’s classification of screen detected CRCs between the two trials (p=0.437). With biennial FOBT screening, screen-detected CRC was 39% Duke’s A, 39 percent B, 15 percent C, distant spread 6 percent. In non-screen-detected CRC in screened persons this was 16 percent Duke’s A, 30 percent B, 22 percent C, distant spread 24 percent.

CRC morbidity/mortality

There was no data reported from this trial regarding CRC morbidity/mortality.

Incidence of CRC

The ongoing biennial FOBT screening Funen-1 programme had a diagnostic yield of screen-detected CRC of 9.9 per 1000 persons screened (95% CI 8.4-11.5) after eight rounds. This was compared with a diagnostic yield of 6.6 per 1,000 persons screened (95% CI 3.4-11.5) in the Funen-2 trial with once-only combined FOBT + FS, no significant difference (p<0.17).

Overall mortality

There was no data reported from this trial regarding overall mortality.

Outcomes related to screening

The Funen-3 study was a comparison of results from two different RCTs. Both screening samples were drawn from a similar population in Funen County, Denmark, but the combined FOBT + FS (Funen-2) trial included a slightly older (50-75 years versus 45-75 years) age group. The main limitations in the Funen-3 study were the comparison of two screening trials started seven years apart and CRC mortality data were not reported. Although diagnostic yield for each of the eight screening rounds, and cumulatively, were reported for the biennial FOBT testing, this was compared with the diagnostic yield of one-time FOBT + FS screening. There was a large difference in colonoscopy-detected CRC rates: one CRC was detected every 30.9 colonoscopies in the once-only combined FOBT + FS study-arm (Funen-2 trial) compared to one CRC detected every 5.1 colonoscopies in the biennial FOBT study-arm (Funen-1 trial). A change in colonoscopy criteria in the detection and removal of small adenomas (≤10 mm) would help redress this high resource use of once-only FS + FOBT screening programme but at a cost of a reduction in CRC yield. When comparing CRC detection rates on a screened-persons basis, once-only combined FOBT + FS offered no advantage over biennial FOBT screening, however on an invited persons basis, much higher CRC detection rates were apparent in the biennial FOBT programme and a similar number of advanced adenomas were removed. The compliance rate for FS and the combined FOBT + FS group was poor compared to FOBT only. As previously stated, participant acceptability may be the issue here with the invasiveness of FS, bowel preparation and type of screening invitation are known to influence compliance. There were 26 persons with negative FS and positive FOBT, but no CRCs and only one advanced adenoma was detected, with follow-up colonoscopy indicating little benefit with the addition of once only FOBT to FS.
ONGOING TRIALS

Baseline screening compliance and diagnostic yield findings are presented for the Norwegian Colorectal Cancer Prevention (NORCCAP) trial. However, health outcomes data on CRC incidence and mortality are not expected until late 2007.

FOBT and FS combined screening compared with FS screening alone

The NORCCAP trial (Norwegian Colorectal Cancer Prevention study)

Study setting and sample

The baseline findings of the ongoing NORCCAP study were presented in this study by Gondal et al. (2003). The purpose of the study was to ascertain the diagnostic yield of screening with FS or a combination of FS and FOBT. The design, organisation and management of the trial have been previously published in the study by Bretthauer et al., 2002. A total of 20,780 average-risk persons, aged 50-64 years from the population registers of Oslo city and Telemark County ‘Norway’ were invited for once-only CRC screening. They were randomised to undergo either FS, using a 140cm Olympus colonoscope or for some cases a 60cm Vision Sciences disposable Endosheath®, or a combination of FS and FOBT, using Flexure® OBT (a immunochemical blood test). Bowel preparation for FS was restricted to an enema, given in attendance before the FS examination. At FS, any lesion identified was biopsied and not removed, and the diagnosis of adenomas was based on histopathological examination following modified WHO guidelines. A positive FS was defined as a finding of any neoplasia or any polyp ≥10mm. A positive FOBT was defined as at least one positive OBT test window. Participants with positive findings were given an appointment to undergo colonoscopy. Colonoscopic recovery of lesions diagnosed at FS was based on localisation, size and histology. Recently, performed colonoscopy of good quality was considered a surrogate for FS.

Outcomes

There was a significantly lower attendance rate of 63 percent in the FS and FOBT combination group compared to the FS alone group with 67 percent attendance, (p<0.01). Compliance for FOBT was 51 percent. Non-attendance was accounted for and overall 12,960/20,033 (65%) of screenees attended. On an intention-to-diagnose basis, 12 percent of screenees had neoplasia diagnosed in the FS only group and 11% in the combination FS + FOBT group. There was no difference in the diagnostic yield between the two groups of CRC or high-risk adenomas. In the FS group, 21 (0.3%) cases of CRC were diagnosed and in the FS and FOBT group 20 (0.3%) cases of CRC were diagnosed, (p=0.95). There were four CRC (20%) and eight (3.1%) high-risk adenomas diagnosed in the group with positive FOBT but negative FS, four CRC (20%) and 168 (59%) high-risk adenomas diagnosed in the group with negative FOBT but positive FS, and nine CRC (22%) and 56 (22%) high-risk adenomas diagnosed in the group with both positive FOBT and FS. In the FS group 19 percent and in the combined FOBT + FS group 21% of participants were recommended colonoscopy. There were no serious complications at FS and minor events were reported in 0.2 percent. A total of 6/2,821 colonoscopies reported perforations, all following polypectomy. The overall study adenoma prevalence rate was estimated to be 17 percent and for advanced neoplasms, 4.5 percent.

The compliance rates were significantly lower in the combined group than the FS-alone group, and those invited for combination screening could only return their FOBT on condition of accepting FS as well. The addition of FOBT reduced attendance by 4 percent. This trial used 140cm colonoscopes as the standard screening tool, and it was reported that 14 percent of screenees were examined beyond the 60cm. This extended range places limitations on the value of this study when applying to the potential diagnostic performance of FS in a screening programme. The addition of FOBT to FS screening in this study did not significantly increase diagnostic yield in the target population. On an intention to diagnose basis similar proportions of participants in both groups were diagnosed with neoplasia. There were few high-risk adenomas diagnosed (3.1%) due to positive FOBT where FS was negative, but 20 percent of CRCs were diagnosed in this way. The addition of FS to FOBT added significantly to the diagnostic yield with 20 percent of CRCs and 59 percent high-risk adenomas diagnosed. The study
methodology was robust in terms of randomisation, patient selection, sample size determination and outcomes, and was also adequately described in the paper by Bretthauer et al, (2002). The present study analysis was limited to the results of one time baseline screening. CRC incidence and mortality data have not yet been published from this trial. The authors concluded that further follow-up results are needed and will become available to provide evidence that the addition of FOBT to once-only FS will add benefit to colorectal screening programmes.

Data on CRC mortality and incidence five-years post-screening for all participants will need to be analysed. Since the last participant was screened in December 2001, results will not be published until late 2007 (personal correspondence, Professor G. Hoff, Investigator, NORCCAP Screening Trial, 8 March 2005).
Table 14. Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT screening alone

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasmussen et al. (1999)</td>
<td>RCT</td>
<td>Grade II</td>
<td>Combined FOBT + FS screening group (Hemoccult II)</td>
<td>From 10,978 general population of Funen, Denmark aged 50-75 years (husbands/wives allocated to same group). Those eligible were randomised to:</td>
<td>Status at screening between 1992 and 1995 and clinical status assessed at 31.12.1997, 26-62 months out from inclusion in programme.</td>
<td>Health outcomes: Interval CRC incidence/mortality. There were 11 CRC-related deaths in the combined screening arm and 14 in the FOBT alone arm. Incidence of CRC was the same in (38 persons) even group. At screening this was 12 and 4 persons.</td>
<td>Limitations: Not designed as a mortality study, low compliance in combined-screening trial arm, non-repeat screening and short follow-up.</td>
</tr>
<tr>
<td>(Funen-2 trial)</td>
<td></td>
<td></td>
<td>Once-only Guaiac FOBT (Hemoccult II) Non-rehydrated, diet-restricted physician not knowing FOBT result, followed with once only flexible sigmoidoscopy (FS).</td>
<td></td>
<td>Diagnostic yield of CRC and adenomas.</td>
<td></td>
<td>Author's conclusions: The addition of FS to FOBT compared with FOBT alone in a one-time screening regime is not an optimal use of resources with high rates of colonoscopy follow-up that requires further cost-benefit analysis. The diagnostic yield of neoplasia was higher with combined FS + FOBT despite poor compliance, and fewer CRCs were detected after screening compared to FOBT alone.</td>
</tr>
<tr>
<td>Funen, Denmark</td>
<td></td>
<td></td>
<td>Non-rehydrated, diet-restricted and physician not knowing FOBT result, followed with once only flexible sigmoidoscopy (FS).</td>
<td></td>
<td>CRC mortality incidence of CRC and adenomas in follow-up interval period. For screening group:</td>
<td></td>
<td>Reviewers' conclusions: Well-conducted study but limited CRC incidence/mortality statistical analysis and participant information. Limited incidence/mortality data means it is difficult to be conclusive about health outcomes and any benefit with combined screening. Although higher diagnostic yield with combined FOBT + FS screening there was a lack of compliance, non-repeat screening and high resource use in follow-up colonoscopies. The true effect of screening with FS is likely to be underestimated because of low compliance. Require more trials to better ascertain benefits/costs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No polyectomy was performed during FS.</td>
<td></td>
<td>Complete FS Rate of positive FOBT and/or FOBT. Synchronous detection Proportion undergoing colonoscopy. Control group:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A positive FS was defined as one or more of six slides positive.</td>
<td></td>
<td>Rate of positive FS and/or FS.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Those with positive FOBT and/or FS were invited for medical review and full colonoscopy with polypectomy (or DCBE if colonoscopy refused or incomplete). Those with CRC or adenomas detected were then invited to surveillance programme.</td>
<td></td>
<td>Participation of those invited for screening round.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FOBT only screening group</td>
<td></td>
<td>Number of complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Once only Guaiac FOBT (Hemoccult II) Non-rehydrated, diet-restricted</td>
<td></td>
<td>Verified diagnosis of CRC/adenoma from Funen patient database, National Cancer Registry, Danish Cancer Registry, National Death Cause Registry.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Study participants from both groups invited for screening between September 1992 and November 1995. Follow-up of clinical status as at 31&lt;sup&gt;st&lt;/sup&gt; December 1997 (24-62 months after inclusion in the screening programme).</td>
<td></td>
<td>Cause of mortality as CRC from pre-specified criteria set to assign CRC (including complications attributable to CRC) as cause of death. Investigator review of death certificates was conducted blind to participant screening status, with independent committee review if unclear whether CRC was cause of death.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 14. Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT screening alone (continued)

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasmussen et al. (2003)</td>
<td>RCT</td>
<td>Combined screening study arm from the Funen-2 RCT</td>
<td>Intervention – combined screening study arm from the Funen-2 RCT</td>
<td>Diagnostic yield of screen-detected CRC and adenomas</td>
<td>Incidence and diagnostic yield of CRC Screen-detected CRC of 9.9 per 1000 persons (95% CI 8.4-9.5) in biennial screening trial after 8 rounds and 6.6 per 1000 persons (95% CI 3.4-11.5) in the combined FOBT + FS trial, p&lt;0.017.</td>
<td>Limitations: CRC mortality data not included. Comparison of two trials initiated seven years apart. Only once combined FOBT + FS was compared to first and cumulative screening rounds of a biennial FOBT programme. No statistical differences in age/sex distribution but adjustments to samples from matching technique and pseudo-randomisation method. Low compliance, non-repeat screening, high follow-up colonoscopy and short follow-up in combined FOBT + FS screening arm. No reported data from direct screen harms.</td>
</tr>
<tr>
<td>(Funen-3 study)</td>
<td></td>
<td>Once-only Guaiac FOBT (Hemoccult II) Non-rehydrated, diet-restricted and physician not knowing FOBT result, followed with once only Flexible Sigmoidoscopy (FS).</td>
<td></td>
<td>Effect of compliance with screening on CRC detection for those from screening group who accepted specific screening rounds relative to those who refused screening.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funen, Denmark</td>
<td></td>
<td>The definitions for a positive FS and positive FOBT were previously described in Rasmussen et al. (1999) Table 14.</td>
<td></td>
<td>Risk of positive FOBT and positive FOBT and/or FS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Those with positive FOBT and/or FS were invited for medical review and full colonoscopy as previously described in Rasmussen et al., 1999 Table 14.</td>
<td></td>
<td>Proportion undergoing colonoscopy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The eighth round of the biennial FOBT screening programme study-arm from the Funen-1 RCT Guaiac FOBT (Hemoccult II) Non-rehydrated, diet-restricted</td>
<td></td>
<td>Participation of those invited for each screening round</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biennial testing (8 rounds total) presented here and ninth commenced in August 2001 in results in Table 8, Kronborg et al., 2004.</td>
<td></td>
<td>For both trials verified diagnosis of CRC/adenoma from Funen patient database, and National Patient Registry, Danish Cancer Registry, National Death Cause Registry. Verified cause of mortality as CRC from pre-specified criteria set to assign CRC, including complications attributable to CRC as cause of death. Investigator review of death certificates was conducted blind to participant screening status, with independent committee review if unclear whether CRC was cause of death.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Only screening who participated in previous rounds (and without CRC or adenomas) were reinvited.</td>
<td></td>
<td>Age matched random sampling technique to exclude those under 50 years in the biennial FOBT programme (Funen-1 RCT) to leave 25,151 persons and 4460 persons in the combined screening study-arm (Funen-2 RCT).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interventions and dates</td>
<td></td>
<td>Outcomes and verification</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Study design</td>
<td></td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Country</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER**
SECONDARY RESEARCH

The search strategy identified one relevant systematic review that considered the effectiveness of FOBT and FS combined screening. The methods and conclusions are described in Table 15. As discussed in the Methodology chapter above, this paper did not employ the same inclusion and exclusion criteria as applied in this review and the results must be interpreted with care.

This review (Walsh and Terdiman, 2003) was also included under secondary research in Chapter 3 on the effectiveness of FOBT screening on health outcomes and Chapter 5 on the effectiveness of FS screening on health outcomes. This review, published in 2003, considered the status of the evidence for colorectal cancer screening, including methods of screening that may become available in future. For FOBT and FS combined screening, only one non-RCT (Winawer et al, 1993 as cited by AHTAC 1997) considered CRC mortality outcomes with rigid sigmoidoscopy (RS) combined with FOBT, with RS only. This study is older and used rigid sigmoidoscopy which has been replaced by flexible sigmoidoscopy, and the results may not be generalisable to flexible sigmoidoscopy. The study found a colorectal cancer mortality reduction of 43 percent compared to RS only after 5-11 years follow-up. Three other RCTs (Berry et al, 1997; Rasmussen et al, 1999; Verne et al, 1998) were included in the analysis but only screening diagnostic yield outcomes were reported. These studies were also included in the current review, but only the RCT by Rasmussen et al. (1999) was included in Chapter 6 on the effectiveness of combined FOBT and FS screening. The authors’ conclusions were that there was a lack of studies evaluating the health outcomes of FOBT and FS combined screening and that the studies that were available considered single FOBT testing in addition to one-time sigmoidoscopy. Direct and indirect evidence indicated that the available tests for CRC screening were effective, but differed in their sensitivity, specificity, cost, and safety, and that the available evidence did not support choosing one test over another at the time of the review.

Several other systematic reviews are briefly commented on but not included for appraisal as they did not fully meet inclusion criteria or the studies considered in these reviews were already included.

The systematic review previously included in Chapter 3 conducted by McLeod with the Canadian Task Force on Preventive Health Care (2001a) and adapted for journal publication (McLeod, 2001b), evaluated the effectiveness of specific screening techniques for CRC in asymptomatic individuals at normal risk. The analysis referred to three of the same RCTs (Berry et al, 1997; Verne et al, 1998; Rasmussen et al, 1999) included in the review by Walsh and Terdiman (2003) above. The systematic review by McLeod (2001a) concluded that evidence from case-control studies indicated that FS may reduce the risk of CRC mortality but that the RCT studies indicated that the benefits of one or both FS and FOBT cannot be determined since these studies did not report mortality data. However, they concluded FS may be superior to FOBT in the detection of adenomas and possibly cancers more work is needed to determine the data regarding the frequency of examinations, compliance, screening FS and follow-up colonoscopy feasibility along with the significance of small polyps (<0.5 cm) identified at FS and the need for colonoscopic follow-up.

The comprehensive systematic review and associated publications by Pignone et al. (2002a, 2002b, 2002c) and linked with this the most recent US Preventive Services Task Force recommendations and rationale for CRC screening (US Preventive Services Task Force, 2002), previously included in Chapter 3 also addressed the effectiveness and accuracy of performing combined FOBT and FS. No RCTs at this time were identified with CRC mortality as an end point that had compared combined testing with either FS alone or FOBT alone (Pignone et al. 2002b). One non-randomised trial by Winawer et al. (1993), also included above (Walsh and Terdiman, 2003) showed a marginally statistically significant 43 percent reduction in CRC mortality when FOBT was added to rigid sigmoidoscopy, but these results are of questionable relevance to combined FOBT and FS screening. The same three RCTs were identified in these reviews, as in the systematic reviews above by Walsh and Terdiman (2003) and McLeod (2001). The conclusions of this review were that for each study the addition of FS to FOBT increased the identification of significant adenomas or cancer by at least a factor of two over FOBT alone and that the addition of FOBT to FS did not identify any significant lesions compared to FS alone in one-time screening. With adverse events, combined screening by FS and FOBT was equivalent to each test alone but outcomes data were not presented.
CONCLUSIONS

The Working Party on Screening for Colorectal Cancer (1998) concluded that there was a lack of RCT evidence for a mortality reduction for a national screening programme based on flexible sigmoidoscopy. However, this should be reinvestigated once new data from ongoing RCTs became available. Flexible sigmoidoscopy was considered as a potential screening tool but many issues surrounding efficacy and cost-effectiveness, and its context and employment in an asymptomatic middle-aged population remained to be resolved. The view of the working party was that New Zealand did not have the endoscopic resources to instigate a large-scale population-based FS screening programme.

Evidence for a combined FOBT and FS screening strategy compared to a strategy of either test alone was not evaluated by the Working Party but the conclusions made by the working party on a FS screening strategy are also valid for any potential combined FOBT and FS screening strategy. The evidence considered in this chapter does not support a combined screening strategy over a screening strategy involving either FOBT or FS alone, as few RCTs were available and these had limited health outcome data. Two RCTs, the Funen-2 (Rasmussen et al, 1999) and Funen-3 (Rasmussen et al, 2003) studies, comparing combined screening with FOBT screening alone were identified and appraised in this current review.

The end-point analysis (24-62 months post-screening) in the Funen-2 trial showed no difference in the CRC clinical status of participants between the two groups of screenees. There was no significant difference in CRC incidence or mortality between the two groups for screenees and non-responders alike. This one-time screening trial with a short follow-up period, though of reasonable quality, was not designed as a mortality study and focused on compliance, diagnostic yield and colonoscopic follow-up. The Funen-3 study, a derivative study from two different trials, did not evaluate CRC mortality as an end-point outcome. There was an evaluation of CRC and advanced adenoma incidence which compared cumulative FOB testing over eight biennial screening rounds to the diagnostic yield of once only combined FOBT and FS screening. This showed no significant difference between the two groups in screen, detected CRC on a screened person’s basis, but on an invited persons basis this was significantly higher in the biennial FOBT programme. There was significantly higher detection of advanced adenomas in the combined screening group. These results were limited by the comparison of two temporally separated (they were begun seven years apart) screening programmes. One was a multi round programme while the other was a single-round programme. CRC incidence and mortality data are not expected from the Norwegian NORCCAP trial (Gondal et al, 2003) until late 2007. This information may provide important data on the effectiveness of combined FOBT and FS screening compared to FS screening alone.

Both studies and also the ongoing Norwegian NORCCAP trial (Gondal et al, 2003) for which baseline screening data was available consistently showed little additional diagnostic yield from adding once-only FOBT to FS screening. There was high additional diagnostic yield of CRC and advanced adenomas from adding once-only FS to FOBT screening compared to FOBT alone. There was no difference in screen-detected CRC and high-risk adenomas for combined screening compared to FS alone. High numbers of colonoscopies were required to confirm the diagnosis of one CRC from FS screening detection compared to FOBT screening detection. Low compliance was reported for FS screening, especially with combination screening where participants were aware of their FOBT results. These studies reported no difference in complications associated with combined screening compared to either FOBT or FS screening alone.

One relevant systematic review (Walsh and Terdiman, 2003) was identified and appraised. This review included one older non-RCT that reported a mortality reduction using rigid sigmoidoscopy and FOBT but could not be generalised to flexible sigmoidoscopy. More recent RCTs were also included but these did not evaluate CRC incidence or mortality. These more recent RCTs are included in this review. The conclusion of the review by Walsh and Terdiman (2003) was that there was a lack of studies evaluating the health outcomes of FOBT and FS combined screening, and that the studies that were available considered single FOB testing in addition to one time sigmoidoscopy. The review also concluded that the evidence did not support the selection of one test over another. Other secondary research identified the same studies as did Walsh and Terdiman, and similar conclusions were made regarding the lack of available RCTs with health outcomes data to determine the benefits of combined

---

EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER
FOBT and FS screening compared to either modality alone. These reviews did conclude that FS may be superior to FOBT in terms of diagnostic yield in the identification of CRC or advanced adenomas but that important issues regarding the frequency of examinations, FS and FOBT compliance, the significance of small polyps and rationale for follow-up colonoscopy needed to be addressed.

Combined FS and FOBT as a first-line screening strategy in asymptomatic middle-aged populations is not supported by the current available evidence in the literature. There is a lack of RCTs evaluating longer-term health outcomes of CRC incidence and mortality. The RCTs considered focused primarily on screening participant compliance, the diagnostic yield of combined screening compared to either test alone and the utilisation of colonoscopy and detection in screening follow-up. The conclusions made by the Working Party on Screening for Colorectal Cancer (1998) related to flexible sigmoidoscopy screening and a combined screening strategy were not considered. However, many of the conclusions made about FS screening still hold for combined FOBT and FS screening as a first-line screening strategy. Such a strategy cannot be justified nor is it feasible, based on the available evidence. Among others, there are also unresolved issues about low compliance, polyp significance and follow-up colonoscopy protocols and endoscopic resource constraints within the New Zealand public health system.
<table>
<thead>
<tr>
<th>Source</th>
<th>Search method</th>
<th>Selection criteria</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walsh and Terdiman (2003) University of California, San Francisco, USA.</td>
<td>Search: January 1966-August 2002. Databases searched: PubMed. Also identified articles from published meta-analyses, reference lists of key articles and expert consultation. Supplemented search with hand searching and proceedings from national professional organisation meetings. Key Words: colorectal neoplasia, occult blood, sigmoidoscopy, barium enema, colonoscopy all combined with screening. Population: Asymptomatic subjects who were at average-risk for CRC.</td>
<td>For combined sigmoidoscopy and FOBT screening: One RCT examined mortality outcomes and three other RCTs and two observational studies were screening studies mostly reporting diagnostic yield and uptake at screening. Mortality Results of 12,479 subjects in a controlled trial (Winawer et al, 1993) comparing combined rigid sigmoidoscopy (RS) and FOBT with RS-only screening after 5-11 years follow-up: • Reduction in colorectal cancer mortality in combined screening group 43% compared to RS alone, p=0.053. Diagnostic yield Two RCTs (Berry et al, 1997, Rasmussen et al, 1999 Funen-2 trial) demonstrated that combination screening tests detected 4-5 times more large polyps and CRCs than FOBT alone. Another RCT (Verne et al, 1998) reported that the diagnostic yield was lower for combination screening than FS alone. A large observational study calculated that more than 600 individuals would need to be screened by immunochemical FOBT to detect one additional advanced adenoma or cancer that otherwise would have been missed by FS alone and another study indicated a small incremental diagnostic benefit with the addition of FOBT to sigmoidoscopy.</td>
<td>Not restricted to published literature. Several internal review validity processes were not reported including how study selection was undertaken, how data was extracted, or how assessment of internal validity of included studies was undertaken. Lack of studies evaluating health outcomes (CRC incidence/mortality) for combined screening programmes. Most studies examined uptake and diagnostic yield for once-only screening programmes. Authors’ conclusions There is little literature supporting combination testing and the studies available consider single FOBT testing in addition to sigmoidoscopy testing. The recommendation that all men and women aged 50 years or older undergo screening for colorectal cancer is supported by a large body of direct and indirect evidence. At present, the available evidence does not currently support choosing one test over another. Reviewers’ conclusions Lack of RCTs on combination screening to assess whether or not this reduces CRC mortality and incidence compared to sigmoidoscopy or FOBT screening alone. Only one study included mortality data and this was combination screening with rigid sigmoidoscopy. Three other RCTs were included but these were reported here with diagnostic yield results. These three studies have been included in this systematic review, with the study by Rasmussen et al. (1999) included in the primary research section of this chapter and the other two by Berry et al. (1997) and Verne et al. (1998) included in Appendix 6 on the accuracy of FOBT and FS combined screening.</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 7: Cost-effectiveness of CRC screening

PRIMARY RESEARCH STUDY DESIGNS AND QUALITY

The search identified 15 eligible primary research studies. Below is an overview of study designs and aspects of quality represented by these studies.

Study designs and quality assessments
Of the 15 eligible primary research studies, the search identified eight studies based on randomised controlled trials comparing FOBT screening with no screening. These were based on the Funen RCT, the Nottingham RCT, and the Minnesota RCT. One study, the Goteburg RCT, was also a source of data. The search also identified three studies based in part on randomised controlled trial data comparing guaiac-based FOBT with immunochemical FOBT. Four studies comparing flexible sigmoidoscopy with no screening were identified. One of these also compared flexible sigmoidoscopy plus FOBT with no screening.

The included studies were all of high quality in their design and conduct.

Methods
Earlier studies of FOBT analysed preliminary RCT outcomes and costs and used simulations to project longer-term results. Later studies have a stronger grounding in real outcomes derived from the longer running time of the RCTs and, therefore, are less reliant on simulations for the longer term estimates of cost-effectiveness.

Ten of the total 15 studies used Markov-based models, or transitional probabilistic models to project longer-term results and two used microsimulation to achieve the same objective (Microsimulation models are computer models that simulate large representative populations of individuals or families to draw conclusions that apply to an entire community. Microsimulation models used in health typically simulate representative populations over time and allow the rational comparison of competing health intervention alternatives, in a framework that captures the effects of disease interactions). One study presented a cost-effectiveness analysis based on directly observable RCT results, supplemented where necessary by values derived from the literature. Two studies were concerned with what trial results might have been if the same trial had been used in a different population: One of these directly projected trial results to the US population; the other used microsimulation to simulate the costs and effects for the Canadian population. Another used microsimulation to generate willingness-to-pay thresholds for a screening modality based on US cost data.

Data
Data for the FOBT studies were drawn primarily from the three major RCTs. However, values from the literature were used in all studies that relied on projections of longer-term results. National incidence data and other relevant data were also used in studies where trial results were projected to different populations. Data on immunochemical FOBT was drawn from the literature and, in one case, from an RCT run in Florence, Italy. Data on flexible sigmoidoscopy was principally drawn from the literature but also from a flexible sigmoidoscopy screening programme in Australia.

Interventions and comparators
Eight studies compared guaiac-based FOBT screening with no screening. The specific type of FOB test was dependent on the RCT on which the study was based. Several of these studies considered
different target age groups and different frequencies of screening in simulations than what had been used in the RCTs in order to identify the optimal screening protocol.

Three studies compared immunochemical FOBT screening with guaiac-based FOBT screening. One of these also included a comparison of non-rehydration with rehydration of the guaiac-based FOB tests (based, in part, on the Minnesota and Goteburg trials).

Four studies compared flexible sigmoidoscopy with annual FOBT. This included flexible sigmoidoscopy every five years and flexible sigmoidoscopy every 10 years. One of these also compared the combined strategy of flexible sigmoidoscopy and FOBT with FOBT alone, flexible sigmoidoscopy alone, and with no screening.

**Outcomes**

The outcome used in an economic analysis is generally designed to provide an indication of value-for-money of the intervention. In order to appreciate value for money, it is important for the outcome measure to be comparable not only with other results for the same intervention and in the same context, but also with results for different interventions and in different contexts (i.e. to be able to compare the outcome measure for CRC screening with that of breast cancer screening).

Thirteen studies used cost per life year gained/saved as an outcome measure. Two studies used cost per quality-adjusted life year (QALY) gained as an outcome measure. One study used cost per cancer detected as an outcome measure, and one study used cost per disability-adjusted life year (DALY) saved as an outcome measure.

**Study results**

The main approach, findings, and limitations of each economic evaluation are summarised below.

**Currency conversion**

Currencies in this report have been converted to New Zealand dollars (NZD) using the average of Westpac buy and sell exchange rates on 3 March 2005. This conversion is not intended to provide accurate estimates of what costs would be in New Zealand. Converting estimates expressed in foreign currencies to New Zealand dollars in a way that captures the true value being expressed in the currency of origin is problematic. However, a simple conversion using a given exchange rate can provide a better appreciation of a value expressed in a currency with which the reader may not be familiar. In this report, estimates expressed in New Zealand dollars are intended only as a rough indication of the magnitude of the monetary values to guide the reader in interpreting the results. These estimates should not be assumed to be the true costs.

FOBT COMPARED WITH NO SCREENING: APPRAISAL OF STUDIES

The search identified eight eligible primary research studies relevant to FOBT compared with no screening (see Table 16 for full details). Below is an overview of each study’s methods, outcomes, identified limitations and author conclusions.
Whynes et al. (1998) and Whynes (1999)

Methods

These studies were based on the experience of the Nottingham RCT and provided a cost-effectiveness analysis of biennial FOBT screening using guaiac-based FOBT, haemoccult and, although published separately, were based on the same methods and data. The incremental costs of participation in screening relative to the costs of diagnosing and treating patients who present symptomatically were calculated using the cost data available from the trial.

Estimates of incremental cost per quality-adjusted life year (QALY) gained as a result of screening according to the Nottingham trial protocol were generated directly from trial data up to a median follow-up of eight years. A Markov-based model with a cohort of 100,000 hypothetical subjects was used to generate estimates of longer-term incremental cost-effectiveness ratios (ICERs). The model encompassed four possibilities for disease progression: Screening asymptomatic individuals was expected to reveal the presence of

- a pre-symptomatic cancer or adenoma which, without screening, would have become clinically detected
- a slowly-developing adenoma or carcinoma which, in the absence of screening, would never have been detected in the subject’s lifetime
- adenomas which would not have progressed to carcinomas, or
- no abnormality.

Net costs would be incurred in all four scenarios, although screening may only benefit those to whom the first applies.

Investigation of positive FOBT results was assumed to follow the pattern observed in the trial of primarily colonoscopy but also some DCBE.

The UK breast cancer screening programme was used as a comparator to allow contextualisation of the results.

Data

The data for this study were derived from the Nottingham trial and supplemented, where necessary, by values from the literature. The primary cost data were derived from a series of audits of resource usage for trial subjects which were conducted for each stage of the screening and treatment process – invitation and FOBT testing; diagnosis and investigation; and, treatment and follow-up.

The estimates of life years gained as a result of screening were derived from trial data but the quality adjustment required to produce the QALY outcome measure was derived from other empirical investigations.

Results

Based on the median eight-year follow-up of the Nottingham trial at the time the analysis was conducted, cost per QALY gained as a result of CRC screening using haemoccult was approximately £5700 (1995-96 prices) (approximately NZD 15,014) for males and £5000 (1995-96 prices) (approximately NZD 13,170) for females. Longer-term estimates, which relied on modelling, took into account the lifetime benefits and costs of continued screening. These estimates were approximately £2000 (approximately NZD 5,268) per QALY gained for males and approximately £1400 (approximately NZD 3688) per QALY gained for females. For all simulations the screening of females was more cost-effective than the screening of males owing to their longer life expectancy.
Sensitivity analysis

The sensitivity analysis included investigating the effect of changes in test and investigation costs as well as in the premium resulting from early- as opposed to late-stage treatment. The analysis also considered the effect of annual rather than biennial screening due to the high rate of cases in the Nottingham and Funen-1 trials of subjects recording a negative result and subsequently presenting symptomatically prior to the next offer of a test. Poorer survival gains were also considered. Estimates were also generated for benefits being discounted at a lower rate than costs. The effects of changes in specificity and sensitivity of the FOB test were also incorporated.

Different compliance assumptions were found to make little difference to results: Increased compliance increased cancer yield and survival gains, but this came at the expense of additional detection, treatment, and follow-up costs, and the effects appeared to be compensated in the calculation of the ICERs.

Doubling of FOBT costs was found to raise the ICER by 30 percent relative to the base estimate.

High specificity of the FOB test was found to be instrumental in avoiding the high costs of investigating false positives: The cost per QALY doubled if FOBT specificity decreased by 10 percent.

Annual screening rather than biennial screening was found to increase both the cost and yield of screening compared with the biennial approach, and it was found that this had little impact on the ICER.

Limitations

The perspective of this study was that of third-party payer. Consequently, no indirect costs such as patient time and travel costs, informal carer costs, or anxiety were included. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Beyond the immediate investigation of positive FOBT results, the study did not specify the assumed nature or frequency of follow-up investigations.

Conclusions

The authors concluded that compared with estimates of the cost per QALY gained as a result of the breast cancer screening programme, which were estimated to be £3500 (approximately NZD 9219) to £6000 (approximately NZD 15,804) (1995 prices), the estimates of cost per QALY gained from CRC screening using FOBT versus no screening were low.

Methods

These studies, although published separately, were based on the same methodology and data. The studies estimated the cost-effectiveness of 60 possible CRC screening programmes using FOBT Haemoccult-II, based on the outcomes of the Funen-1 RCT. A Markov-based model was used in which core variables were estimated within the model and were input into a simulation process where the number of cancers detected per age group at each screening round was estimated for hypothetical screening programmes with varying screening intervals and target age groups.

Costs and effects were based on screening an unscreened population and, therefore, the first few years were expected to generate a higher number of detected cancers. Consequently, the screening programmes were run through the model for 36 years to obtain the estimates associated with the permanent rate of cancer detection.

Positive FOBT results were assumed to be investigated by colonoscopy or DCBE where colonoscopy failed or was not acceptable. A follow-up programme was assumed to consist of colonoscopy every three years.

Data

Data was derived from the Funen-1 RCT.

Results

The six most efficient programmes evaluated included biennial screening of 65-74 year olds, of 60-74 year-olds, and of 55-74 year olds, with cost per life year saved estimates between 17,000 DKK and 18,800 DKK (approximately NZD 4132 and NZD 4570, respectively). Also on the list of six most efficient programmes were: screening 55-74 year olds every 1.5 years, with an ICER of 20,200 DKK; and annual screening of 55-74 year-olds and 50-74 year-olds, with ICERs of 23,000 DKK and 26,000 DKK, respectively (approximately NZD 4910, NZD 5590 and NZD 6320, respectively).

Sensitivity analysis

Plausible variations were introduced for the values of test cost, colonoscopy cost, the effect of adenoma follow-up, and excess survival rate. The sensitivity of the results was also tested with regard to discounting, the inclusion of production loss, and the inclusion of future unrelated health costs.

The cost of colonoscopy and variations in the excess survival rate were found to have significant effects on the estimate of cost per life year saved: Tripling the cost of colonoscopy was found to increase the ICERs by 40-45 percent and alter the list of efficient programmes. A 1 percent decrease in the excess survival rate was found to generate a 4-4.9 percent increase in incremental costs.

Changes in the discounting method, namely not discounting benefits, resulted in more favourable results for younger target age groups.

The inclusion of future unrelated health care costs markedly affected the cost of screening (ICERs were in the range of 42,800 DKK to 63,800 DKK (approximately NZD 10,403 to NZD 15,507)) and targeting younger age groups became more favourable.

Limitations

The perspective of the study was that of third-party payer and, in the sensitivity analysis, the perspective of the public was considered but not well-described. Indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. The methods used for estimation from the societal perspective, including production loss, were unclear.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.
Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

The effect of adenoma removal on cancer incidence was not considered.

The excess survival rate of individuals whose cancer is detected by screening was assumed constant and independent of screening interval.

The disutility and potential negative health effects associated with complications of colonoscopy were not included in the analysis.

Conclusions

The authors concluded that CRC screening programmes based on FOBT Haemoccult II are cost-effective and compare favourably with screening programmes for breast and cervical cancer. In the Danish context, it would be optimal to introduce annual colorectal cancer screening of 50-74 year olds.

Helm et al. (2000)

Methods

This study was based on the experience of the three major RCTs: the Nottingham RCT, the Funen-1 RCT and the Minnesota RCT. Rather than using mathematical simulation, this study directly projected the published outcomes of the three RCTs to the US population and determined what the benefits and costs would be if each clinical trial had been applied to a relevant cohort of the US population: a cohort aged 45-75 years in 1997 (approximately 73.3 million) for the Nottingham and Funen-1 trial results; and a cohort aged 50-80 years in 1997 (approximately 61.7 million) for the Minnesota trial results. Compliance rates observed in the trials were applied, except in the case of projecting the Minnesota results where a lower compliance rate of 60 percent was used.

Three outcomes were considered: The number of cancers detected, the number of lives saved, and the number of life-years saved.

It was assumed that positive FOBT results are investigated primarily by colonoscopy and by DCBE where colonoscopy cannot be completed, according to the experience in the Funen-1 trial. Follow-up colonoscopy was assumed to take place three years after finding a polyp.

Data

Test performance data were derived primarily from the Nottingham, Funen-1, and Minnesota trials. Incidence data were from US sources. Cost data were based on Medicare and large Health Maintenance Organisation data.

Results

Estimates of cost per life year saved were: USD 2500 (approximately NZD 3440) for screening based on the Nottingham protocol; USD 2700 (approximately NZD 3716) for screening based on the Funen-1 protocol; and USD 20,500 (approximately NZD 28,212) for screening based on the Minnesota protocol. The high estimate for the Minnesota-based result is possibly explained by the smaller survival benefit associated with the trial’s ‘healthy volunteer’ recruitment and the practice of FOBT rehydration, which increased the number of false positive results and generated substantial numbers of unnecessary endoscopic investigations.

Sensitivity analysis

Applying plausible variations to key parameters yielded ranges of cost-effectiveness ratios: USD 1300 to USD 4200 (approximately NZD 1789 to NZD 5780) per life year saved based on the Nottingham results; USD 1600 to USD 4200 (approximately NZD 2202 to NZD 5780) per life year saved based on
the Funen-1 results; and, USD 11,400 to USD 32,500 (approximately NZD 15,688 to NZD 44,726) per life year saved based on the Minnesota results.

Limitations

The perspective of the study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Life years saved were not adjusted for quality of life.

The lack of simulated results meant that outcomes beyond the 10-year follow-up used could not be accounted for and neither could the possible benefits that arise from incidental polypectomy at the time of colonoscopy.

Conclusions

The authors concluded that, although the effectiveness of FOBT in detecting cancers and saving lives is limited, it was comparable to the effectiveness of accepted methods of screening for other kinds of cancer and its costs were relatively low. Overall, the cost-effectiveness of screening for CRC with FOBT versus no screening was favourable, regardless of the RCT on which estimates were based and regardless of whether specimens were rehydrated, as in the Minnesota trial.

Flanagan et al. (2003)

Methods

This study used a micro-simulation model adapted to simulate biennial CRC screening using FOBT on 67 percent of individuals aged 50-74 years in the year 2000. Estimates were generated for two cohorts: A fixed cohort for the year 2000, which allowed the simulation of clinical trial conditions; and dynamic cohorts for the period 2000-2024, which allowed the longer-term effects and potential impact on resources of a population-based CRC screening programme using unrehydrated Haemoccult II to be estimated. The model was validated against the evidence from the Funen-1 trial before being applied to the Canadian population to generate estimates of cost per life year saved.

It was assumed that all positive FOBT results are investigated by colonoscopy. Negative colonoscopy would then exempt subjects from screening for 10 years, provided no polyps were found. Follow-up colonoscopies would be performed at three, five, and 10-year intervals if polyps were found. Polyp removal was assumed to have no impact on the incidence of CRC, since none was observed in the 10-year follow-up of the trials.

Projected results over 25 years also accounted for the effect of the ageing baby boom generation.

Data

Canadian data and RCT results for sensitivity, specificity, participation, incidence, staging, progression of disease, and mortality were used. Cost data were derived from Canadian health system costs.

Results

CRC screening was found to result in a 10-year mortality reduction of 16.7 percent and a 15 percent increase in the demand for colonoscopy. Cost per life year saved was estimated at CDN 11,907 (approximately NZD 13,216).
Sensitivity analysis

Applying plausible variations to key parameters yielded modest variations on the base estimate of cost per life year saved. The ratio remained favourable under high-cost scenarios.

Annual screening would be associated with a cost per life year saved of approximately CDN 13,497 (approximately NZD 14,981).

High compliance, as observed in the Funen-1 trial, was found to be pivotal in generating favourable estimates of cost per life year saved.

Limitations

The perspective of the study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

The analysis did not include potential reduction in CRC associated with polyp removal, inclusion of which may result in a more favourable cost-effectiveness ratio.

Conclusions

The authors concluded that screening for colorectal cancer with FOBT followed by colonoscopy for subjects with positive FOB test results was cost-effective for the Canadian scenarios simulated by the model. The potential cost-effectiveness of the screening programme would, however, depend on reaching the level of compliance observed in the trial. It was noted that participation in organised breast screening programmes in Canada was low relative to the level of participation assumed for the CRC screening simulation.

Whynes (2004)

Methods

This study was based on the experience of the Nottingham RCT and provided a cost-effectiveness analysis of FOBT screening using guaiac-based FOBT three-day and six-day Haemoccult II and three-day Fecatwin/Feca EIA (Feca), the tests which were used in the Nottingham RCT. The incremental costs of participation in screening relative to the costs of diagnosing and treating patients who present symptomatically were calculated by combining the results of a comprehensive audit of resource use on the part of subjects within the trial and previously established unit costs for each of the procedures involved. The total cost of detecting cancers by screening was the sum of the products of resource use and resource costs as observed in the resource use patterns of trial subjects. Each resource type was subject to detailed cost analysis during the trial. Costs were adjusted to 2002 prices using the GDP deflator in the overall cost calculations. Life expectancy gains were estimated from a survival analysis of those trial subjects who had been diagnosed with cancer. It was assumed that all positive FOBT results were investigated by colonoscopy.

Data

The main source of both clinical and cost data was the Nottingham trial.
Results

Estimates of cost per cancer detected were presented for three-day Feca; three-day Feca and Haemoccult; three-day Haemoccult; and, 6-day Haemoccult. Costs ranged from £3,113 (approximately NZD 8,200) (2002 prices) per cancer detected for 3-day Haemoccult to £11,924 (approximately NZD 31,408) (2002 prices) per cancer detected for the combined use of Feca and Haemoccult.

Overall, a screening programme based on the Nottingham trial protocol was estimated to have an incremental cost-effectiveness ratio of £1584 (approximately NZD 4172) (2002 prices) per life year gained. Using Monte Carlo simulation, a 95 percent confidence interval of £717 to £8612 (approximately NZD 1889 to NZD 22,684) per life year gained was generated for this estimate.

Sensitivity analysis

The results were found to be relatively insensitive to plausible variations in the assumed discount rate for costs but more sensitive to variations in the discount rate for benefits. Discounting benefits by the same rate as costs raised the ICER by 77.4 percent, relative to the base estimate which was based on 6.0 percent per annum discounting of costs and 2.0 percent per annum discounting of benefits.

Using the highest Kaplan-Meier survival estimate resulted in a 23.3 percent decrease in the ICER relative to the base estimate.

Doubling testing, investigation, and treatment costs resulted in increases in the ICER of 59.6, 27.5, and 12.9 percent respectively.

Limitations

The perspective of this study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

It was not clear to what extent the cost data from the Nottingham trial allowed for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Costs in this study were discounted at 6.0 percent p.a. while benefits were discounted at 2.0 percent p.a., which resulted in favourable results compared with other studies, where costs and benefits have generally been discounted at the same rate.

Beyond the immediate investigation of positive FOBT results, the study did not specify the assumed nature or frequency of follow-up investigations.

Conclusions

The authors concluded that the most important contribution of this study was the presentation of the longest-term results of the Nottingham trial estimates which did not rely on modelling.

The authors also concluded that, overall, the cost-effectiveness of screening for CRC with FOBT versus no screening was favourable. Although, the experience of the Nottingham trial indicated that the number of cases detected through FOBT screening and colonoscopic investigation was relatively small compared to the number of subjects invited for screening, the survival gains were sufficiently high and costs sufficiently low to generate a low cost-effectiveness ratio overall.

Specifically, of all the reported trial phases, those associated with Feca and with re-invitation generated the highest costs per cancer detected due to low specificity of the Feca test and low compliance in the re-invitation cases.

Furthermore, there was no evidence of increasing costs per detection as the number of tests taken increased. Indeed, the costs per detection for the fourth and fifth test were the lowest estimated. To some degree, this is explained by the fact that subjects completing more tests were older.
The ICER for FOBT screening in Nottingham was lower than the equivalent ratio for the national breast cancer screening programme.

**Stone et al. (2004)**

**Methods**

This study evaluated CRC screening in Australia of 55-69 year olds through the use of biennial guaiac-based FOBT (based on meta-analysis of the properties of the FOBTs used in the Nottingham and Funen-1 trials) compared with the status quo of minimal opportunistic screening. The analysis employed a Markov-based model, which assumed a hypothetical nationally-coordinated screening programme based on the average-risk Australian population of 1996. It was assumed that the programme was in steady-state, in order to provide estimates of ongoing annual costs and exclude the higher implementation costs, as well as increased detection of cancer associated with the introduction of a screening programme.

Extension to younger and older age groups was evaluated using marginal analysis.

The health benefit estimated was disability-adjusted life years (DALYs).

It was assumed that a screening programme reduces mortality by 14 percent but does not reduce CRC incidence.

**Data**

The performance of the test was based on a meta-analysis of the Nottingham RCT, the Funen-1 RCT, and the Goteburg RCT. The Minnesota RCT was excluded due to the high positivity rate of the FOBT in this trial.

National data were used for incidence, mortality, age of onset, mean survival time for each age group, and duration of each phase of the model. Data from the South Australian Cancer Registry was used to estimate the stage distribution and five-year survival rates. The distribution of patients across stages was close to the combined distributions in the control groups from the four RCTs.

Cost data were derived from Australian health services expenditure data.

**Results**

A screening programme based on screening subjects aged 55-69 years would be associated with an incremental cost-effectiveness ratio of $A17,000 (approximately NZD 18,310) per DALY. Extension of the programme to 70-74 year olds would be more cost effective than extension to 50-54 year olds.

**Sensitivity analysis**

Multi-way sensitivity analysis was performed and included uncertainty distributions based on a combination of the reported confidence intervals, the range of reported values in the literature, and expert opinion of the range of likely values in the Australian context.

The major influences on the uncertainty of the health benefits were the size of the mortality reduction and the participation rate.

The possible range for mortality reduction resulted in a range of ICERs from AUD 13,000 per DALY gained to AUD 52,000 (approximately NZD 14,002 to NZD 56,007) per DALY gained.

Annual screening would be associated with an ICER of AUD 20,000 (approximately NZD 21,541) per DALY gained.
Limitations

The perspective of the study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

It does not appear that the cost data allowed for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis. Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Sensitivity analysis was not well described.

Conclusions

The authors concluded that, in Australia, the optimal cost-effectiveness would be reached by a screening programme starting at age 55, rather than the recommended 50. If less than AUD 50,000 (approximately NZD 53,853) per DALY is used as the yardstick for acceptable cost-effectiveness, then a biennial population screening programme using FOBT provides value-for-money, particularly if screening starts at the age of 55 rather than the most commonly advocated starting age of 50.

FOBT compared with no screening: Summary

The studies reviewed above were of high quality and captured most of the costs associated with a CRC screening programme based on biennial or annual FOBT with positive results investigated primarily by colonoscopy. The results suggest that such a screening programme based on biennial or annual FOBT would likely be cost-effective by commonly accepted cost-effectiveness thresholds. The cost-effectiveness of such a programme also would compare favourably with the cost-effectiveness of breast cancer and cervical cancer screening programmes, which are currently funded in many countries.

It should be noted, however, that even the most complete accounting for costs does not fully account for programme-related expenses, so the true costs of a CRC screening programme based on FOBT are likely to be somewhat higher than the estimates presented. It is also unclear whether such costs are included in the estimates of cost-effectiveness for breast and cervical cancer screening programmes, with which CRC screening is often compared. The estimates presented by the studies, however, are frequently sufficiently low (mostly less than NZD 20,000) that there is a wide margin of error in which unaccounted costs could potentially be included without pushing the cost-effectiveness of such a programme over the commonly accepted cost-effectiveness threshold of $50,000. Nevertheless, no study has been able to demonstrate this.

Many studies found that the pivotal factor in achieving a favourable cost-effectiveness ratio for this type of screening programme is likely to be the participation rate. Furthermore, to the extent that clinical data is based on inaccurate estimates and assumptions regarding the natural history of CRC and other clinical factors, estimates of cancers detected and prevented as well as mortality reductions may be biased. This uncertainty introduces a potential error of unknown magnitude to cost-effectiveness estimates.

GUAIAC-BASED FOBT COMPARED WITH IMMUNOCHEMICAL FOBT: APPRAISAL OF STUDIES

The search identified three eligible primary research studies relevant to guaiac-based FOBT compared with immunochemical FOBT. Below is an overview of each study’s methods, outcomes, identified limitations and author conclusions (see Table 17 for full details).

Methods

This study used a Markov-type model to compare the cost-effectiveness of the unhydrated Haemoccult II test based on the experience of the Funen-1 trial with the potential cost-effectiveness of alternative FOB tests (rehydrated Haemoccult II, Hemeselect, and Haemoccult II Sensa), as if these FOBTs were to be used in a similar population-screening programme. The cost-effectiveness of screening 55-74 year olds at one- and two-year intervals, as well as screening 50-74 year-olds annually, was estimated.

In order to compare the sensitivity of the tests, the investigator chose not to use the results of studies in which the sensitivity of a test was dependent on the detectable preclinical phase (DPCP) of the diagnostic test or studies based on cancer patients or high-risk subjects. Studies used to establish sensitivity were based on follow-up periods of one to two years, which allowed comparison with the estimates derived from the Funen-1 trial.

Data

Data on unrehydrated Haemoccult II was derived from the Funen-1 RCT. Data on the rehydrated Haemoccult II was derived from the Minnesota RCT and the Göteborg trial. Data on HemeSelect and Haemoccult II Sensa was derived from the literature.

Sources of cost data were not well described.

Results

Higher sensitivities of the rehydrated H-II test, the Hemoccult Sensa test, and the HemeSelect test were at a cost of lower specificity. The most cost-effective screening programmes were:

- biennial screening of 55-74 year olds using unrehydrated Haemoccult II, at a cost per life year gained of 17,500 DKK (approximately NZD 4254)
- annual screening of 55-74 year olds using unrehydrated Haemoccult II, at a cost per life year gained of 30,000 DKK (approximately NZD 7292)
- annual screening of 50-74 year olds using unrehydrated Haemoccult II, at a cost per life year gained of 39,000 DKK (approximately NZD 9479)
- annual screening of 50-74 year olds using HemeSelect, at a cost per life year gained of 71,300 DKK (approximately NZD 17,330)
- annual screening of 50-74 year olds using rehydrated Haemoccult II, at a cost per life year gained of 138,100 DKK (approximately NZD 33,567).

Sensitivity analysis

The major source of uncertainty lay around the cost-effectiveness of the rehydrated Haemoccult II test. If the Minnesota trial results were used rather than the Göteborg results, from which the base estimates were derived, the rehydrated Haemoccult II test was more cost-effective than the alternative tests.

Limitations

The perspective of the study was that of third-party payer. A lack of detail on cost data means it is not clear whether the cost data allowed for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis. Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.
The lack of simulated results meant that outcomes beyond the 10-year follow-up used could not be accounted for and neither could the possible benefits that arose from incidental polypectomy at the time of colonoscopy.

The sensitivity analysis was not well described.

Conclusions

The authors concluded that the unrehydrated Haemoccult II test was the most cost-effective option. The rehydrated H-II test may be an alternative but evidence of the test’s characteristics remained unclear. Use of the immunochemical FOBT, HemeSelect, would result in a significantly higher cost per life year saved.

Van Ballegooijen et al. (2003)

Methods

This study compared the cost-effectiveness of the guaiac-based FOBTs Haemoccult II and HemoccultSENSA with a hypothetical immunochemical FOBT assumed to have comparable sensitivity to HemoccultSENSA but with higher specificity. The study used a microsimulation model, which simulates the natural history of the adenoma-carcinoma sequence in a population with and without colorectal cancer screening. The model simulated a screening programme based on annual FOBT for approximately 72,000 individuals aged 65-79, over 30 years of screening.

The report estimated the cost-effectiveness of the immunochemical FOBT with test performance parameters that were equivalent or better than those associated with the guaiac FOBT. The report also generated the threshold payment for the immunological FOBT relative to the guaiac test that would result in equivalent cost-effectiveness.

In the base case it was assumed that the immunochemical FOBT had a sensitivity comparable to HaemoccultSENSA but a higher specificity; 100 per cent compliance was assumed.

The perspective was that of third-party payer.

Data

Data for this study were mainly derived from US. studies including the Minnesota RCT. Costs were based on US Medicare data.

Results

The cost-effectiveness ratios for all FOBTs were favourable, with the least favourable ratio generated for HemoccultSENSA (due to its low specificity): The CE ratio was USD 1071 (approximately NZD 1474) per life year saved for Haemoccult II; USD 4500 (approximately NZD 6193) per life year saved for IFOBTs with 98 percent specificity; and, USD 5827 (approximately NZD 8019) per life year saved for Haemoccult SENSA.

At payment levels of USD 28 (approximately NZD 39) for the IFOBT and USD 4.50 (approximately NZD 6) for Haemoccult II, the incremental cost-effectiveness ratio for IFOBT was USD 11,000 (approximately NZD 15,138) per additional life year saved if specificity was 98 per cent and USD 21,000 (approximately NZD 28,900) per life year saved if specificity was 95 percent.

Sensitivity analysis

Using HaemoccultSENSA as the base case rather than Haemoccult II, resulted in more favourable ICERs for the immunochemical FOBT, particularly if the IFOBT was assumed to have the higher level of specificity (98 percent).
Limitations

The major limitation of this study is the use of a hypothetical immunochemical FOBT with properties that do not currently exist amongst immunochemical FOBTs. This approach severely limits the practical use of the study. However, the study does show how future improvements in the specificity of IFOBTs, assuming such improvements can be made, may be associated with significant improvements in cost-effectiveness.

The perspective of the study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

Conclusions

The authors concluded that FOBTs are a cost-effective intervention for reducing colorectal cancer incidence and mortality. Guaiac FOBT was the most cost-effective. Comparable cost-effectiveness for immunochemical FOBTs may be reached if an IFOBT maintained the high specificity of Haemoccult II (98 percent) and had an increased sensitivity for CRC of 70 percent, over that of Haemoccult II (40 percent). In this case, the unit cost of USD 13 (approximately NZD 18) for the IFOBT would result in as comparable a cost-effectiveness ratio as for Haemoccult II, at a unit cost of USD 4.50 (approximately NZD 6). However, evidence of the relative specificity and sensitivity of IFOBTs in comparison to Haemoccult II and HaemoccultSENSA is sparse and highly uncertain.

Berchi et al. (2004)

Methods

This study compared the cost-effectiveness of the guaiac-based FOBT Haemoccult with the cost-effectiveness of immunochemical FOBT Magstream. The model used to simulate the costs and effects of the tests was Markov-based and estimated the costs and effectiveness of successive biennial CRC screening campaigns on a hypothetical cohort of 165,000 subjects aged 50-74 years screened over 20 years with 43.7 percent compliance, the rate observed in a screening programme run in Calvados, France.

The perspective was that of third-party payer.

Data

The epidemiological and cost data for this study were based on a screening programme run in Calvados, France. The data on characteristics of the guaiac FOBTs were based on the Funen-1 RCT. The data on the characteristics of the immunochemical test were derived from Zappa et al. (2001), based on a trial in Florence, Italy.

Cost data were derived from the Calvados experiment and included the costs of organising and managing the campaign as well as testing, follow-up, and treatment costs.

Results

The incremental cost-effectiveness of substituting Magstream for Haemoccult was estimated to be 7458 (approximately NZD 13,471) euros per life year saved after 10 years of screening and 2980 (approximately NZD 5382) euros per life year saved after 20 years of screening.
Sensitivity analysis

Incremental cost-effectiveness ratios were found to be positively and significantly correlated with compliance rates, suggesting that achieving targeted compliance is pivotal to cost-effectiveness.

For a given specificity, cost-effectiveness ratios were negatively correlated to the sensitivity of the immunologic FOBT. For a given sensitivity, cost-effectiveness ratios were negatively correlated with the specificity of the immunochemical FOBT.

Cost-effectiveness ratios were also found to be very sensitive to variations in assumptions regarding the natural history of CRC and the cost of colonoscopy.

Limitations

The perspective of the study was that of third-party payer and, as such, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. Also not included were the effects on general consumption and productivity, which would be relevant from a societal perspective.

It was not clear whether the cost data allowed for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis. However, because the cost data were derived from an actual screening programme, these are likely to be the most comprehensive programme cost data available from these studies.

The model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Life years saved were not adjusted for quality of life.

Undiscounted results were presented but were not useful. Discounted results involve only discounting costs and not effects, resulting in more favourable ICERs than if costs and effects were discounted at the same rate or effects are discounted at any positive rate.

Conclusions

Using an immunochemical test like the Magstream FOBT could provide a substantial increase in sensitivity over the Haemoccult FOBT and, as a result, increase the effectiveness of CRC screening at a reasonable cost to society.

**Guaiac-based FOBT compared with immunochemical FOBT: Conclusions**

The studies reviewed above were of high quality but noted the scarcity of evidence on immunochemical FOBT as a weakness. The studies captured most of the costs associated with CRC screening programmes based on the two types of FOBT. The results suggest that such a screening programme based on guaiac FOBT is likely to be more cost-effective than a screening programme based on immunochemical FOBT, but that there may be some potential for immunochemical FOBT to save additional lives for a reasonable incremental cost over that of guaiac FOBT.

It should be noted that even the most complete accounting for costs does not fully account for programme related expenses so the true costs of a CRC screening programme based on annual FOBT are likely to be somewhat higher, than the estimates presented.

Most importantly, however, the evidence on immunochemical FOBT is weak and there is, therefore, considerable uncertainty surrounding any estimate of its cost-effectiveness.
FLEXIBLE SIGMOIDOSCOPY COMPARED WITH NO SCREENING: APPRAISAL OF STUDIES

The search identified four eligible primary research studies relevant to flexible sigmoidoscopy compared with no screening. Below is an overview of each study’s methods, outcomes, identified limitations and author conclusions (see Table 18 for full details).

Frazier et al. (2000)

Methods

This study compared the cost-effectiveness of five-yearly and 10-yearly flexible sigmoidoscopy with a strategy of no screening. The analysis was based on a Markov-based model, which produced cost-effectiveness results for the screening of hypothetical subjects representative of the 50-year old U.S. population at average risk for CRC. The model assumed that a polypectomy would be performed during the procedure and included two possibilities for follow-up: Under one scenario, only persons found to have a high-risk polyp were referred for follow-up colonoscopy; under the other scenario, persons found to have any adenomatous polyp, regardless of size or histology, were referred for follow-up colonoscopy.

The analysis was based on screening white males, with 60 percent compliance.

Data

The data for this study were based on US figures obtained from the literature, including the published results of the Minnesota RCT, and also the Surveillance Epidemiology and End Results (SEER). Cost data was obtained from a large Health Maintenance Organization (HMO).

Results

Compared with no screening, screening with flexible sigmoidoscopy every 10 years (regardless of follow-up regime) resulted in incremental cost-effectiveness ratios of less than USD 17,000 (approximately NZD 23,295) per life year gained (USD 15,800 (approximately NZD 21,744) associated with the less aggressive approach to follow-up, and USD 16,100 (approximately NZD 22,156) associated with the more aggressive approach to follow-up). The more aggressive approach to follow-up, however, was associated with a substantial increase in effectiveness (2.8 days of life gained, compared with 1.6 for the less aggressive approach). Screening with flexible sigmoidoscopy every five years was dominated by the 10-year interval.

Sensitivity analysis

Sensitivity analysis showed that under more optimistic, but still plausible, scenarios it could be less costly to screen than not to screen, due to averted cancer treatment costs: If the cost of sigmoidoscopy were 15 percent less than assumed (USD 279 (approximately NZD 384)) or the cost of cancer treatment were 20 percent higher than assumed (USD 22,000 to USD 58,300 (approximately NZD 30,276 to NZD 80,231)), then screening was cost-saving. Screening would also be cost-saving if the incidence of polyps at 50 years of age was 20 percent higher than estimated (the incidence of polyps estimated for the base case was 21 percent).

Limitations

The study claimed to be from a societal perspective although only costs to the health care system were included. The perspective was equivalent to those studies which assume a publicly-funded health care system and take the perspective of third-party payer. As a result, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. The study also did not include the effect on general consumption and productivity, which would be relevant from a societal perspective.
The cost data did not allow for programme costs such as the costs of health promotion, recall systems and administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Life years saved were not adjusted for quality of life.

The results were presented for white males only. Because of increased life expectancy (white females) or increased cancer mortality (blacks), CRC screening was simply reported to be more cost-effective for other groups.

Conclusions

The authors concluded that screening for CRC, even in a context of low compliance, was expected to significantly reduce mortality at a cost that is comparable to other cancer screening programmes. Nevertheless, compliance significantly affected the cost-effectiveness ratios. Overall, flexible sigmoidoscopy every 10 years was preferred to flexible sigmoidoscopy every five years as a cost-effective CRC screening strategy and its cost-effectiveness was not significantly changed by adding a relatively aggressive follow-up strategy based on colonoscopy.

Sonnenberg et al. (2000)

Methods

This study compared the cost-effectiveness of flexible sigmoidoscopy every five years with a strategy of no screening, with annual FOBT, and with colonoscopy every 10 years. The analysis was based on a Markov-based model, which produced cost-effectiveness results for the screening of 100,000 hypothetical subjects of 50 years of age. Positive results of FOBT and flexible sigmoidoscopy were assumed to be investigated by colonoscopy. After polypectomy, colonoscopy was assumed to be repeated every three years until no polyps were found.

Data

Clinical data were derived from the literature, including the Nottingham and Funen-1 RCTs as well as large autopsy studies. Transition rates were also based on US vital statistics and cancer statistics. Cost data were derived from Medicare reimbursement data.

Results

Compared with no screening, screening with flexible sigmoidoscopy every five years resulted in an incremental cost-effectiveness ratio of USD 36,509 (approximately NZD 50,243) per life year saved.

Compared with annual FOBT, screening every five years with flexible sigmoidoscopy resulted in an incremental cost-effectiveness ratio of USD 65,704 (approximately NZD 90,420) per life year saved.

Overall, screening every five years with flexible sigmoidoscopy was dominated by annual FOBT and by colonoscopy every 10 years.

Sensitivity analysis

Sensitivity analysis involved testing the sensitivity of results mainly for the two more favourable strategies of annual FOBT and colonoscopy every 10 years. Varying the interval of flexible sigmoidoscopy to 10-yearly, however, was found to be associated with lower costs per life year saved. It remained the case, however, that flexible sigmoidoscopy would be dominated as a CRC screening strategy by annual FOBT and colonoscopy every 10 years.

Limitations

The study’s main focus was on colonoscopy and, therefore, results on flexible sigmoidoscopy are not well detailed. The perspective was that of third-party payer. As a result, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates.
The study also did not include the effect on general consumption and productivity, which would be relevant from a societal perspective.

Because cost data were based on Medicare reimbursement data, they did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

Conclusions

The authors concluded that screening for CRC using flexible sigmoidoscopy is associated with a favourable (below USD 50,000 (approximately NZD 68,809) per life year saved) cost-effectiveness ratio, especially if the strategy is flexible sigmoidoscopy every 10 years rather than every five years. However, flexible sigmoidoscopy at both screening intervals is dominated in terms of cost-effectiveness by both annual FOBT and colonoscopy every 10 years.

Loeve et al. (2000)

Methods

This study estimated the costs and effects of endoscopic screening for CRC, including flexible sigmoidoscopy every five years and colonoscopy every 10 years, compared with no screening. Results were generated by a microsimulation model previously used in the evaluation of breast and cervical cancer screening. It was assumed that all positive flexible sigmoidoscopy tests would be followed by colonoscopy and that all lesions would be removed within a short time. Persons with lesions of 6mm or greater would be followed-up by colonoscopy every five years until no lesions were found. Sensitivity of flexible sigmoidoscopy was assumed to be the same as that of colonoscopy (85 percent for adenomas of six to nine millimetres and 95 percent for adenomas greater than or equal to ten millimetres and cancers) for lesions within reach, except in the case of adenomas of less than or equal to 5mm where the 80 percent sensitivity of colonoscopy was reduced to 75 percent sensitivity for flexible sigmoidoscopy. Estimates were based on screening by flexible sigmoidoscopy every five years from age 50-75 years.

Data

The prevalence of adenomas was based on values derived from autopsy and colonoscopy studies. Stage-specific survival data were derived from the Surveillance, Epidemiology, and End Results registry data from 1975 to 1993. Other clinical data were derived from the literature, including the Nottingham, Funen-1 and Minnesota RCTs. Cost data were based on published estimates and included values for testing, treatment, follow-up and surveillance.

Results

Screening for CRC by flexible sigmoidoscopy every five years was found to be associated with savings in treatment costs compared with a strategy of no screening from the fifth year of the screening programme. If the screening of a cohort of patients ended at 30 years and total costs were estimated, there would be no savings. However, by five years after the end of screening, the savings associated with lower treatment costs would be sufficient to allow for break-even when weighed against the costs of the programme for the cohort. If the programme were to continue indefinitely, break-even would occur at the 44-year point.

Sensitivity analysis

Sensitivity analysis was used to test the sensitivity of results to variations in uncertain parameters including: adenoma-cancer sequence, the mean and variance of dwelling time, the cost of screening, and the cost of surveillance. Results of the sensitivity analysis showed that estimated savings
associated with flexible sigmoidoscopy do not occur under plausible higher cost scenarios, with only 55 percent of induced costs compensated for by induced savings. Assumption variations regarding the adenoma-cancer sequence and dwelling time still showed potential for savings over a long-term screening programme, although these could be considerably less than estimated in the base case.

Limitations

The perspective was that of third-party payer. As a result, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. The study also did not include the effect on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems and administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide as well as training for staff.

Life years saved were not adjusted for quality of life.

Conclusions

The authors concluded that in screening for CRC, costs are induced many years before reduced costs of treatment are observed and discounting future savings makes the screening programme even less favourable. As a result, discounted savings must be considerably larger than costs. Endoscopic CRC screening by colonoscopy every 10 years or flexible sigmoidoscopy every five years may be one instance of secondary prevention where discounted savings are, in fact, greater than costs, resulting in a potential for net savings associated with a screening programme for CRC based on endoscopic screening. Reductions in test costs or improvements in test performance will increase net savings, whereas reductions in treatment costs will reduce net savings.

O'Leary et al. (2004)

Methods

Based on the published results of an Australian flexible sigmoidoscopy screening programme, this study compared the cost-effectiveness of flexible sigmoidoscopy every 10 years with annual and biennial rehydrated FOBT Hemoccult, with colonoscopy every 10 years, and with no screening. The analysis was based on a Markov-type model, which produced cost-effectiveness results for the screening of hypothetical asymptomatic subjects aged 55-64 years. The analysis compared outcomes and costs over 10 years from the perspective of third-party payer. Incremental cost-effectiveness of each screening strategy was derived compared with a strategy of no screening. The validity of the model with regards to the natural history of CRC was tested by comparing the projected outcomes with CRC incidence and mortality recorded by the Western Australian Cancer Registry. The analysis assumed that 42 percent of the eligible population would comply with screening by flexible sigmoidoscopy and that 60 percent would comply with FOBT.

Data

The clinical data for this study were based on the literature, including the results of an Australian flexible sigmoidoscopy screening programme, the Nottingham RCT and the Minnesota RCT, as well as a community-based project involving flexible sigmoidoscopy. Cost data for the cost of the FOBT test, pathology examinations, and medical attendance were approximated by the respective fees in the Medicare Benefits Schedule. The cost of the screening tests, colonoscopy, flexible sigmoidoscopy, and treatment of cancer were based on a survey of endoscopic centres in Melbourne.

Results

Compared with no screening, screening with flexible sigmoidoscopy every 10 years was found to be associated with an incremental cost-effectiveness ratio of AUD 16,801 (approximately NZD 18,096)
per life year saved. Biennial and annual FOBT were found to be less cost-effective when compared with no screening, with estimated cost-effectiveness ratios of AUD 41,183 and AUD 46,900 (approximately NZD 44,357 and NZD 50,514) per life year saved, respectively.

Sensitivity analysis

Sensitivity analysis was used to test the sensitivity of results to variations in uncertain parameters, including cost parameters and assumptions about the natural history of CRC. Costs were varied by 20 percent. Higher compliance rates were tested for their effect on cost-effectiveness, but lower compliance rates were not considered.

The sensitivity analysis indicated that the cost-effectiveness of flexible sigmoidoscopy is particularly sensitive to assumptions about the progression to cancer of adenomas of at least 10mm. Base case figures included a 5.0 percent probability for this variable and this was varied to 2.0 percent and ten percent in the sensitivity analysis. This variation yielded a range of incremental cost-effectiveness ratios of AUD 8762 (approximately NZD 9437) for a 10 percent probability to AUD 32,362 (approximately NZD 34,856) for a two percent probability.

Limitations

The study was from a third-party payer perspective and, as a result, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. The study also did not include the effect on general consumption and productivity, which would be relevant from a societal perspective.

The study was unclear as to whether the cost data allowed for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

Conclusions

The authors concluded that screening for CRC using flexible sigmoidoscopy is cost-effective compared with no screening, and is associated with a more favourable cost-effectiveness ratio than annual or biennial rehydrated FOBT compared with no screening. Flexible sigmoidoscopy remained the most favourable strategy for CRC screening amongst the strategies considered under a range of sensitivity analyses.

Flexible sigmoidoscopy compared with no screening: Conclusions

The studies reviewed above were of high quality and captured most of the costs associated with a CRC screening programme based on flexible sigmoidoscopy every five or 10 years. The results suggest that such a screening programme would likely be cost-effective by commonly accepted cost-effectiveness thresholds.

It should be noted that even the most complete accounting for costs does not fully account for programme-related expenses so the true costs of a CRC screening programme based on annual FOBT are likely to be somewhat higher than the estimates presented. Some studies suggest that a small amount of cost savings may be possible with a CRC screening programme based on flexible sigmoidoscopy. However, the addition of unaccounted costs is likely to significantly reduce or even eliminate this possibility.

Furthermore, the evidence on effectiveness of flexible sigmoidoscopy is scarce and, therefore, the results are heavily reliant on assumptions of sensitivity and specificity.
The studies also note that the participation rate may be a pivotal factor in achieving a favourable cost-effectiveness ratio.

**FLEXIBLE SIGMOIDOSCOPY PLUS FOBT COMPARED WITH FLEXIBLE SIGMOIDOSCOPY OR FOBT OR NO SCREENING: APPRAISAL OF STUDIES**

The search identified one eligible primary research study relevant to flexible sigmoidoscopy plus FOBT compared with flexible sigmoidoscopy or FOBT or no screening. Below is an overview of the study’s methods, outcomes, identified limitations and author conclusions (see Table 19 for full details).

**Frazier et al. (2000)**

**Methods**

This study compared the cost-effectiveness of flexible sigmoidoscopy every five or 10 years plus annual FOBT to, among others, a strategy of no screening, a strategy of flexible sigmoidoscopy alone every five or 10 years, and a strategy of annual FOBT alone. The analysis was based on a Markov-type model, which produced cost-effectiveness results for the screening of hypothetical subjects representative of the 50-year-old US population at average risk for CRC. The model assumed that a polypectomy would be performed during a flexible sigmoidoscopy and included two possibilities for follow-up: Under one scenario, only persons found to have a high-risk polyp were referred for follow-up colonoscopy; under the other scenario, persons found to have any adenomatous polyp, regardless of size or histology, were referred for follow-up colonoscopy. Two types of guaiac FOBT were considered: rehydrated FOBT and unrehydrated FOBT.

**Data**

The data for this study was based on US figures obtained from the literature, including published results from the Minnesota RCT, and also the Surveillance Epidemiology and End Results (SEER) program. Cost data were obtained from a large Health Maintenance Organization (HMO).

**Results**

Compared with flexible sigmoidoscopy every 10 years, screening with annual unrehydrated FOBT plus flexible sigmoidoscopy every 10 years (with aggressive follow-up regime) resulted in an incremental cost-effectiveness ratio of USD 21,200 (approximately NZD 29,175) per life year gained.

The strategy of annual rehydrated FOBT plus flexible sigmoidoscopy every five years (with aggressive follow-up) yielded an incremental cost-effectiveness ratio of USD 92,900 (approximately NZD 127,847) per life year gained compared with a strategy of unrehydrated FOBT plus flexible sigmoidoscopy every five years.

The strategy of annual unrehydrated FOBT plus flexible sigmoidoscopy every five years (with aggressive follow-up) yielded an incremental cost-effectiveness ratio of USD 51,200 (approximately NZD 70,460) per life year gained compared with a strategy of unrehydrated FOBT plus flexible sigmoidoscopy every 10 years.

The following strategies were reported to dominate:

- annual rehydrated FOBT
- annual unrehydrated FOBT
- annual unrehydrated FOBT plus flexible sigmoidoscopy every five years
- annual rehydrated FOBT plus flexible sigmoidoscopy every five years (with less aggressive follow-up), and
- annual rehydrated FOBT plus flexible sigmoidoscopy every 10 years.
Sensitivity analysis

Sensitivity analysis showed that under more optimistic, but still plausible scenarios, it could be less costly to screen than not to screen (due to averted cancer treatment costs): If the costs of sigmoidoscopy were 15 percent less than assumed (USD 279 (approximately NZD 384)) or the costs of cancer treatment were 20 percent higher than assumed (USD 22,000 to USD 58,300 (approximately NZD 30,276 to NZD 80,231)), then screening was cost-saving. Screening would also be cost-saving if the incidence of polyps at 50 years of age was 20 percent higher than estimated (the incidence of polyps at 50 years of age estimated for the base case was 21 percent).

Limitations

The study claimed to be from a societal perspective although only costs to the health care system were included. The perspective was equivalent to those studies which assume a publicly-funded health care system and take the perspective of third-party payer. As a result, indirect costs such as patient time and travel costs, informal carer costs, and the effect of anxiety were not included in the estimates. The study also did not include the effect on general consumption and productivity, which would be relevant from a societal perspective.

The cost data did not allow for programme costs such as the costs of health promotion, recall systems, and for administrative overheads to be included in the analysis.

Furthermore, the model did not allow for inclusion of estimates of the substantial additional capital investment which may be required for endoscopy facilities nationwide, as well as training for staff.

Life years saved were not adjusted for quality of life.

Conclusions

The authors concluded that screening for CRC, even in a context of low compliance, was expected to significantly reduce mortality at a cost that is comparable to other cancer screening programmes. Nevertheless, compliance significantly affected the cost-effectiveness ratios. When compared with a strategy of no screening, flexible sigmoidoscopy alone was associated with a lower incremental cost-effectiveness ratio than annual FOBT plus flexible sigmoidoscopy at five or 10 years.

Flexible sigmoidoscopy plus FOBT compared with flexible sigmoidoscopy or FOBT or no screening: Conclusions

The study reviewed above was of high quality and captured most of the costs associated with a CRC screening programme based on flexible sigmoidoscopy every five or 10 years plus annual FOBT. The results suggest that combining flexible sigmoidoscopy every five or 10 years with annual FOBT is likely to be associated with a more favourable cost-effectiveness ratio than flexible sigmoidoscopy alone and that the cost-effectiveness of such a programme would be comparable to that of other cancer screening programmes.

It should be noted that the estimated costs do not fully account for programme-related expenses so the true costs of a CRC screening programme based on flexible sigmoidoscopy plus annual FOBT are likely to be somewhat higher than the estimates presented.

Furthermore, the evidence on effectiveness of flexible sigmoidoscopy is scarce and, therefore, the results are heavily reliant on assumptions of sensitivity and specificity.

The study also notes that the participation rate may be a pivotal factor in achieving a favourable cost-effectiveness ratio.
Table 16: Primary research appraised relevant to cost-effectiveness of FOBT screening compared with no screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whynes et al. (1999)</td>
<td>A Markov model was used to compare the use of Guaiac-based FOBT, unrehydrated Haemoccult II, with no screening programme. Based on a median 8-year follow-up, and a longer-term follow-up simulated by modeling, the study estimated cost per QALY gained as a result of screening according to the Nottingham protocol versus no screening programme. The model differentiated between patients with cancers or adenomas which would have become detected and received treatment and patients with cancers or adenomas which would never have been detected in the patient’s lifetime. Benefits of screening can only be applied to the former.</td>
<td>Clinical and cost data were drawn from the Nottingham trial and supplemented by values from the literature to allow a quality adjustment to life years saved.</td>
<td>Cost per QALY gained from using Haemoccult II according to the Nottingham protocol was estimated to be approximately £5,700 (approx. NZD 15,014) for males and £5000 (approx. NZD 13,170) for females (1995-96 prices). Longer term estimates based on modeling suggested approximately £2000 (approx. NZD 3268) per QALY gained for males and £1,400 (approx. NZD 3688) per QALY gained for females.</td>
<td>Results were shown to be sensitive to the cost of the FOBT. Results were particularly sensitive to the specificity of the FOBT. A specificity that is 10% lower than the assumed specificity of FOBT doubles the cost per QALY gained. Although shorter screening intervals (annual vs. biennial screening; considered) increase costs, outcomes also tend to improve such that there is little impact on the ICER.</td>
<td>Compared with estimates of cost per QALY gained as a result of the breast cancer screening programme, which are estimated to be £3500 to £6000 (approx. NZD 9219 to NZD 15,804) (1995 prices), the estimates of cost per QALY gained from CRC screening using FOBT compared with no screening are low.</td>
</tr>
<tr>
<td>Gyrd-Hansen et al. (1998) and Gyrd-Hansen (1999)</td>
<td>A Markov model was used to estimate the cost-effectiveness of 60 different screening programmes, involving varying screening intervals and target age groups, for CRC based on the Haemoccult II FOBT and the outcomes of the Funen-1 trial. The observed costs and effects of the Funen-1 trial were based on screening an unscreened population and were, therefore, representative of the first few years of screening. The model estimated costs and effects for 36 years in order to generate costs and effects associated with a permanent rate of cancer detection. Positive FOBT was assumed to be investigated by colonoscopy or by DCBE where colonoscopy was not possible.</td>
<td>Clinical and cost data were derived from the Funen-1 RCT.</td>
<td>The six most efficient programmes evaluated were: biennial screening of 65-74 year olds, of 60-74 year olds, and of 55-74 year olds, with incremental cost per life year saved ranging from 17,000 (approx. NZD 4132) DKK to 18,800 DKK (approx. NZD 4570); and, screening 55-74 year olds every 1.5 years (ICER=20,200 DKK (approx. NZD 4910)), screening 55-74 year olds annually (ICER=23,000 DKK (approx. NZD 5,590)), and screening 50-74 year olds annually (ICER=26,000 DKK (approx. NZD 6320)).</td>
<td>The cost of colonoscopy and variations in the excess survival rate were found to have significant effects on the ICER. The inclusion of future unrelated health care costs, which approximates a societal perspective, marked these affected the cost of screening (ICERs were now in the range of 42,800 DKK to 63,800 DKK (approx. NZD 10,403 to NZD 15,507)) and made targeting younger age groups more favourable.</td>
<td>CRC screening programmes based on FOBT Haemoccult II are cost-effective and compare favourably with screening programmes for breast and cervical cancer. In the Danish context it would be optimal to introduce annual CRC screening of 50-74-year-olds.</td>
</tr>
<tr>
<td>Helm et al. (2000)</td>
<td>A direct projection of the results of 3 RCTs to the US population. The study determined what the effects and costs would have been if each RCT had been applied to a relevant cohort of the US population. Three outcomes and their associated costs were generated: number of lives saved, number of cancers detected, and number of life years saved. Positive FOBT results were assumed to be investigated by colonoscopy or by DCBE where colonoscopy is not possible.</td>
<td>Clinical data were derived from the Minnesota, Funen-1, and Nottingham trials. National data, especially on costs, supplemented trial data where relevant.</td>
<td>Cost per life year saved was estimated to be USD 2500 (approx. NZD 3440) based on the Nottingham protocol, USD 2700 (approx. NZD 3716) based on the Funen-1 protocol, and USD 20,500 (approx. NZD 28,212) based on the Minnesota protocol.</td>
<td>Plausible variations in parameters resulted in ranges of USD 1300 to USD 4200 (approx. NZD 1789 to NZD 5780) per life year saved based on the Nottingham protocol; USD 1600 to $4200 (approx. NZD 2202 to NZD5780) per life year saved based on the Funen-1 protocol; and, $11,400 to $32,500 (approx. NZD 15,688 to NZD 44,726) per life year saved based on the Minnesota protocol.</td>
<td>Although the effectiveness of FOBT in saving lives is limited, it is comparable to the effectiveness of accepted methods of screening for other kinds of cancers and its costs are relatively low. Overall the cost-effectiveness of screening for CRC using FOBT versus not screening is favourable.</td>
</tr>
</tbody>
</table>
Table 16. Primary research appraised relevant to cost-effectiveness of FOBT screening compared with no screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flanagan et al. (2003)</td>
<td>A microsimulation model was used to simulate biennial screening of subjects aged 50-74 years in the year 2000 in Canada. A hypothetical fixed cohort allowed simulation of clinical trial conditions in order to validate the results of dynamic cohorts from 2000 to 2024, which allowed the longer-term effects of a population-based CRC screening programme using FOBT unrehydrated Haemoccult II to be estimated.</td>
<td>Canadian population data supplemented RCT data on sensitivity, specificity, participation, incidence, staging, progression of disease, and mortality. Canadian health system cost data was used.</td>
<td>CRC screening using FOBT Haemoccult II was found to result in a ten year mortality reduction of 1.7 percent and a 15 percent increase in the demand for colonoscopies. Incremental cost per life year saved, compared with no screening, was estimated at CDN 11,907 (approx. NZD 13,216).</td>
<td>Plausible variations on the key parameters yielded only modest variations in the estimated ICER. The ratio remained favourable even under high-cost scenarios.</td>
<td>Screening for CRC using FOBT followed by colonoscopic investigation of positive results is cost-effective for the Canadian context. However, the cost-effectiveness relies on achieving the level of compliance observed in the trials. Existing screening programmes in Canada (for breast and cervical cancer) have much lower participation rates.</td>
</tr>
<tr>
<td>Whynes (2004)</td>
<td>Presented a cost-effectiveness analysis of Nottingham trial results for a CRC screening programme based on FOBTs Haemoccult and Feca. Estimates of cost per cancer detected were derived directly. Estimates of cost per life year gained were also generated. In this study, benefits were discounted at a lower rate than costs (2% vs. 6% p.a.).</td>
<td>Clinical and cost data were based on the most recent published results of the Nottingham trial.</td>
<td>Cost per cancer detected ranged from £3113 (approx. NZD 8200) for 3-day haemoccult to £11,924 (approx. NZD 31,408) for combined use of Feca and Haemoccult. Overall incremental cost effectiveness ratio of a screening programme based on the Nottingham trial protocol was estimated at £1584 (approx. NZD 4172) per life year gained.</td>
<td>Discounting benefits at the same rate as costs raised the ICER by 77.4% relative to the base estimate. Testing costs were the greatest cost-related source of sensitivity: doubling testing costs resulted in a 59.6% increase in the ICER.</td>
<td>Overall cost-effectiveness of screening for CRC using FOBT was favourable.</td>
</tr>
<tr>
<td>Stone et al. (2004)</td>
<td>An evaluation of CRC screening in Australia, targeting 55-69 year-olds through the use of biennial Guaiac-based FOBT compared with the status quo of minimalist opportunistic screening, using a Markov model and assuming a hypothetical nationally coordinated screening programme for average-risk Australians in 1996. The programme was assumed to be in steady state to provide estimates of ongoing costs.</td>
<td>Clinical data was based on meta-analysis of three RCTs: Nottingham, Funen-1, and Minnesota. National data was used for incidence, mortality, age of onset, mean survival duration of each phase. Medicare cost data.</td>
<td>A base programme of screening 55-69 year-olds would have an incremental cost per DALY saved of AUD 17,000 (approx. NZD 18,310). Extension of the programme to include 70-74 year-olds would be more cost-effective than including the 50-54 year-olds.</td>
<td>Multi-way sensitivity analysis was performed and included uncertainty distributions based on a combination of the reported confidence intervals, the range of reported values in the literature, and expert opinion of the range of likely values in the Australian context. Less and more optimistic assumptions regarding mortality reduction resulted in a range of ICER from AUD 13,000 (approx. NZD 14,002) per DALY gained to AUD 52,000 (approx. NZD 56,007) per DALY gained.</td>
<td>The findings support the case for a national programme directed at the 55-69 year age group with extension to the 70-74 year age group if resources allow.</td>
</tr>
</tbody>
</table>
Table 17. Primary research appraised relevant to cost-effectiveness of Guaiac-based FOBT screening compared with immunological FOBT screening

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyrd-Hansen (1998)</td>
<td>A cost-effectiveness analysis of the use of Guaiac based FOBTs, Haemoccult II based on the experience of the Funen-1 trial, compared with alternative FOBT tests: rehydrated Haemoccult II, HemeSelect, and Haemoccult II Sensa as if these were to be used in a similar population-based screening programme. The cost-effectiveness of screening 55-74 year olds at one- and two-year intervals as well as screening 50-74 year olds annually is estimated. Incremental cost per life year gained compared with no screening was estimated.</td>
<td>Data on unrehydrated Haemoccult II was drawn from the Funen-1 trial. Data on unrehydrated Haemoccult II was drawn from the Minnesota and Göteborg trials. Data on HemeSelect and Haemoccult II Sensa was derived from the literature.</td>
<td>The higher sensitivity of the rehydrated Haemoccult II test, the Haemoccult Sensa, and the HemeSelect test was at a cost of lower specificity. As a result, the most cost-effective screening programmes were: biennial screening of 55-74 year olds using unrehydrated Haemoccult II (ICER=17,500DKK (approx. NZD 4254)); annual screening of 50-74 year olds using unrehydrated Haemoccult II (ICER=39,000DKK (approx. NZD 9479)); annual screening of 50-74 year olds using HemeSelect (ICER=71,300DKK (approx. NZD 17,330)); and, annual screening of 50-74 year olds using rehydrated Haemoccult II (ICER=138,100DKK (approx. NZD 33,567)).</td>
<td>The major source of uncertainty lies around the cost-effectiveness of the rehydrated Haemoccult II test. If the Minnesota trial results are used rather than the Göteborg results, from which the base estimates were derived, the rehydrated Haemoccult II test is more cost-effective than the alternatives.</td>
<td>The unhydrated Haemoccult II test is likely to be the most cost-effective option. The rehydrated test may be an alternative but evidence of the test's characteristics remain unclear. Use of the immunological FOBT would result in a significantly higher cost per life year saved.</td>
</tr>
<tr>
<td>Van Ballegooijen et al. (2003)</td>
<td>The cost-effectiveness of the guaiac-based FOBTs, Haemoccult II and HemoccultSENSA were compared with that of a hypothetical immunochemical FOBT assumed to have comparable sensitivity to HemoccultSENSA but with higher specificity. A microsimulation model was used to simulate the natural history of the adenoma-carcinoma sequence and applied the hypothetical screening programme using annual FOBT to 72,000 individuals aged 65-79 over 30 years of screening.</td>
<td>Clinical data were derived from the Minnesota RCT. Costs were based on Medicare data.</td>
<td>Cost-effectiveness ratios for all FOBTs were found to be favourable. HemoccultSENSA was found to be more favourable due to low specificity. The ICER associated with the hypothetical immunochemical test was $11,000 (approx. NZD 15,138) per additional year of life saved over Haemoccult II if specificity of the IFOBT is assumed to be 98%; $21,000 (approx. NZD 28,900) if the IFOBT is assumed to be 95% specific.</td>
<td>Sensitivity analysis included varying compliance rates, which suggested that achieving target compliance was pivotal to cost-effectiveness. All cost-effectiveness ratios are found to be sensitive to assumptions regarding the natural history of CRC and the cost of colonoscopy.</td>
<td>FOBTs in general were a cost-effective intervention for reducing CRC mortality. Guaiac FOBT was the most cost-effective</td>
</tr>
<tr>
<td>Berchi et al. (2004)</td>
<td>Compared the cost-effectiveness of the guaiac-based FOBT Haemoccult with the cost-effectiveness of immunochemical FOBT Magstream. Costs and effects were estimated using a Markov-type model for successive biennial CRC screening campaigns on a hypothetical cohort of 165,000 subjects aged 50-74 years screened over 20 years with 43.7% compliance.</td>
<td>Clinical data were based on the Funen-1 RCT as well as the trial in Calvados, France. Cost data were derived from the Calvados trial.</td>
<td>The incremental cost-effectiveness of substituting Magstream for Haemoccult was estimated to be 7458 euros (approx. NZD 13,471) per life year saved after 10 years of screening and 2980 euros (approx. NZD 5382) per life year saved after 20 years of screening.</td>
<td>Sensitivity analysis included varying compliance rates, which suggested that implementing target compliance was pivotal to cost-effectiveness. All cost-effectiveness ratios are found to be sensitive to assumptions regarding the natural history of CRC and the cost of colonoscopy.</td>
<td>Using an immunochemical FOBT like Magstream could provide a substantial increase in sensitivity over the Haemoccult FOBT and, as a result, increase the cost-effectiveness of CRC screening at a reasonable cost to society.</td>
</tr>
</tbody>
</table>
Table 18. Primary research appraised relevant to cost-effectiveness of flexible sigmoidoscopy alone

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frazier et al. (2000)</td>
<td>A Markov-based model based on CRC screening from age 50 years using several possible strategies, including five-yearly and 10-yearly flexible sigmoidoscopy. Follow-up to flexible sigmoidoscopy takes two possible approaches: (1) persons found to have a high-risk polyp were referred for follow-up colonoscopy and (2) persons found to have any adenomatous polyp, regardless of size or histology, were referred for follow-up colonoscopy. The analysis was based on screening white males with 60% compliance.</td>
<td>Clinical data were derived from the Minnesota RCT, SEER, and the literature. Cost data were derived from costs of a large HMO.</td>
<td>Compared with no screening, screening with flexible sigmoidoscopy every 10 years, regardless of follow-up regime, resulted in an ICER of less than USD 17,000 (approx. NZD 23,395) per life year gained. The more aggressive approach was also associated with a substantial increase in effectiveness. Flexible sigmoidoscopy every five years was dominated by the 10 year interval.</td>
<td>Under more optimistic scenarios, it was found to be possible that screening would be less costly than not screening due to averted treatment costs.</td>
<td>Screening for CRC, even in a context of low compliance, was expected to significantly reduce mortality at a cost that is comparable to other screening programmes. Flexible sigmoidoscopy every 10 years was preferable to flexible sigmoidoscopy every five years as a cost-effective strategy and the cost-effectiveness was not significantly affected by aggressive follow-up regimes.</td>
</tr>
<tr>
<td>Sonnenberg et al. (2000)</td>
<td>This study compared the cost-effectiveness of flexible sigmoidoscopy every five years with a strategy of no screening, with annual FOBT, and with colonoscopy every 10 years. The analysis was based on a Markov-type model simulating results for subjects 50 years of age. Positive results of FOBT or flexible sigmoidoscopy were followed up by colonoscopy.</td>
<td>Clinical data were derived from the Funen-1 and Nottingham RCTs and large autopsy studies as well as US vital and cancer statistics. Costs were based on Medicare reimbursement data.</td>
<td>Compared with no screening, screening with flexible sigmoidoscopy every five years resulted in an ICER of USD 36,509 (approx. NZD 50,243) per life year saved. Compared with annual FOBT, screening with flexible sigmoidoscopy resulted in an ICER of USD 65,704 (approx. NZD 90,420) per life year saved.</td>
<td>The sensitivity of results was assessed with respect to annual FOBT and colonoscopy as these were found in the base case to be the most favourable strategies. Varying the interval of flexible sigmoidoscopy to 10 years did not change the result that flexible sigmoidoscopy would be dominated by annual FOBT and colonoscopy every 10 years.</td>
<td>Screening with flexible sigmoidoscopy every five years was dominated in terms of cost-effectiveness by annual FOBT and by colonoscopy every 10 years.</td>
</tr>
</tbody>
</table>
Table 18. Primary research appraised relevant to cost-effectiveness of flexible sigmoidoscopy alone (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loeve et al. (2000)</td>
<td>This study estimated the costs and effects of screening for CRC using flexible sigmoidoscopy every five years versus colonoscopy every 10 years and versus no screening. Microsimulation is used to generate results for subjects aged between 50 and 75 years. The model relied on assumptions regarding the sensitivity of flexible sigmoidoscopy.</td>
<td>Clinical data were derived from autopsy and colonoscopy studies. Data were also derived from SEER and the literature, including RCT results. Cost data were based on published estimates.</td>
<td>Screening for CRC with flexible sigmoidoscopy every five years was found to be associated with savings in treatment costs compared with a strategy of no screening from the fifth year of the screening programme.</td>
<td>Under higher cost scenarios, savings do not occur at the same rate as found in the base case. Assumption variations applied to the adenoma-carcinoma sequence and dwelling time still showed potential for savings over a long-term screening programme although these could be considerably less than estimated in the base case.</td>
<td>There is potential for savings in treatment costs to allow break-even with programme costs in a CRC screening programme based on colonoscopy or flexible sigmoidoscopy, although this break-even would occur only over the long term and is associated with considerable uncertainty.</td>
</tr>
<tr>
<td>O’Leary et al. (2004)</td>
<td>This study compared the cost-effectiveness of flexible sigmoidoscopy every 10 years with annual and biennial rehydrated FOBT Hemoccult, with colonoscopy every 10 years, and with no screening. A Markov-type model was used to generate estimates for subjects aged 55-64 over 10 years and with compliance rates of 42% for flexible sigmoidoscopy and 60% for FOBT.</td>
<td>Clinical data were derived from the literature, including the results of an Australian flexible sigmoidoscopy screening programme. Costs were derived from the Medicare Benefits Schedule.</td>
<td>Compared with no screening, screening with flexible sigmoidoscopy every 10 years was found to be associated with an incremental cost-effectiveness ratio of AUD 16,801 (approx. NZD 18,096) per life year saved. This was a lower ICER compared with no screening than for biennial and annual FOBT.</td>
<td>Sensitivity analysis was conducted varying uncertain parameters, including assumptions about the natural history of CRC, costs, and compliance rates. Cost-effectiveness of flexible sigmoidoscopy was found to be particularly sensitive to assumptions about the progression to cancer of adenomas of at least 10mm. Base case probability of progression was 5%. Reducing the probability to 2% increased the ICER to AUD 32,362 (approx. NZD 34,856).</td>
<td>Screening for CRC with flexible sigmoidoscopy was cost-effective compared with no screening and was associated with a more favourable cost-effectiveness ratio than FOBT.</td>
</tr>
</tbody>
</table>
Table 19. Primary research appraised relevant to cost-effectiveness of flexible sigmoidoscopy with FOBT

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Data</th>
<th>Results</th>
<th>Sensitivity Analysis</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frazier et al. (2000)</td>
<td>A Markov-based model based on CRC screening from age 50 years using several possible strategies, including five-yearly and 10-yearly flexible sigmoidoscopy plus annual unrehydrated FOBT. Follow-up to flexible sigmoidoscopy takes two possible approaches: (1) persons found to have a high-risk polyp were referred for follow-up colonoscopy and (2) persons found to have any adenomatous polyp, regardless of size or histology, were referred for follow-up colonoscopy. The analysis was based on screening white males with 60% compliance.</td>
<td>Clinical data were derived from the Minnesota RCT, SEER, and the literature. Cost data were derived from costs of a large HMO.</td>
<td>Compared with no screening, screening with annual unrehydrated FOBT plus flexible sigmoidoscopy every 10 years and an aggressive follow-up regime resulted in an ICER of USD 21,200 (approx. NZD 29,175) per life year saved. The corresponding ICER for annual unrehydrated FOBT plus flexible sigmoidoscopy every five years is USD 51,200 (approx. NZD 70,460) and USD 92,200 (approx. NZD 126,884) if the FOBT is rehydrated.</td>
<td>Under more optimistic scenarios, it was found to be possible that screening would be less costly than not screening due to averted treatment costs.</td>
<td>Screening for CRC with annual FOBT plus flexible sigmoidoscopy every 10 years was cost-effective, but generated higher cost-effectiveness ratios than flexible sigmoidoscopy alone.</td>
</tr>
</tbody>
</table>
SECONDARY RESEARCH: STUDY DESIGNS AND QUALITY

The search identified three eligible secondary research studies. An overview of these studies is provided here. As discussed in the Methodology chapter, these papers may not employ the same inclusion criteria as applied in this review and the results must be interpreted with care.

The Australian Health Technology Advisory Committee (1997) produced a report for the Commonwealth Department of Health and Family Services in Australia, which included a review of cost-effectiveness studies (published before 1997) of CRC screening. The committee noted the lack of applicability of the majority of studies due to their overseas context. The committee also suggested that knowledge of the economic issues relating to CRC screening was deficient in a number of areas including: uncertainty surrounding the effectiveness of screening; the marginal cost-effectiveness of screening particular population sub-groups; data relating to the cost of diagnostic work-up, the cost of surveillance and follow-up, the cost of treatment, and personal costs to the patient; the likely compliance of the Australian population with CRC screening and the relative acceptability of various screening modalities; and, quality of life for people with CRC. The report also recommended that a major area of uncertainty affecting the reliability of estimates of life years saved and cost per life year saved, the adenoma-carcinoma sequence, needs to be better understood.

The Conseil d’Évaluation des Technologies de la Santé du Québec (2000) report to the Québec Minister of Research, Science and Technology included a review of studies, some of which were published in or after 1997, including Gyrd-Hansen et al. (1997), Hristova and Hakama (1997), and Whynes et al. (1998).

The Conseil also adapted the model used by Walker and Whynes (1992) in order to generate estimates of cost-effectiveness of CRC screening for a range of modalities in the Quebec population. Estimates are presented for cost per cancer detected for 60 percent compliance in one round of screening but not for cost per life-year saved or for ongoing screening. Results were:

- a cost of between CDN 10,258 and CDN 15,167 (approximately NZD 11,386 to NZD 16,834) per cancer detected for unrehydrated Haemoccult II;
- a cost of between CDN 19,372 and CDN 25,504 (approximately NZD 21,500 to NZD 28,308) per cancer detected for rehydrated Haemoccult II;
- a cost of CDN 14,121 (approximately NZD 15,673) per cancer detected for HemeSENSA;
- a cost of between CDN 46,082 and CDN 56,594 (approximately NZD 51,148 and NZD 62,816) per cancer detected for HemeSelect, and;
- a cost of CDN 100,742 (approximately NZD 111,818) per cancer detected for flexible sigmoidoscopy.

The Conseil noted that the models of cost-effectiveness presented were often limited in scope and that results were quite sensitive to variations in parameters around which there was still considerable uncertainty. However, the Conseil also suggested that although the estimated results were not likely to be accurate, the range of results indicated that there was sufficient evidence for consideration to be given to implementing biennial CRC screening in Québec using a non-rehydrated FOB test.

The Pignone, Saha et al. (2002) study consisted of an evaluation of the cost-effectiveness of various screening methods used in population-based screening, with results derived from a systematic review of cost-effectiveness studies published between January 1993 and September 2001. The purpose of the study was to inform the US Preventative Services Task Force’s deliberation on recommendations regarding colorectal cancer screening.

Several major RCT-based cost-effectiveness studies (those based on the Funen-1 trial and the Nottingham trial) were excluded due to the authors’ assessment that the costs, incidence, and prevalence estimates used were not comparable to US figures. The five studies for which most detail is provided were Wagner et al. (1996), Frazier et al. (2000), Khandker et al. (2000), Sonnenberg et al.
All studies included found that screening for colorectal cancer by any of the screening strategies considered (FOBT, flexible sigmoidoscopy, FOBT plus flexible sigmoidoscopy, DCBE, and colonoscopy) reduced colorectal cancer mortality for average risk adults older than 50 years. Most strategies were found to have average cost-effectiveness ratios between USD 10,000 (approximately NZD 13,762) and USD 25,000 (approximately NZD 34,404) per life year saved when compared with no screening. Worst case scenarios generated by sensitivity analyses produced cost-effectiveness ratios below USD 100,000 (approximately NZD 137,618) per life year saved when compared with no screening.

The combination of annual FOBT and flexible sigmoidoscopy every five years was found by three of the studies to be the most cost-effective strategy. Two studies found colonoscopy every ten years to be the most cost-effective, although only one of these considered the combined strategy of FOBT plus flexible sigmoidoscopy. Four of the five multiple strategy analyses found FOBT alone to be more effective than flexible sigmoidoscopy alone for reduction of CRC mortality.

Compared with no screening, annual FOBT was found to have a cost-effectiveness ratio of between USD 5,691 (approximately NZD 7,832) and USD 17,805 (approximately NZD 24,503) per life year saved; flexible sigmoidoscopy every five years was found to have a cost-effectiveness ratio of between USD 12,477 and USD 39,359 (approximately NZD 17,171 and NZD 54,165) per life year saved; and, annual FOBT and flexible sigmoidoscopy every five years was found to have a cost-effectiveness ratio of between USD 13,792 and USD 22,518 (approximately NZD 18,980 and NZD 30,989) per life year saved.

At low willingness-to-pay thresholds (i.e. less than USD 30,000 (approximately NZD 41,285) per life year saved), the studies reached heterogeneous conclusions about the preferred strategy: two identified annual FOBT; one identified flexible sigmoidoscopy every five years; and two identified colonoscopy every ten years. At higher willingness-to-pay thresholds (i.e. USD 30,000 to USD 100,000 (approximately NZD 41,285 to NZD 137,618) per life year saved), annual FOBT with flexible sigmoidoscopy every five years was preferred and colonoscopy every ten years or colonoscopy at age 55 years and age 65 years only were more favourable.

The investigators concluded that the overall cost-effectiveness of screening for CRC compared favourably with other commonly endorsed preventive health care interventions, such as mammography breast cancer screening in women over 50 years old, or the treatment of moderate hypertension. It was not clear, however, whether one method of screening for colorectal cancer was superior to other methods due to the heterogeneity of study results, which was associated with the differences in assumptions regarding the biologic behaviour of CRC, the adverse events and effectiveness of each strategy, and compliance levels.

CONCLUSIONS

The Working Party on Screening for Colorectal Cancer (1998) reviewed evidence of cost-effectiveness up to 1998. The report noted that studies based on the results of the Nottingham and Funen-1 trials showed that, in terms of cost per life year saved, CRC screening using FOBT was comparable with screening for breast cancer. However, the report also noted that studies did not include intangible costs or the adverse effects of screening. Furthermore, published studies did not fully account for the health service costs of a screening programme, including health promotion, recall systems, administrative overhead, training, and set-up costs. With regard to other screening modalities, the New Zealand report noted that there was no evidence from RCTs as to the effectiveness of other screening options, which prevents their cost-effectiveness from being accurately determined.

In this review, some of the same studies (Whynes 1998 and Gyrd-Hansen 1998) have been considered along with many more recent studies. Many of these have a strong grounding in RCTs. However, it remains the case that only guaiac-based FOBT has been the subject of large RCTs with published results on clinical outcomes as well as costs. Overall the studies reviewed suggest that screening for CRC using annual or biennial guaiac-based FOBT is cost-effective and well within the range of cost-
Effectiveness estimated for other screening programmes, namely breast cancer screening and cervical cancer screening.

Studies of other screening modalities, such as immunochemical FOBT, flexible sigmoidoscopy and annual FOBT plus flexible sigmoidoscopy, are less definitive in their results. A weaker data set to draw from due to the lack of large RCTs is the underlying challenge. As a result, these studies can only suggest that these other screening modalities are potentially cost-effective. Due to the lack of grounding in large RCTs, greater degree of uncertainty is associated with the results of these studies.

It should be noted, however, that all studies of cost-effectiveness reviewed here have found that the estimates of cost-effectiveness are sensitive to assumptions regarding the adenoma-carcinoma sequence and polyp dwell time, suggesting that uncertainty on the clinical side remains a major obstacle to generating robust results on cost-effectiveness.

The perspective of the studies of CRC screening also leaves several issues still open to question: The widest perspective taken by any of the studies reviewed is that of third-party payer or that of government-funded health system. As a result societal costs and benefits were not included. Even allowing for a limited perspective, the programme costs of a colorectal cancer screening programme are not well estimated: Many studies used only testing, follow-up and treatment costs while others have, at best, been able to generate estimates of what a real CRC screening programme would cost to run.

In conclusion, the studies included indicate that significant progress has been made in generating evidence on guaiac-based FOBT since the 1998 working party in that a great deal more evidence is available from RCTs and that robust simulation models have been developed that allow RCT results to be projected credibly onto other populations, and showing that, even in different contexts and with different costs, cost-effectiveness estimates are remarkably consistent. On the other hand, little progress has been made in generating evidence on other screening modalities or in identifying the true costs of a CRC screening programme.
Chapter 8: Discussion

SUMMARY OF EVIDENCE

Aim of the review

This review aimed to identify and appraise the international evidence for the effectiveness and cost-effectiveness of FOBT screening, the accuracy of guaiac compared to immunochemical FOBT, FS screening, and combined FOBT and FS screening. It was commissioned by the National Screening Unit, and was primarily intended to provide an updated review of relevant evidence for CRC screening since that considered by the Working Party on Screening for Colorectal Cancer (1998). Consequently, this report should be read in conjunction with the previous report.

Results of the search strategy

Following an extensive search strategy, approximately 1986 articles were identified by the search strategy, 220 articles were retrieved as full text from which a final group of 56 articles was identified as eligible for appraisal and inclusion in the review. Some of these articles included data relevant to multiple research questions requiring more than one appraisal; appraisals therefore totalled 64:40 considering primary research articles and 24 considering systematic reviews.

Clinical effectiveness of faecal occult blood test (FOBT) screening compared with no screening

Eleven eligible primary research papers and 13 eligible secondary research papers (nine reviews) were identified. The search identified four eligible RCTs comparing FOBT screening with no screening. The RCTs were all graded as level II evidence. Ten of the included papers updated data since 1997 on outcomes from three of these trials, which compared screening with the guaiac test Haemoccult/Haemoccult II with no screening. The other included paper reported from an RCT that compared screening with an immunochemical test plus a health questionnaire to no screening.

As the Working Party on Screening for Colorectal Cancer (1998) found, there is high-quality evidence that FOBT screening with the guaiac FOBT Haemoccult reduces mortality from CRC. Evidence available from ongoing follow-up of the three RCTs suggests that this mortality reduction has been sustained for the populations in which screening has stopped, but decreased slightly for the population to whom screening has continued to be offered. The investigators from the RCT that has continued screening stated that the most probable explanation for this observation was the decrease in the number of subjects being screened with increasing number of rounds (Kronberg et al, 2004). It is noted that such a mortality reduction could have been explained if the control group had also been offered screening, but this has not been reported as happening. The direct evidence available concerning the efficacy of an immunochemical test to reduce CRC mortality was less robust, but did suggest that a reduction in rectal cancer may be achievable with the use of this test.

Some new evidence, as detailed in the report, regarding the benefits and harms from CRC screening with FOBT has become available. However, little new data regarding possible psychological harms became available as a result of this review. Similar conclusions can be drawn from the results of this review as those made by other reviewers, including the Working Party on Screening for Colorectal Cancer (1998). The evidence available since this topic was last considered in New Zealand has not changed substantially.

Accuracy of immunochemical FOBT screening compared with guaiac FOBT screening

Seven eligible primary research papers and five eligible secondary research papers were identified. The studies identified by this review evaluated a diverse range of both types of test. Design and quality
limitations limited the conclusions that could be drawn from the published literature about which test offers the optimal performance. The limited evidence available since the Working Party considered FOBT screening regarding immunochromatographic test performance compared to guaiac, suggested that only HemeSelect performs as well as, or better than the guaiac tests HOII and HOS when compared head to head in cross-sectional studies.

There was good evidence that the simplified testing process and sampling kit of the immunochromatographic test Insure encouraged a greater proportion of people from the general population to participate when invited to complete FOB screening tests.

The conclusions on this topic were similar to those made in other recent reviews. There is limited definitive evidence regarding superior immunochromatographic FOBT performance over the guaiac tests. HemeSelect is the immunochromatographic test that compares most favourably with the guaiac tests, but Insure may be more acceptable.

**Clinical effectiveness of flexible sigmoidoscopy (FS) screening compared with no screening**

No large-scale RCT has been completed providing incidence and mortality data relevant to the impact of flexible sigmoidoscopy screening. Two papers reporting on one small pseudo-randomised controlled trial were identified in the current review, the Telemark Polyp Study (Thiis-Evensen et al, 1999, Thiis-Evensen et al, 2001), graded as level III-1 evidence. This trial offered screening using flexible sigmoidoscopy to 400 men and women aged 50-59 years who were drawn from a population registry in Norway. A control group of 399 people received usual care and were unaware of their enrolment at baseline. At 13-year follow-up, Thiis-Evensen (2001) reported that there was no difference in the prevalence of adenomas between screening and control groups, however there were more large adenomas and a trend towards more high-risk adenomas found for participants in the screening group compared with the control group. Investigating CRC incidence and mortality outcome, Thiis-Evensen et al. (1999) reported that there was one-fifth the incidence of CRC at 13-year follow-up in the screening group compared with the control group. However higher overall mortality was observed in the screening group than the control group (14% versus 9%) (p=0.02). There has been some discussion as to whether those screened were more relaxed about lifestyle factors (such as smoking and obesity) following screening, which increased their likelihood of premature death (Hoff et al, 2001). However this trial was limited by its relatively small size; differences in how the samples were drawn for the screening and control groups affecting comparability; and that 13-year follow-up colonoscopy was not offered to those non-attending for FS in the screening group.

As acknowledged by the investigators of the Telemark Polyp Study, larger, fully randomised, trials are necessary to investigate more reliably the impact of FS screening on CRC incidence and mortality. Three such studies are currently underway. Two multi-centre trials using similar protocols in the UK (UK Flexible Sigmoidoscopy Screening Trial; Atkin et al, 2001) and Italy (the SCORE Trial; Segnan et al, 2002) offered one-time FS screening for those aged between 55 and 64 years. Incidence and mortality data will not become available before 2008. The other RCT underway in the United States is the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO), a community-based, multi-centre study evaluating the effectiveness of repeated cancer screening tests (including FS screening every 3-5 years) on site-specific cancer mortality for 55-74 year olds (Schoen et al, 2003). This trial will not report incidence and mortality data until 2010-2012. Baseline and procedural data from all three trials suggest that screening FS is likely to be feasible and acceptable, and the care and thought put into the planning and conduct of these trials augurs well for the robustness of their results, once published.

With respect to secondary literature, two systematic reviews (AHTAC, 1997; Walsh and Terdiman, 2003) reporting on evidence from RCTs were identified. The AHTAC review (1997) identified only one pseudo-randomised, controlled trial, the Telemark Polyp Study. Unpublished 13-year follow-up CRC incidence data were obtained following correspondence with the authors and when the trial’s small sample size and concerns about the comparability of the screening and control groups were raised. The review concluded that definitive evidence from randomised controlled trials was required before widespread screening could be seriously considered. Walsh and Terdiman’s (2003) review briefly (and largely uncritically) described the results of the Telemark Polyp Study in Norway that demonstrated the benefits of one-time FS screening in terms of reduced colorectal cancer incidence, but
not CRC mortality. Larger trials underway were mentioned although no outcome data were yet available.

**Clinical effectiveness of FS and FOBT combined screening compared with FOBT screening, FS screening, or no screening.**

Two eligible RCTs, the Funen-2 and Funen-3 trials (Rasmussen et al, 1999; Rasmussen et al, 2003), were identified comparing combined FS and FOBT with FOBT-only screening. These studies provided limited evidence regarding the effectiveness of CRC screening on health outcomes from data on CRC incidence and/or mortality. One other ongoing RCT was identified comparing combined FS and FOBT with FS screening. The study from the ongoing Norwegian NORCCAP trial (Gondal et al, 2003) presented baseline screening compliance and diagnostic yield findings. Data on CRC incidence and mortality from this RCT are expected in late 2007. No studies were identified as being eligible for inclusion comparing FOBT and FS combined, compared with no screening. Three other RCTs (Berry et al, 1997; Brevinge et al, 1997; Verne et al, 1998) comparing the diagnostic accuracy of combined screening compared to either modality alone were identified and included in Appendix 6. These studies did not include an analysis of the effectiveness of combined FOBT and FS screening on CRC incidence and/or mortality and focused on screening compliance, diagnostic yield and colonoscopic findings from once-only screening.

The Funen-2 trial (Rasmussen et al, 1999) had an adequate design and conduct but a short follow-up period, non-repeat screening and only limited CRC incidence and mortality data. Outcomes at 24-62 months post-screening showed no significant difference in the CRC incidence and mortality status of participants between the two groups of screenees. Although the diagnostic yield was greater in the combined screening group despite low FS compliance the addition of once only FOBT to once-only FS added nothing to the predictive value of FS. There were no serious complications reported after the 2235 sigmoidoscopies performed. The Funen-3 trial (Rasmussen et al, 2003) was of lower quality in terms of design and conduct due to limitations in the study design and conduct including pseudo-randomisation, subject selection bias, and limitations in the comparative analysis. The evidence for the effectiveness of combined screening over FOBT-alone screening from this trial was limited by the comparison of two screening trials (Funen-2 and Funen-3) started seven years apart. Also, the comparison of cumulative test results from biennial FOB testing compared with the diagnostic yield of one-time FOBT + FS screening, and CRC mortality data were not reported. The low compliance rate for FS in the combined FOBT + FS screening group was likely to have affected CRC detection rates on a screened-persons basis as once-only combined FOBT + FS offered no advantage over biennial FOBT screening and on an invited-persons basis, much higher CRC detection rates were apparent in the biennial FOBT programme. Similar to the Funen-2 trial, there was little diagnostic benefit with the addition of once-only FOBT to FS. There was a large difference in colonoscopy-detected CRC rates between the study arms due to the colonoscopy criteria used in the detection and removal of small adenomas (≤ 10 mm). One study (Gondal et al, 2003) from the Norwegian Colorectal Cancer Prevention (NORCCAP) trial underway was identified comparing combined FOBT and FS compared to FS alone, and baseline screening compliance and diagnostic yield findings were presented. Health outcomes data on five-year CRC incidence and mortality will become available in late 2007, according to the contacted trial investigator. The study methodology for the NORCCAP trial was robust in terms of design and execution, and this bodes well for the validity of the trial results on CRC incidence and mortality to be published in late 2007.

Additional to the three RCTs included in the section on clinical effectiveness of combined screening, three other eligible RCTs were also identified and appraised. These RCTs provided data on the diagnostic accuracy of combined FOBT and FS screening compared to either modality alone (appraisals presented in Appendix 6). There was no CRC incidence or mortality data were evaluated in these RCTs. Combination testing had significantly higher detection rates of neoplasms compared to FOBT alone, even with low FS compliance. There was insignificant additional diagnostic benefit from adding FOBT to FS compared to FS alone. Participant compliance for FS in combination testing with FOBT was low compared to FOBT alone or FS alone. This low compliance was associated with acceptability issues and participants knowing their FOBT results prior to undergoing FS.

For the review of secondary literature, one relevant systematic review (Walsh and Terdiman, 2003) was identified and appraised. The authors concluded that there was a lack of RCTs evaluating the health outcomes of FOBT and FS combined screening compared with either modality alone or to no
screening. The studies that were available considered single FOB testing in addition to one-time sigmoidoscopy, with a focus on comparing compliance and the diagnostic yield between testing modes. With the lack of evidence on CRC incidence and mortality outcomes, it was not possible to support combined screening or the selection of one test over another. From the studies available to the reviewers it was concluded that FS is probably better than FOB in terms of the identification of CRC or advanced adenomas. However, issues regarding the frequency of examinations, compliance, the significance of small polyps and protocols for follow-up colonoscopy remain. Reported complications associated with combined FS and FOB screening were equivalent to each test alone. Other secondary research that was identified included the same primary studies as Walsh and Terdiman (2003) and made similar conclusions.

The Working Party on Screening for Colorectal Cancer (1998) concluded that there was a lack of RCT evidence for a mortality reduction for a national screening programme based on flexible sigmoidoscopy. Evidence for a combined FOB and FS screening strategy compared to a strategy of either test alone was not evaluated by the working party. The evidence considered does not support a combined screening strategy in asymptomatic middle-aged populations over a screening strategy involving either FOB or FS alone. Such a strategy cannot be justified nor is it feasible based on the available evidence. Few RCTs were available and those that were had limited health outcome data and were primarily focused on screening compliance, the diagnostic yield of combined screening compared to either test alone, and the utilisation of colonoscopy and detection in screening follow-up. Once the results of the NORCCAP trial are known, this conclusion should be reviewed.

Cost-effectiveness of CRC screening modalities

The Working Party on Screening for Colorectal Cancer (1998) reviewed evidence of cost-effectiveness up to 1998. The report noted that studies based on the results of the Nottingham and Funen-1 trials showed that, in terms of cost per life year saved, CRC screening using FOB was comparable with screening for breast cancer. However, the report also noted that studies did not include intangible costs or the adverse effects of screening. Furthermore, published studies did not fully account for the health service costs of a screening programme, including health promotion, recall systems, administrative overhead, training, and set-up costs. With regard to other screening modalities, the New Zealand report noted that there was no evidence from RCTs as to the effectiveness of other screening options, which prevents their cost-effectiveness from being accurately determined.

In this review more recent studies have been considered. Many of these have a strong grounding in RCTs. However, it remains the case that only guaiac-based FOB has been the subject of large RCTs with published results on clinical outcomes as well as costs. Overall, the studies reviewed suggest that screening for CRC using annual or biennial guaiac-based FOB is cost-effective and well within the range of cost-effectiveness estimated for other screening programmes, namely breast cancer screening and cervical cancer screening.

Studies of other screening modalities, such as immunochemical FOB, flexible sigmoidoscopy, and annual FOB plus flexible sigmoidoscopy are less definitive in their results. A weaker data set to draw from due to the lack of large RCTs is the underlying challenge. Due to the lack of grounding in large RCTs, greater degree of uncertainty is associated with the results of these studies.

It should be noted, however, that all studies of cost-effectiveness reviewed here have found that the estimates of cost-effectiveness are sensitive to assumptions regarding the adenoma-carcinoma sequence and polyp dwell time, suggesting that uncertainty on the clinical side remains a major obstacle to generating robust results on cost-effectiveness.

The perspective of the studies of CRC screening also leaves several issues still open to question: the widest perspective taken by any of the studies reviewed is that of third-party payer or that of a government-funded health system. As a result, societal costs and benefits were not included. Even allowing for a limited perspective, the programme costs of a colorectal cancer screening programme are not well estimated: many studies used only testing, follow-up and treatment costs and others have, at best, been able to generate estimates of what a real CRC screening programme would cost to run.

In conclusion, the studies included indicate that significant progress has been made in generating evidence on guaiac-based FOB since the 1998 working party but that little progress has been made in
CONCLUSIONS

Consistent with the findings of the Working Party on Screening for Colorectal Cancer (1998), high-quality evidence was found that FOBT screening with the guaiac-based FOBT Haemoccult reduces mortality from CRC. Specifically, evidence from ongoing follow-up for three major RCTs suggests that this mortality reduction has been sustained for the populations in which screening has stopped, but decreased slightly for the population to whom screening has continued to be offered. Whilst less robust, the direct evidence available suggests that a reduction in rectal cancer may be achievable with the use of an immunochemical test. The discussion presented in Chapter 1 of this report about FOBT as a screening test raised several issues concerning other aspects of how to conduct an FOBT screening programme, such as how many positive slides should be considered as a positive test and what dietary advice should be given to screenees. As a result of this report, it appears that little additional guidance on these matters has emerged from the evidence since it was considered by the working party in 1998.

There is limited definitive evidence regarding superior immunochemical FOBT performance over the guaiac tests. However, evidence from cross-sectional studies suggests that the immunochemical test HemeSelect is comparable, or superior, to guaiac testing. Simplified FOBT tests may be more acceptable, as there was good evidence that the simplified testing process and sampling kit of the immunochemical test Insure encouraged greater participation rates. The conclusions on this topic should be revisited if further reliable evidence on the comparative performance of screening FOBTs becomes available.

International interest in establishing FOBT screening programmes for CRC remains high. A recently published report from a pilot screening programme conducted in five regions of the UK, using a guaiac test with identical biochemical characteristics to those used in the Nottingham and Funen trials, concluded that screening for CRC using FOBT testing was feasible within the context of the UK’s National Health Service (Steele and UK Colorectal Cancer Screening Pilot, 2004). Since that report, the UK’s health secretary has announced that a national screening programme based on FOBT testing for CRC will be introduced in England from April 2006 (Mayor, 2004). Australia is also currently evaluating FOBT screening, in a pilot programme that is being conducted in three states using two types of immunochemical FOBTs: Inform and Bayer Detect (http://www.cancerscreening.gov.au/bowel/index.htm, accessed 27 August 2004).

One small randomised controlled trial designed as a feasibility study was identified investigating the impact of flexible sigmoidoscopy screening, however no large sampled RCT has been completed providing incidence and mortality data. Three large ongoing trials are investigating flexible sigmoidoscopy as either one-off or repeated screening modalities for average-risk men and women aged from their mid-50s. Preliminary results are promising in terms of the feasibility and acceptability of this screening modality, however long-term incidence and mortality data will not become available for these trials for three to seven years. Whilst the introduction of FS in a national screening programme cannot currently be justified, this recommendation should be reviewed once the results of these trials are available. The generalisability of these findings to the New Zealand population would also need to be considered.

There was limited evidence available regarding the effectiveness of combined FOBT and FS screening on health outcomes in terms of CRC incidence and/or mortality. Two RCTs were appraised. One demonstrated no significant difference in the CRC incidence and mortality status of participants between the two groups of screenees at 24-62 months post-screening. The other trial, though of lower methodological quality, found that once-only combined FOBT + FS screening offered no advantage over biennial FOBT screening, but on an invited-persons basis, much higher CRC detection rates were apparent in the biennial FOBT programme. No CRC mortality data was reported in this trial. An ongoing RCT was identified investigating this issue, the Norwegian NORCCAP trial, but data on CRC incidence and mortality from this RCT are not expected until late 2007. Three other RCTs, without CRC incidence and/or mortality data, compared the diagnostic accuracy of combined screening compared to either modality alone. One relevant systematic review commented on the lack of RCTs and concluded that it was not possible to support combined screening or the selection of one test over
another. Given the lack of RCT evidence for a mortality reduction, a national screening programme offering combined FS and FOBT screening cannot be justified currently.

The 15 primary research studies and three secondary research studies reviewed in this report provided a good indication of the current state of knowledge as to the cost-effectiveness of CRC screening programmes. Stronger evidence is available on the use of guaiac-based FOBT as the results of major RCTs based on guaiac-based FOBT have been made available, including reports of directly observed cost information. The available evidence provided a strong basis for modeling the longer-term costs and effects of a CRC screening programme based on annual FOBT. The results of the economic analysis of model results are generally favourable, suggesting that a screening programme based on annual guaiac-based FOBT is likely to be cost-effective. There are however, several major areas of uncertainty: there is a lack of understanding of the natural history of disease and the adenoma-carcinoma sequence; participation rates may be pivotal to achieving a favourable cost-effectiveness ratio and it is not known what participation rates will be; and, due to the lack of evidence on programme costs, the magnitude of administrative overhead costs, recall system costs, and health promotion costs are unknown.

Other screening modalities, such as immunochemical FOBT and flexible sigmoidoscopy, are subject to significantly more uncertainty than guaiac-based FOBT. These modalities have not been the subject of major RCTs and, consequently, evidence of their effectiveness and of the cost of running a screening programme based on these modalities is weak.

Although most studies suggested that screening for CRC is likely to be more cost-effective at saving life-years than cervical cancer screening or breast cancer screening, several studies pointed out that a CRC screening programme based on any screening modality will put additional pressure on the health system to increase the availability of colonoscopy and services used to investigate or follow-up on positive results. Any country considering a CRC screening programme must, therefore, consider whether the supply of such services can be increased to meet the expected increase in demand.
References


National Health and Medical Research Council (1999). *A guide to the development, implementation and evaluation of clinical practice guidelines*. Canberra: NHMRC.

National Health and Medical Research Council (2000). *How to use the evidence: assessment and application of scientific evidence*. Canberra: NHMRC.


StataCorp (2001). *Stata statistical software: release 7.0*. College Station, TX.: Stata Corporation.


Appendix 1

METHODOLOGY FOR CALCULATING DIAGNOSTIC TEST PERFORMANCE

The diagnostic test performance includes consideration of validity and reliability of the test. Specifically, sensitivity, specificity, and positive and negative predictive values were calculated when possible to assess the validity of each screening test. These measures were calculated based on presentation of results as shown in Table 20 below.

Table 20. Assessment of validity of a diagnostic test

<table>
<thead>
<tr>
<th>Diagnostic test</th>
<th>Reference test or true disease state</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>a</td>
<td>b</td>
<td>(n_1)</td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>c</td>
<td>d</td>
<td>(n_2)</td>
</tr>
<tr>
<td>Total sample size</td>
<td></td>
<td>(n_1)</td>
<td>(n_2)</td>
<td></td>
</tr>
</tbody>
</table>

Based on Table 20 above, measures of validity and 95 percent confidence intervals were calculated using the following formulae:

**Sensitivity**

\[
\text{Sensitivity} = \frac{a}{a+c} = \frac{a}{n_1}
\]

Confidence interval for sensitivity: 

\[
p \pm 1.96(pq/n_1)^{1/2}
\]

Where

\[
p = \frac{a}{a+c} \quad q = \frac{c}{a+c}
\]

**Specificity**

\[
\text{Specificity} = \frac{d}{b+d} = \frac{d}{n_2}
\]

Confidence interval for specificity: 

\[
p \pm 1.96(pq/n_2)^{1/2}
\]

Where

\[
p = \frac{d}{b+d} \quad q = \frac{b}{a+b}
\]

**Positive predictive value (PPV)**

\[
\text{Positive predictive value (PPV)} = \frac{a}{a+b}
\]

Confidence interval for PPV: 

\[
p \pm 1.96(pq/n_1)^{1/2}
\]

Where

\[
p = \frac{a}{a+b} \quad q = \frac{b}{a+b}
\]
Negative predictive value (NPV) = d/(c+d)

Confidence interval for NPV: p ± 1.96(pq/n)1/2

Where

\[ p = \frac{d}{(c+d)} \]
\[ q = \frac{c}{(c+d)} \]

If either n*p or n(1-p) were less than five for sensitivity or specificity, confidence intervals based on the normal approximation to the binomial distribution using the formulae above were considered unreliable and exact methods based on the binomial distribution were used to calculate the confidence interval. Stata version 7.0 was used for these calculations (StataCorp, 2001).
Appendix 2

SEARCH STRATEGIES

A core search strategy covering colorectal screening as comprehensively as possible was developed for each of the Medline and Embase databases using relevant subject headings and additional text words. A series of additional strategies looking at aspects of particular relevance was then linked to the core strategy. The Medline core strategy was also used for Cinahl and the Cochrane Central Register of Controlled Trials. The additional strategies were almost entirely text-word based and were therefore able to be used in all databases.

**Medline core search**

1. exp colorectal neoplasms/
2. (colorectal adj (carcino$ or adeno$)).tw.
3. (colon$ adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
4. (bowel adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
5. (sigmoid adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
6. ((caecum or cecum) adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
8. or/1-7
9. mass screening/ or screen$.
10. 12 and 13

**Embase core search**

1. (colorectal adj (carcino$ or adeno$)).tw.
2. (colon$ adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
3. (bowel adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
4. (sigmoid adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
5. ((caecal or cecal) adj (cancer or neoplas$ or carcino$ or adenocarcino$ or malignan$ or polyp$)).tw.
6. ((rectal or rectum) adj adenoma).tw.
7. (colorectal adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw.
8. (caecum adj (cancer or neoplas$ or carcin$ or adenocarcino$ or malignan$ or polyp$)).tw.
9. 12 or/1-9
10. mass screening/ or cancer screening/
11. or/14-16
12. 13 and 16
Medline trials/systematic reviews/meta-analyses filter

1 randomized controlled trial.pt.
2 meta-analysis.pt.
3 randomized controlled trials/ or meta-analysis/
4 controlled clinical trials/ or controlled clinical trial.pt.
5 exp clinical trials/ or clinical trial.pt.
6 random allocation/ or (random$ adj2 allocat$).tw.
7 single blind method/ or double blind method/
8 (clinical$ adj trial$).tw.
9 ((singl$ or doubl$ or tripl$) adj (blind$ or mask$ or dumm$)).tw.
10 (systematic adj (review$ or overview)).tw.
11 (meta-analysis$ or metaanaly$.tw.
12 exp review literature/
13 (hand search$ or relevant journals or manual search$ or selection criteria or data extraction).ab.
14 or/1-13
15 letter.pt.
16 case report.tw.
17 (historical article or review of reported cases or review, multicase).pt.
18 or/15-17
19 14 not 18
20 animal/
21 human/
22 20 not (20 and 21)
23 19 not 22

Cost-effectiveness strategy

1 exp "costs and cost analysis"/
2 ec.fs.
3 (economic evaluation$ or economic analy$ or health economi$).tw.
4 (cost containment$ or cost benefit$ or cost minimi$ or cost utili$ or cost effectiv$).tw.
5 or/1-4
6 5 and (core search results)
7 (letter or news).pt.
8 6 not 7
9 limit 8 to yr=1997-2004
Immunochemical tests strategy

1 Immunochemistry/
2 exp Immunologic Tests/
3 immunochemi$.tw.)
4 hemeselect.tw. (28)
5 (imdia-hem or imdiahem or (imdia adj hem)).tw.
6 bayerdetect.tw.
7 bayer detect.tw.
8 hemsp.tw.
9 immudia-hem Sp.tw.
10 monohem.tw.
11 iatro hemchek.tw.
12 la hemohasert.tw.
13 oc hemodia.tw.
14 flex sure obt.tw.
15 quicktest.tw.
16 dima fob-10.tw.
17 magstream hemsp.tw.
18 insure.tw.
19 bm-test colon albumin.tw.
20 inform.tw.
21 or/1-20
22 21 and (core search results)
23 animal/
24 animal/ and human/
25 23 not (23 and 24)
26 22 not 25

Guaiac tests strategy

1 guaiac/ or guaiac.mp.
2 hemoccult$.af.
3 hemo fec.mp.
4 haemoccult$.mp.
5 fecatest.mp.)
6 coloscreen$.mp.
7 fecatest$.mp.
8 fecatwin.mp.
9 hema fecia.mp.
10 hema fecia.mp.
11 haemafecia.mp.
12 shionogi b.mp
13 guajac.mp.
14 occult blood/
15 (occult blood or (fecal adj2 blood)).tw.
16 (blood adj2 feces) or (blood adj2 faeces)).tw.
17 fobt.tw
18 or/1-17
19 18 and (core search results)
20 (news or letter or editorial).pt.
21 19 not 20
22 limit 21 to yr=1997-2004

Flexible sigmoidoscopy strategy

1 exp colorectal neoplasms/ (34069)
2 (colorectal adj (carci$ or aden$)).tw. (5553)
(colon$ adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw. (1206)
(bowel adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw. (467)
(sigmoid adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw. (73)
((caecal or caecum) adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw. (28)
((rectal or rectum) adj (cancer or neoplas$ or malignan$ or carcino$ or adenocarcino$ or polyp$)).tw. (3821)
(colorectal adj (cancer or neoplas$ or malignan$ or polyp$)).tw. (13548)
((caecal or cecal) adj (cancer or neoplas$ or carcino$ or adenocarcino$ or malignan$ or polyp$)).tw. (69)
((rectal or rectum) adj adenoma).tw. (31)
((colorectal or colon or sigmoid or bowel or caecal or caecum or cecal or cecum) adj adenoma).tw. (343)
(or/1-11 (40495)
mass screening/ or screen$.tw. (110643)
12 and 13 (3590)
sigmoidoscopy/ (945)
(flexible adj5 sigmoid$).mp. (508)
15 or 16 (1120)
14 and 17 (589)
randomized controlled trial.pt. (97827)
meta-analysis.pt. (7442)
randomized controlled trials/ or meta-analysis/ (28842)
controlled clinical trials/ or controlled clinical trial.pt. (23865)
exp clinical trials/ or clinical trial.pt. (253098)
random allocation/ or (random$ adj2 allocat$).tw. (21953)
single blind method/ or double blind method/ (40127)
(clinic$ adj trial$).tw. (47834)
((singl$ or doubl$ or trebl$ or tripl$) adj (blind$ or mask$ or dumm$)).tw. (32165)
(systematic$ adj3 (review$ or overview$)).tw. (6798)
(meta-analy$ or metaanaly$).tw. (9285)
exp review literature/ (1696)
(hand search$ or relevant journals or manual search$ or selection criteria or data extraction).ab. (7690)
or/19-31 (309932)
letter.pt. (210235)
case report.tw. (41662)
(historical article or review of reported cases or review, multicase).pt. (80288)
or/33-35 (322324)
animal/ (1112832)
human/ (2921882)
37 not (37 and 38) (754966)
32 not (36 or 39) (284558)
exp epidemiologic studies/ (417939)
exp case control studies/ (167277)
exp cohort studies/ (252485)
cross-sectional studies/ (35789)
(case control or cohort analy$ or cross sectional).tw. (50178)
(longitudinal or retrospective).tw. (89808)
(cohort adj (study or studies)).tw. (15627)
((follow up or observational) adj (study or studies)).tw. (16036)
or/41-48 (459215)
18 and 40 (130)
18 and 49 (138)
50 or 51 (225)
Current Contents, Science/Social Science Citation Index searches

The platform used for these databases does not support lengthy search strategies and the databases do not have indexing. A simplified core strategy was developed as follows:

1 screen*
2 adenoma OR adenocarcino* OR polyp*
3 malignan* OR neoplas* OR carcino*
4 bowel OR sigmoid OR cecum OR cecal OR caecum OR caecal
5 rectum OR rectal OR colon* OR colorectal
6 #5 OR #6
7 #1 AND (#2 OR #3) AND #6

This basic search was linked with a simplified version of the additional strategies shown above for cost-effectiveness, immunochemical tests, flexible sigmoidoscopy and guaiac tests.

SEARCHES FROM OTHER SOURCES

In databases and all other sources without controlled vocabulary combinations of the index terms and additional keywords from the above strategies were used in the search.
Appendix 3

SOURCES SEARCHED

Bibliographic databases
Medline
Embase
Cinahl
Science Citation Index
Social Science Citation Index
Current Contents
Cochrane Central Register of Controlled Trials

Review databases
Cochrane Database of Systematic Reviews
ACP Journal Club
Database of Abstracts of Reviews of Effectiveness
NHS Economic Evaluation database
Health Technology Assessment database

Guidelines and Health Technology Assessment agencies
US National Guidelines Clearing House
UK National Electronic Library for Health Guidelines Finder
American Cancer Society
Health Council of the Netherlands
Swedish Council on Technology Assessment in Health Care
Finnish Office for Health Technology Assessment
Norwegian Center for Health Technology Assessment
Institute for Clinical Systems Improvement

Other websites
Current Controlled Trials
Controlled Trials.gov
Australian National Cancer Initiative
Australian Department of Health and Ageing Bowel Screening Pilot
NHS Cancer Screening Programmes
Bayer Australia/New Zealand
Australian Technology Showcase Annual Bowel Check

SOURCES SEARCHED

Bibliographic databases
Medline
Embase
Cinahl
Science Citation Index
Social Science Citation Index
Current Contents
Cochrane Central Register of Controlled Trials
**Review databases**
Cochrane Database of Systematic Reviews
ACP Journal Club
Database of Abstracts of Reviews of Effectiveness
NHS Economic Evaluation database
Health Technology Assessment database

**Guidelines and Health Technology Assessment agencies**
US National Guidelines Clearing House
UK National Electronic Library for Health Guidelines Finder
American Cancer Society
Health Council of the Netherlands
Swedish Council on Technology Assessment in Health Care
Finnish Office for Health Technology Assessment
Norwegian Center for Health Technology Assessment
Institute for Clinical Systems Improvement

**Other websites**
Current Controlled Trials
Controlled Trials.gov
Australian National Cancer Initiative
Australian Department of Health and Ageing Bowel Screening Pilot
NHS Cancer Screening Programmes
Bayer Australia/New Zealand
Australian Technology Showcase Annual Bowel Check
Appendix 4

RETRIEVED STUDIES EXCLUDED FOR REVIEW


---

**EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER**


Appendix 5

RETRIEVED STUDIES APPRAISED FOR REVIEW


EFFECTIVENESS AND COST-EFFECTIVENESS OF POPULATION SCREENING FOR COLORECTAL CANCER


Appendix 6

ACCURACY OF FOBT AND FS COMBINED SCREENING

Three RCTs were identified that examined the effect of combined FOBT and FS screening programmes on screening compliance, diagnostic yield and follow-up colonoscopic findings from once-only screening. No data were presented on CRC incidence and/or mortality in these studies. Below is an overview of study designs and aspects of quality represented by these studies. Full details of the three studies appraised, including methods, key results, limitations and conclusions, are provided in evidence tables (see Table 21 for studies comparing FOBT and FS combined with FOBT alone and Table 22 for the studies comparing FOBT and FS combined with FS alone and FOBT alone). There were no studies included comparing FOBT and FS combined screening with non screening. Studies are presented in chronological order of publication.

Study designs and quality assessments

Two RCTs investigated FOBT and FS combined, compared with FOBT alone. One RCT (Berry et al, 1997) was graded as level II evidence, with adequate design and conduct although the study methodology was not described in detail. The other RCT (Brevinge et al, 1997) was graded as level III-1 evidence due to limitations in study design and conduct, including pseudorandomisation, possible subject selection bias and a lack of methodological details. One RCT (Verne et al, 1998) compared FOBT and FS combined screening with FS alone, and FOBT alone. This RCT was graded as level II evidence and was of adequate quality with good design and conduct and methodology description but participant group results were not presented in a way to adequately compare effectiveness of the screening modalities.

FOBT and FS combined screening compared with FOBT screening alone

Berry et al. trial (1997)

Study setting and sample

This prospective RCT invited a sample of 6371 asymptomatic persons aged 50-74 years for CRC screening over a four-year period. Persons were identified from GP registers and were randomised by household into two groups. After exclusions, 3243 persons were invited to undergo one FOBT (Haemoccult, YS1, Farnborough, UK) and one flexible sigmoidoscopy (FS) and the other group of 3128 persons invited to undergo one FOBT (Haemoccult, YS1, Farnborough, UK) alone. Those invited to undergo CRC screening and found to have a positive FS and/or positive FOBT were invited to undergo full colonoscopy. The at-home FOB testing was done without dietary restriction or rehydration of the test. Those invited for a surgery-based FS as part of the combined FOBT and FS programme were given a single phosphate enema as bowel preparation and FS was performed without sedation.

Outcomes

The aim of this study was to ascertain the compliance and diagnostic yield of FOBT and FS compared with FOBT alone in asymptomatic individuals. Compliance with FOB testing was 50 percent in the FOB only group and 48 percent in the combined FOBT + FS group. Compliance with FS was only 20 percent.

Test positivity rate

Significant neoplasms were detected by FOBT alone in two persons per 1,000 screened (4.2 per 1000 in those who completed screening). Combination testing of FOBT and FS detected neoplasms in 8.9 persons per 1000 screened (44.2 per 1000 in those who completed screening).
Diagnostic test performance

There were no data reported from this trial regarding diagnostic test performance.

Stage CRC detected

There were no data reported from this trial regarding stage of CRC detected.

Outcomes related to screening

The focus of this trial was on ascertaining baseline diagnostic yield and compliance for combination FOBT + FS screening compared to FOBT alone. Even with low compliance, the diagnostic yield was at least four times greater with combined FOBT + FS screening than FOBT screening alone. Longer term follow-up data on screening complications and incidence and mortality data was not available from this trial. All lesions detected by positive FS in the combined FOBT + FS group were within reach of the flexible sigmoidoscope and no further lesions were detected within 60cm of the anal verge at colonoscopy following positive FS.

The Göteborg trial, Brevinge et al. (1997)

Study setting and sample

The purpose of this study was to compare screening with FOBT and FS of those with positive tests versus a direct primary examination with FS. Both compliance and diagnostic yield findings were evaluated. The trial, known as the Goteborg trial, invited a birth cohort of 6,397 citizens from Göteborg, Sweden born in 1938 and the first half of 1941 for CRC screening between September 1993 and the first six months of 1996 (thus all persons aged 55-56 years), (Brevinge et al, 1997). The trial cohort was randomised into two groups with one group of 3183 persons invited to undergo rehydrated FOB (Haemoccult II) tests and the other group of 3184 persons to undergo direct flexible sigmoidoscopy (FS). Those invited to undergo direct FS (n=3184) received an invitation and those from this group born in 1938 (n=2,113) also received a set of FOB slides for testing. The purpose of this was that a person should receive both FOB testing and FS or either one independently. Persons obtained FOB test results and depending on these decided, whether or not to undergo FS. Because of this no FOB testing was undertaken on those in the FS group born in 1941 to remove the bias caused by FOB testing on FS compliance. All persons received two reminder letters if required but those in the FOB testing group born in 1941 received only one reminder letter. The demographic characteristics of the two groups and sub-groups were not described and it was not specified if these were compared to assess potential differences between these groups.

The FOB testing group received one invitation letter and Haemoccoul II test kit (3 x 2 test spots). If a FOBT was positive in at least one of the slides, persons were asked for another set of FOBTs with more rigorous dietary restrictions. If the retest was positive, persons were offered FS examination and double-contrast enema (DCE). For the FS procedure (65cm Olympus CF 100 S video-endoscope), bowel preparation included colon cleaning with Bisacodyl (Toilax, Orion Farmos, Sweden) and low-fibre diet two days prior to examination with an additional microenema administered the evening/morning before the exam. A positive test was identified through removal of all polyps (≥5 mm) sent for histology and these patients were classified by their largest neoplasia. If two or more adenomas (≥5 mm) were removed and/or cancer found, the entire colon was examined by colonoscopy. A short questionnaire on medical history was also requested.

Outcomes

Compliance with screening in the FOB testing group with at least one test was 59 percent. In the FS group, 39 percent of those offered FOBT associated with FS (born in 1938) underwent FS compared with 49 percent for those for whom FOBT was omitted (born in 1941), p<0.001. Compliance rates for those invited for FOB testing compared to FS were significantly higher in both cohorts (those born in 1938, 61% versus 39%, p<0.001; those born in 1941, 55% versus 49%, p<0.001). Of those offered FOBT (n=2,113) in the FS group only 368 persons participated.
Test positivity rate

Of the 1893 persons in the FOB testing group who attended, 4 percent had a positive test and of these, 13 percent (95% CI 9-26%) (10/78) had a neoplasm $\geq$1cm diagnosed by endoscopy in the rectum or sigmoid colon. For the FS group, the proportion of positive tests was (31/1353 who attended FS) or 2.3 percent (95% CI 1.5-3.1%). The authors also examined the assumption that if the distribution of neoplasms ($\geq$1 cm) was the same in attending/not attending FS after FOBT the diagnostic yield would be 2.6 percent versus 2.8 percent actual in the FS group (cohort born 1938).

Diagnostic test performance

Using data from the FS group screeners who had both FOBT and FS it was possible to determine sensitivity and specificity for FOBT for a person with neoplasms $\geq$1cm within reach of FS. The sensitivity for retesting for FOBT with neoplasia of $\geq$1cm was 26 percent and specificity was 95.6 percent. This was 39 percent and 85.3 percent respectively for the first test only in the FOBT retest group. The risk ratio for finding a neoplasm $\geq$1cm by FS directly compared with first selecting the person to be examined by retesting with FOBT was 0.33 (95% CI 0.15-0.67). To find one neoplasm of $\geq$1cm in the rectum/sigmoid colon, 44 examinations were needed when FS was used directly, 13 FS examinations with one FOBT and seven FS examinations (with retesting) when any with positive FOBTs were examined.

Stage CRC detected

The number of persons diagnosed by FS with carcinomas was one (Duke’s A) in the FOBT group and five (four Duke’s A, one Duke’s B) in the FS group. The number of adenomas $\geq$1 cm was nine in the FOBT group and 26 in the FS group, and for adenomas $\geq$0.5cm - <1 cm, four and 44 respectively. For neoplasia $\geq$1 cm this was 10 and 31 cases respectively.

Outcomes related to screening

In this screening programme the number of subjects in the FOBT group undergoing flexible sigmoidoscopy was approximately one in 17 (78/1358) through the selection of persons based on positive FOBT (rehydrated/retesting) prior to FS. The authors concluded that FOBT (rehydrated/retesting) is a viable alternative to direct FS for all invited persons for screening. However, the lower sensitivity of FOBT for neoplasms $\geq$1 cm reduced advanced adenoma findings to only one third of that for diagnosis directly by FS. The rate of attendance for FS increased when not combined with a FOBT. Additionally compliance in the FOBT retesting group was lower in the 1941 birth cohort, probably because only one reminder letter was sent. In the FS group, selection bias was likely as subjects knew their test results prior to opting for FS and those with a negative FOBT who did not attend FS would affect the yield of positive FS findings. This RCT was limited in its evaluation as it included no cost-benefit nor comprehensive health outcomes analysis, e.g. incidence and mortality data, and was conducted over a short time period. The method of randomisation, concealment, and investigator involvement in outcomes evaluation was not described. Post-randomisation, the FS group was further broken down for screening intervention and evaluation purposes based on birth cohort into two groups of persons, one receiving FOBT and FS and the other only FS to help reduce non-compliance bias in the FOBT + FS group. However, this methodology may introduce bias into the randomisation status and comparisons.

FOBT and FS combined screening compared with FS only screening, and FOBT only screening

The Verne et al. trial (1998)

Study setting and sample

The purpose of this study was to assess the feasibility of combined screening by FS with FOB testing against each of FS and FOBT individually. Both compliance and diagnostic findings were evaluated. The trial (Verne et al, 1998) invited one general practice population of 3744 asymptomatic persons aged
between 50 and 75 years for CRC screening. The trial subjects (households) were randomised into three
groups with one group of 1249 persons invited by letter to undergo FS, another group of 1245 persons
invited to undergo FOB (non-rehydrated-Haemoccult II, Rohm Pharma) testing and another group of
1250 persons invited to undergo combined FOBT and FS. Reminders were not routinely sent. The
demographic characteristics of the three groups were not described and it was not specified if these were
compared to assess potential differences.

For FOB testing a three-day, six sample, diet-restricted, non-rehydrated regime was provided. Patients
were recalled if one or more samples yielded a positive result. For FS, an appointment was made two
weeks in advance and bowel preparation was undertaken with laxatives. Small polyps (<5 mm)
appearing hyperplastic and found only in the rectum were removed at screening and subjects recalled if
histology showed an adenoma. Those with other lesions were recalled for colonoscopy or surgery
depending on the finding. For the combined FS and FOB test screen group, subjects underwent FOB
before attending FS. The FOBT was developed blind to the results of the FS examination and vice
versa. Colonoscopy was performed by a GP and patients were classified as high/low risk based on the
lesions identified.

Outcomes

The crude uptake rate of screening with FS was 46.6 percent, which was, significantly higher than
FOBT with an uptake rate of 31.6 percent; combined (both) 30.1 percent and either FS, 37.6 percent or
FOBT 32 percent (all \( p < 0.001 \)).

Test positivity rate

Of the patients undergoing FOBT alone or in combination with FS, 0.8 percent (95% CI 0.2%-1.4%) of
patients had positive results and all underwent colonoscopy. Eleven (10%) of those undergoing
colonoscopy were found to have adenomas proximal to the sigmoid colon.

Diagnostic test performance

Neoplasm prevalence rates detected at screening were 0.4 percent for cancer and 6.8 percent for
adenomas. The FOB testing missed at least one CRC and 30 cases of adenoma which were found by FS
in the combined group. In the combined FOBT + FS group of 401 subjects who underwent both tests a
comparison of test performance was made, only one patient with a polyp seen at FS had a positive
FOBT, and in 81 subjects with negative FOBT results, polyps were identified at FS, including 30 with
adenomas and one with early cancer.

Stage CRC detected

Of 197 patients who had polyps and cancers removed from the sigmoid and rectum during either FS or
colonoscopy or both, histological diagnoses were 2.0 percent cancer, 38.8 percent adenomas, 47.2
hyperplastic and 12.2 percent miscellaneous abnormalities. Polyps were found in 19.3 percent (95% CI
17.0%-21.6%) of FS examinations, including 2.4 percent (95% CI 1.5%-3.3%) classified as ‘high risk’
adenomas and CRC was detected in four subjects. Of the four subjects had carcinoma (three Duke’s A,
one Duke’s B) of the sigmoid colon or rectum. Of the patients undergoing FOBT alone or in
combination with FS, of the seven with positive results, one had a Duke’s stage C rectal carcinoma, one
a 2cm adenoma and two had diminutive adenomas (one detected at screening, the other at colonoscopy).
The other three remaining subjects did not have neoplasia.

Outcomes related to screening

The authors concluded that the poor uptake and high numbers of colonoscopies with FS screening are
challenges which can be overcome through appropriate subject selection, follow-up and GP
involvement in the screening programme. Although there was a high colonoscopy rate in this study,
they argue that this can be reduced through the biopsy of polyps at screening, with subsequent recall
restricted to patients with adenomas detected. Recalling only ‘high risk’ patients would reduce the
colonoscopy rate further. The results showed that FS has significantly greater detection capability for
distal neoplasia than FOBT and the proportion of positive FOBT was low at 0.8 percent. The crude
uptake rates for FS were significantly higher than FOBT, indicating its acceptability in this study population. The addition of FOBT to FS detection rates was insignificant, adding nothing to neoplasia detection by FS. It was not possible to fully compare the screening effectiveness of FS versus combined FOBT + FS or FOBT only screening, as data on FOBT (alone or combination) were combined, as were diagnoses by FS or colonoscopy, or both. The method of randomisation and concealment and blinding of FOBT and FS results was well described in this study but no participant demographic information was presented and there was no comparison of baseline characteristics between the three randomised groups. The trial was primarily a small feasibility pilot RCT with a limited (one-off) screening programme evaluating uptake and diagnostic yield. The authors did not present longer-term data on CRC incidence and mortality.

CONCLUSIONS

Three RCTs investigating the diagnostic accuracy of one-time screening programmes with combined FOBT and FS screening compared to either test alone, were included. The outcomes focus of these trials was participant compliance, diagnostic yield, and colonoscopy utilisation and findings. There were no CRC incidence or mortality data evaluated in these RCTs. These RCTs were of average quality as there were some limitations in study design, conduct and presentation of findings.

Participant compliance for FS in combination testing with FOBT was low compared to FOBT alone or FS alone. This low compliance was associated with participants knowing their FOBT results prior to undergoing FS. Combination testing had significantly higher detection rates of neoplasms compared to FOBT alone, even with low FS compliance. There was insignificant additional diagnostic benefit from adding FOBT to FS compared to FS alone. The RCT by Brevinge et al. (1997) demonstrated that there was some benefit in reducing the rate of FS testing by prior testing with FOBT first before FS, compared to direct FS. In the RCT by Verne et al. (1998) the authors concluded that low compliance and high numbers of colonoscopies could be reduced through appropriate patient selection, with biopsy of polyps at screening, and the recall of patients with detected adenomas, and even reduced further through the recall of only ‘high risk’ patients.
Table 21. Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT only screening

<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Berry et al. (1997) United Kingdom | RCT Grade II | Combined FOBT + FS-based screening group  
Once-only Guaiac FOBT (Hemoccult, YS1, Farnborough) 
Non-rehydrated and not diet-restricted followed with once-only Flexible Sigmoidoscopy (FS). 
Those with positive FOBT and/or FS were invited for medical review and full colonoscopy.  
FOBT-only screening group  
Once-only Guaiac FOBT (Hemoccult, YS1, Farnborough) 
Non-rehydrated, not diet-restricted | Over a four-year period in Newport, South Wales, 6371 asymptomatic individuals aged 50-74 years identified from GP registers were randomised using standard random number generation by household into either once-only combined FOBT + FS screening (3243 persons) or once-only FOBT alone (3128 persons) screening.  
No differences in the age/sex demographic profile of the two groups. 
Contact with screenees was by GP letter and also included the FOBT tests. Subjects randomised for FS were invited to the surgery for information on the test. Two follow-up reminders were sent to non-responders. 
Excluded: 
GPs excluded unsuitable subjects from the register such as persons with CRC, those under abdominal investigation and people with advanced disease. | Diagnostic yield of invited/screen detected CRC and adenomas  
Numbers undergoing colonoscopy  
Participation of those invited for screening round for each of FOBT alone, FS, FOBT. | Yield of CRC/adenomas  
Similar numbers of patients detected by FOBT in both groups with large adenomas (>1cm) and/or carcinomas: 
FOBT + FS combined group seven patients (7/3243 with 11 lesions) 
FOBT alone group six patients (6/3128 with nine lesions). 
All patients with positive FOBT results in the combined group went on to colonoscopy, as did a further 22 patients in this group with negative FOBT results who had 24 lesions identified by FS. 
Significant neoplasms were detected by FOBT alone in 2.0 persons per 1,000 screened (4.2 per 1000 in those who completed screening). Combination testing of FOBT and FS detected neoplasms in 8.9 persons per 1,000 screened (44.2 per 1,000 in those who completed screening) | Limitations  
No mortality or incidence data, non-repeat screening, short screening period, no description of test validation methods. No reported data on direct screen harms. No detailed description of sample power calculations, randomisation, concealment and investigator blinding from test results.  
Authors’ conclusions  
Despite poor compliance, FS increases the diagnostic yield of neoplasia and there is potential to increase this further by implementing strategies to increase compliance with FS.  
Reviewers’ conclusions  
Small study assessing one-time screening regimen. Clearly higher diagnostic yield with combined FOBT + FS screening than FOBT alone, despite a lack of FS compliance. Limitations in methodology and study design and no mortality and incidence data because of short-term screening period means it is not possible to adequately assess the impact of combined screening on health outcomes such as CRC incidence/mortality. |
<table>
<thead>
<tr>
<th>Source Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brevinge et al. (1997) Sweden</td>
<td>RCT</td>
<td>Grade III-1</td>
<td>Combined FOBT + FS screening group Direct flexible sigmoidoscopy (FS). From those invited to undergo FS those from this group born in 1938 also received a set of FOB slides for testing. The purpose of this was that a person should receive both FOBT testing and FS or either one independently. Persons obtained FOBT results, and depending on these, decided whether or not to undergo FS. No FOBT testing was undertaken on those in the FS group born in 1941 as a way to remove the bias caused by FOBT testing on FS compliance. Those with positive FOBT and/or FS were invited for medical review and full colonoscopy. FOB-only screening group Guaiac-based FOBT (Hemoccult II, Smith Kline Diagnostics, Ca, US) test kit with 3 x 2 guaiac spots. Rehydrated and not diet-restricted, followed with repeat testing with more rigorous dietary restrictions for those with a positive first test. If the retest was positive, persons were offered FS examination and double-contrast enema (DCE). CRC screening was undertaken between September 1993 and first six-months of 1996 thus all enrolled persons were aged 55-56 years.</td>
<td>The trial invited a birth cohort of 6397 citizens from Göteborg, Sweden born in 1938 and the first half of 1941 for CRC screening. The trial cohort was randomised into two groups with one group of 3183 persons invited to undergo rehydrated FOBT (Haemoccult II) tests and the other group of 3184 persons to undergo direct flexible sigmoidoscopy (FS). Those invited to undergo FS were further assigned based on a subgroup born in the 1938 cohort (n=2113) to receive a set of FOB slides for testing as well. No assessment of differences in the age/sex demographic profile of the two groups and subgroups. All persons were invited by letter and received two reminder letters if required but those in the FOBT testing group born in 1941 received only one reminder letter. Excluded: No criteria specified.</td>
<td>Diagnostic yield of invited/screen detected CRC and adenomas Risk ratio of detecting neoplasia with FS directly and combined FS + FOBT. Sensitivity/specificity of retesting and single FOBT. Compliance rates for FS and for first and second-round FOBT and in conjunction with FS. Screening harms In the FOBT group if a FOBT was positive in at least one of the slides persons were asked for another set of FOBTs with more rigorous dietary restrictions. If the retest was positive, persons were offered FS examination and double-contrast enema (DCE).</td>
<td>Diagnostic yield of CRC/adenomas For the FS group the proportion of positive tests was 31/1353 (who attended FS) or 2.3% (95% CI 1.5-3.1%). Of the 1893 persons in the FOB testing group who attended, 4% had a positive test and of these, 13% (95% CI 9.2-26%) (10/78) had a neoplasm ≥1 cm diagnosed by endoscopy in the rectum or sigmoid colon. Two patients who had profuse bleeding following polypectomy and one patient had moderate diverticulitis after endoscopy. The sensitivity for neoplasia of ≥1 cm with FOBT without retesting was 39% and specificity 85.3% and for retesting with FOBT was 26% and 95.6% respectively. The risk ratio of finding a neoplasm ≥1 cm by FS directly versus after first selecting the person to be examined by retesting with FOBT was 0.33 (95% CI 0.15-0.67%). One neoplasm of ≥1 cm in the rectum/sigmoid colon was identified with every 44 examinations with direct FS.</td>
<td>Limitations: No incidence/mortality data, short screening period, no description of sample power calculations, randomisation, concealment methods, investigator involvement. No comparison made of group/sub-group differences and likely bias from attempt to reduce selection bias from FS group participant’s refusal to undergo FS after negative FOBT. Authors’ conclusions This study compared screening with retesting of positive FOBTs and direct FS and prior FOBT + FS. Compliance and diagnostic yield findings were evaluated. The authors concluded that FOBT (rehydrated/retesting) is a viable alternative to direct FS for all invited persons for screening. With a low prevalence of colorectal neoplasia in the younger 55-56 year old age-group considered, screening of the rectum and sigmoid colon using FS can be reduced to an acceptable rate (1 in 17) through the selection of persons based on a prior positive FOBT testing (rehydrated/retesting). The lower sensitivity of FOBT for neoplasms ≥1cm reduced findings to one third of that for diagnosis directly by FS.</td>
</tr>
</tbody>
</table>
Table 21. Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with FOBT only screening (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Country</th>
<th>Study design Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brevinge et al.</td>
<td>Sweden</td>
<td>(1997)</td>
<td>For the FS procedure (65cm Olympus CF 100 S video-endoscope) bowel preparation included colon cleaning with Bisacodyl (Toilax, Orion Farmos, Sweden) and low-fibre diet two days prior to examination with an additional microenema administered the evening/moning before the exam. A positive test was identified through removal of all polyps (≥5 mm) sent for histology and these patients were classified by their largest neoplasm. If two or more adenomas (≥5 mm) were removed and/or cancer found the entire colon was examined by colonoscopy. A short questionnaire on medical history was also requested.</td>
<td>There were 69 persons (9 refused) from the FOBT (retest) group who had an additional DCE after retesting positive with FOBT and FS. One cancer and one large adenoma was found beyond reach of FS. In the FS group one person out of 36 was found with a large adenoma in the right colon. Compliance to screening in the FOBT testing group with at least one test was 59%. In the FS group 39% of those offered FOBT associated with FS (born in 1938) underwent FS compared with 49% for those for whom FOBT was omitted (born in 1941), p&lt;0.001. Compliance rates for those invited for FOBT testing (FOBT testing group) compared to FS was significantly higher in both cohorts (those born in 1938, 61% versus 39%, p&lt;0.001; those born in 1941, 55% versus 49%, p&lt;0.001). Of those offered FOBT (n=2113) in the FS group only 368 persons participated. The authors examined the assumption that if the distribution of neoplasms (≥1cm) was the same in attending/not attending FS after FOBT the incidence rate would be 2.6% versus 2.8% actual in the FS group (cohort born 1938).</td>
<td>Reviewers’ conclusions: The true effect of screening with FS is likely to be underestimated because of selection bias arising through knowing prior FOBT results. There were limitations in methodology and study design. No mortality data was presented in this study that investigated a short-term screening period. The rate of attendance for FS increased when not combined with a FOBT. This RCT was limited in its evaluation as it included no cost-benefit nor comprehensive health outcomes analysis e.g. CRC incidence/mortality and was conducted over a restricted time period.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 22. Primary research studies appraised investigating FOBT and FS combined screening for CRC compared with, FS only screening, and FOBT only screening

<table>
<thead>
<tr>
<th>Source/Country</th>
<th>Study design</th>
<th>Evidence Grading</th>
<th>Comparison interventions and dates of testing</th>
<th>Sample</th>
<th>Outcomes and verification</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verne et al. (1998)</td>
<td>RCT</td>
<td>Grade II</td>
<td>FS-screening-only group Direct flexible sigmoidoscopy (FS). Combined FOBT + FS screening group Flexible sigmoidoscopy (FS) plus Guaiac-based FOBT (Hemoccult II, Rohm Pharma) test with 3 day, six sample diet restricted non-rehydrated regime. Subjects underwent FOBT prior to FS. FOBT only screening group Guaiac-based FOBT (Hemoccult II, Rohm Pharma) test with 3 day, six sample diet restricted non-rehydrated regime. Those with positive FOBT of one or more samples positive were recalled for full colonoscopy. For FS, subjects were given an appointment 2-weeks in advance and those with small polyps and other lesions were recalled for full colonoscopy based on histology findings. The FOBT was developed blind to the results of FS examination.</td>
<td>3744 asymptomatic subjects aged 50-75 years were randomised into three groups with one group of 1249 persons invited to undergo flexible sigmoidoscopy (FS), another group of 1,245 persons randomised to FOBT only (non-rehydrated, Haemoccult II) and the third group of 1250 persons randomised to undergo combined flexible sigmoidoscopy (FS) and FOBT. No description and assessment of differences in the age/sex demographic profile of the three groups. All persons were invited by letter and reminder letters were not routinely sent. Excluded: Deceased, moved patients, previous diagnosis of CRC, prior endoscopy within past 2-years were removed from GP lists prior to randomisation.</td>
<td>Diagnostic yield of positive and negative tests. Findings at colonoscopy. Diagnostic yield and prevalence of detected CRC and adenomas. Uptake rates for FOBT, both FOBT and FS, either FOBT or FS, FS. Complications With FOBT a test was positive if at least one of the slides was positive. For the FS procedure bowel preparation included colon cleaning with laxatives. Small polyps (&lt;5 mm) appearing hyperplastic and found only in the rectum were removed at screening and subjects were recalled if histology showed an adenoma. Those with other lesions were recalled for colonoscopy or surgery depending on the findings.</td>
<td>Diagnostic yield of CRC/adenomas Of 197 patients who had polyps and cancers removed from the sigmoid and rectum at either FS or colonoscopy or both, diagnoses were 2.0% cancer, 38.8% adenomas, 47.2 hyperplastic and 12.2% misc. Eleven (10%) of those undergoing colonoscopy were found to have adenomas proximal to the sigmoid colon. Neoplasm prevalence rates detected at screening were 0.4% for cancer and 6.8% for adenomas. The crude uptake rates for screening were FS, 46.6% compared to FOBT, 31.6%; combined (both) 30.1% and either FS, 37.6% or FOBT 32% (all p&lt;0.001). A total of 1116 subjects underwent FS without complications. No specific data on complications. Polyps were found in 19.3% (95% CI 17.0%-21.6%) of FS examinations, 6.3% (95% CI 5.3%-8.3%) adenomas and 2.4% (95% CI 1.5%-3.3%) classified as “high risk” adenomas and CRC was detected in four subjects.</td>
<td>Limitations No incidence/mortality data, short screening period, no comparison made of group differences, results not adequately presented to enable comparison of diagnostic yield of FOBT alone, FS alone and FOBT + FS combined screening. Authors’ conclusions This study compared the feasibility of screening FS compared to FOBT and both tests combined. The main outcomes were uptake and diagnostic findings. The poor uptake and high numbers of colonoscopies with FS screening could be addressed with appropriate patient selection, follow-up and GP involvement in the screening programme. A reduction in the high rate of colonoscopies through the biopsy of polyps at screening, with subsequent recall being limited to patients with adenomas, would reduce the colonoscopy rate.</td>
</tr>
<tr>
<td>Source Country</td>
<td>Study design Evidence Grading</td>
<td>Comparison interventions and dates of testing</td>
<td>Sample</td>
<td>Outcomes and verification</td>
<td>Results</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------</td>
<td>---------------------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Verne et al. (1998)</td>
<td>United Kingdom</td>
<td>(Continued)</td>
<td></td>
<td></td>
<td>Of the patients undergoing FOBT alone or in combination with FS, 0.8% (95% CI 0.2%-1.4%) of patients had positive results and all underwent colonoscopy.</td>
<td>The FOBT testing missed at least one CRC and 30 cases of adenoma which were found by FS in the combined group.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The authors estimated that the high rate of 20% for colonoscopy could be reduced to 7% (95% CI 5.3-9.3) if polyps were biopsied at screening and only those patients with adenomas were recalled and lower if only ‘high risk’ patients were recalled.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>For the combined FOBT + FS group of 401, only one patient polyp seen at FS had a positive FOBT, and in 81 subjects with negative FOBT results, polyps were identified at FS.</td>
<td>The trial results were not analysed in a way to adequately evaluate the diagnostic effectiveness of FS versus combined FOBT + FS or FOBT-only screening. FOBT data (alone or combination with FS) were combined, as were diagnoses by FS or colonoscopy, or both. The study methodology was described, with method of randomisation and concealment and blinding of FOBT and FS results. The trial is primarily a small feasibility RCT with a limited ‘one-off’ screening programme with no longer-term data on cost-effectiveness or CRC incidence and mortality to evaluate the effect of combined screening on health outcomes.</td>
<td></td>
</tr>
</tbody>
</table>