CONFIDENTIAL UNTIL PUBLISHED

Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence

Virtual chromoendoscopy for the real-time assessment of colorectal polyps in vivo: a systematic review and economic evaluation

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors

Dr Joanna Picot, Senior Research Fellow, SHTAC

Mr Micah Rose, Research Fellow (Health economics), SHTAC

Dr Keith Cooper, Senior Research Fellow (Health economics), SHTAC

Dr Karen Pickett, Research Fellow, SHTAC

Professor Joanne Lord, Director and Professorial Fellow in Health Economics, SHTAC

Ms Petra Harris, Research Fellow, SHTAC

Dr Sophie Whyte, Research Fellow, School of Health and Related Research (ScHARR), The University of Sheffield

Professor Dankmar Böhning, Professor in Medical Statistics, Director of Southampton Statistical Sciences Research Institute, University of Southampton

Dr Jonathan Shepherd, Principal Research Fellow, SHTAC

Correspondence to

Dr Joanna Picot

Senior Research Fellow

Southampton Health Technology Assessments Centre (SHTAC)

University of Southampton

First Floor, Epsilon House

Enterprise Road, Southampton Science Park

Southampton, SO16 7NS, UK.

Tel: +44(0)23 8059 5921

Fax:+44(0)23 8059 5639

email: j.picot@soton.ac.uk

www.southampton.ac.uk/shtac

Date completed 17 October 2016

Word count: 64,668 (excluding appendices)

Data sharing statement: All data relevant to this technology assessment report are provided in the accompanying appendices, or may be obtained upon request from the corresponding author. Note that the current report does not include confidential data which were considered during the NICE diagnostics assessment. Confidential data cannot be shared, but their implications for the conclusions of the diagnostic assessment are clearly stated in the current report.

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 15/17/05 and will be published in full in *Health Technology Assessment* (http://www.journalslibrary.nihr.ac.uk/hta).

Declared competing interests of authors

All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). Sophie Whyte declares consultancies fees from SHTAC for her work on the economic evaluation. All authors within SHTAC declare 1) no financial support for the submitted work from anyone other than their employer; 2) no financial relationships with commercial entities that might have an interest in the submitted work; 3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; and 4) no non-financial interests that may be relevant to the submitted work.

Acknowledgements

Thanks to Karen Welch, Information Specialist at SHTAC, for developing the search strategy, conducting the literature searches and downloading the search results into a bibliographic database. We are also grateful to Geoff Frampton (SHTAC) for acting as an internal editor for the draft report. We also thank the members of the project's Advisory Group who were contacted for clinical advice and comments on draft versions of the protocol, project report and economic model: Professor Pradeep Bhandari, Consultant Gastroenterologist and Professor of Gastrointestinal Endoscopy, Portsmouth Hospitals NHS Trust; Dr Philip Boger, Consultant in Gastroenterology and Advanced Endoscopy, University Hospitals Trust, Southampton; Professor Brian Saunders, Consultant Gastroenterologist and Specialist Gastrointestinal Endoscopist, Wolfson Unit for Endoscopy, St Mark's Hospital and Academic Institute, London; Professor John Schofield, Consultant Histopathologist, Maidstone and Tunbridge Wells NHS Trust.

Additionally, the following NICE Specialist Committee Members who provided comments on a draft version of the report: Dr James East, Consultant Gastroenterologist and Endoscopist, John Radcliffe Hospital; Mrs Susan McConnell, Nurse Endoscopist/Training Lead, County Durham & Darlington Foundation Trust; Dr Morgan Moorghen, Consultant Histopathologist, St Mark's Hospital; Dr Venkat Subramanian, Clinical Associate Professor and Consultant Gastroenterologist, Leeds Institute of Biomedical and Clinical Sciences/ St James University Hospital.

This project was supported by the Complex Reviews Support Unit, which is funded by the National Institute for Health Research (project number 14/178/29). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health.

Rider on responsibility for the report

The views and opinions expressed in this report are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, NHS or the Department of Health. Any errors are the responsibility of the authors.

This report should be referenced as follows: Virtual chromoendoscopy for the real-time assessment of colorectal polyps in vivo: a systematic review and economic evaluation. Southampton Health Technology Assessments Centre (SHTAC), 2016.

Contribution of authors

Dr Joanna Picot (Senior Research Fellow, evidence synthesis) project managed the study, developed the research protocol, assisted in the development of the search strategy, assessed test accuracy studies for inclusion, performed data extraction and critical appraisal of included test accuracy studies, synthesised evidence including conducting the meta-analyses, drafted and edited the final report.

Mr Micah Rose (Research Fellow, health economics) developed the research protocol, assessed cost-effectiveness and HRQoL studies for inclusion, synthesised evidence, developed the economic model, drafted and edited the final report.

Dr Keith Cooper (Senior Research Fellow, health economics) developed the research protocol, assessed cost-effectiveness and HRQoL studies for inclusion, synthesised evidence, led the development of the economic evaluation, drafted and edited the final report.

Dr Karen Pickett (Research Fellow, evidence synthesis) developed the research protocol, assessed test accuracy studies for inclusion, performed data extraction and critical appraisal of included test accuracy studies, synthesised evidence, drafted and edited the final report.

Professor Joanne Lord (Professorial Fellow in Health Economics) contributed to discussions on the design of the economic model and drafted and edited the final report.

Ms Petra Harris (Research Fellow, evidence synthesis) performed data extraction and critical appraisal of included test accuracy studies, synthesised evidence, drafted and edited the final report.

Dr Sophie Whyte (Research Fellow, health economics) contributed to the development of the economic evaluation, provided data for the economic model, drafted and edited the final report.

Professor Dankmar Böhning (Professor in Medical Statistics) provided training and guidance in the conduct of meta-analyses of diagnostic studies and edited the final report.

Dr Jonathan Shepherd (Principal Research Fellow, evidence synthesis) developed the research protocol, assisted in the development of the search strategy, assessed test accuracy studies for inclusion, performed data extraction and critical appraisal of included test accuracy studies, synthesised evidence, drafted and edited the final report, and acted as the project guarantor.

Keywords
virtual chromoendoscopy; diminutive colorectal polyps; real-time assessment; diagnostic accuracy; cost-effectiveness; economic evaluation; colorectal cancer; NBI; Narrow band imaging; FICE; Flexible Spectral Imaging Colour Enhancement; i-scan

ABSTRACT

Background: Current clinical practice is to remove a colorectal polyp detected during colonoscopy and determine whether it is an adenoma or hyperplastic by histopathology. Identifying adenomas is important because they may eventually become cancerous if untreated, whereas hyperplastic polyps do not usually develop into cancer. Virtual chromoendoscopy (VCE) (an electronic endoscopic imaging technique) could be used by the endoscopist under strictly controlled conditions for real-time optical diagnosis of diminutive (≤5 mm) colorectal polyps to replace histopathological diagnosis.

Objective: To assess the clinical-effectiveness and cost-effectiveness of the VCE technologies Narrow band imaging (NBI), Flexible Spectral Imaging Colour Enhancement (FICE), and i-scan for the characterisation and management of diminutive (≤5mm) colorectal polyps using high definition systems without magnification.

Design: Systematic review and economic analysis

Participants: People undergoing colonoscopy for screening or surveillance or to investigate symptoms suggestive of colorectal cancer

Interventions: NBI, FICE and i-scan

Main outcome measures: diagnostic accuracy; recommended surveillance intervals; health-related quality of life (HRQoL), adverse effects, colorectal cancer, mortality, cost-effectiveness of VCE compared with histopathology.

Data sources: Electronic bibliographic databases including MEDLINE, EMBASE, The Cochrane Library and DARE were searched for English language published studies from inception to June 2016. Bibliographies of related papers, systematic reviews and company information were screened and experts were contacted to identify additional evidence.

Review methods: Systematic reviews of test accuracy and economic evaluations were undertaken according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Meta-analyses were conducted where possible to inform the independent economic model. A cost-utility decision analytic model was developed to estimate the cost-effectiveness of VCE compared with histopathology. The model used a decision tree for patients undergoing endoscopy, combined with estimates of long-term outcomes derived from the ScHARR Bowel Cancer Screening model. The model took a National Health Service (NHS) perspective, with costs and benefits discounted at 3.5% over a lifetime horizon. There were limitations in the data on the distribution of adenomas across risk categories, and recurrence rates post-polypectomy.

Results: Thirty test accuracy studies were included: 24 for NBI, five for i-scan and three for FICE (two studies assessed two interventions). Two economic evaluations were included. NBI and i-scan are dominant strategies compared to histopathology, i.e. they are cost saving and more effective. FICE is cost

effective compared to histopathology. The correct surveillance interval would be given to 95% of patients with NBI, 94% of patients with FICE and 97% of patients with i-scan.

Limitations: Limited evidence was available for i-scan and FICE and there was heterogeneity among the NBI studies. There is a lack of data on longer-term health outcomes of patients undergoing VCE for assessment of diminutive colorectal polyps.

Conclusions: VCE technologies, using high definition systems without magnification, could potentially be used for the real-time assessment of diminutive colorectal polyps, if endoscopists have adequate experience and training

Future work: Head-to-head RCTs of the three VCE technologies and more research on the diagnostic accuracy of FICE and i-scan. Longitudinal data on colorectal cancer incidence, HRQoL and mortality.

Study registration: PROSPERO CRD42016037767

Funding details: The National Institute for Health Research Health Technology Assessment programme

Word count: 508 words

TABLE OF CONTENTS

TABLES

FIGURES

Scientific Summary

Background

Colorectal polyps are small growths on the lining of the colon or rectum. They are common, particularly in people over 60 years of age and they do not usually cause symptoms. Histology can distinguish between polyps that are adenomas and those that are hyperplastic. It is important to identify adenomas because these polyps may eventually become cancerous if undiagnosed and untreated whereas hyperplastic polyps usually do not carry a risk of developing into cancer.

Current clinical practice is to detect colorectal polyps during a colonoscopy when the colon and rectum are examined using conventional white light endoscopy. Dyes may also be used (chromoendoscopy) to enhance visualisation of tissues being inspected. Usually, each detected polyp is removed (by polypectomy) and sent for histopathological examination to determine whether it is an adenoma or hyperplastic.

An addition to conventional white light endoscopy is virtual chromoendoscopy (VCE), an electronic imaging technique that enables the endoscopist to differentiate between adenomatous and hyperplastic colorectal polyps in real-time during colonoscopy (optical assessment). There are three commercial systems of relevance to this diagnostic assessment report: Narrow band imaging (NBI), Flexible Spectral Imaging Colour Enhancement (FICE), and i-scan. There have been proposals suggesting that virtual chromoendoscopy can be used, under strictly controlled conditions, for real-time optical diagnosis of diminutive ($\leq$5 mm) colorectal polyps to replace histopathological diagnosis. The features of these propsals are typically that when the endoscopist has high confidence in the diminutive polyp characterisation, adenomas would be removed and discarded (i.e. not sent to histopathology), whereas hyperplastic polyps would be left in situ (because the risk for colorectal cancer is very low). When the endoscopist cannot confidently characterise a polyp, it would be resected and sent for histopathological examination. The potential benefits of virtual chromoendoscopy, include: fewer polyp resections and possible reduction in associated complications (e.g. bleeding and bowel perforation), patients receiving results faster (so less anxiety associated with waiting for results), and a reduction in health care resource use (e.g. fewer histopathological examinations). However, a potential downside of VCE is that it is not as accurate as histopathology, and so some adenomas may be missed and then left in situ, potentially developing into cancer. For VCE to be incorporated into clinical practice for the real-time assessment of polyps, evidence is needed that it provides an appropriate and efficient standard of care compared to existing practice.

Objectives

To determine, through a systematic review and economic evaluation, the clinical-effectiveness and cost-effectiveness of the virtual chromoendoscopy technologies NBI, FICE, and i-scan for the characterisation and management of diminutive (≤5mm) colorectal polyps.

Methods

Systematic review of clinical-effectiveness

We undertook a systematic review of studies assessing diagnostic accuracy and other health outcomes when NBI, FICE and i-scan are used to characterise the histology of diminutive colorectal polyps in real-time. A comprehensive search strategy was designed to capture relevant clinical-effectiveness and cost-effectiveness studies. We searched the following databases from inception to June 2016: MEDLINE, PreMedline In-Process & Other Non-Indexed Citations, EMBASE, Web of Science, the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effectiveness (DARE), Health Technology Assessment database, and the NHS Economic Evaluation Database (EED). We also identified publications through conference proceedings, websites, bibliographies of included studies and relevant systematic reviews, and our advisory group. Studies were eligible for the review if they were randomised controlled trials (RCTs), prospective longitudinal cohort or cross-sectional studies that evaluated NBI, i-scan or FICE (using high definition endoscopy systems, without magnification) for the real-time diagnosis of diminutive colorectal polyps in people undergoing colonoscopy for screening or surveillance or to investigate symptoms suggestive of colorectal cancer. The reference standard was histopathology with at least one of the following outcomes reported: diagnostic accuracy; number of polyps designated to be left in place, resected, discarded, or sent to histopathology; recommended surveillance intervals; examination time; number of medical consultations; health-related quality of life (HRQoL, including anxiety), adverse effects of polypectomy, colorectal cancer and mortality. We assessed the risks of bias of the included studies using the QUADAS instrument (Quality Assessment Tool for Diagnostic Accuracy Studies) and narratively synthesised included studies. We conducted bivariate meta-analyses, where possible, to provide pooled estimates of diagnostic sensitivity and specificity for each technology. An advisory group of four independent experts was invited to comment on the protocol and draft report.

Systematic review of economic studies

A systematic review of cost-effectiveness studies was conducted to identify relevant evidence to inform the economic evaluation. The review used the same set of references identified in our systematic review of diagnostic accuracy with an additional filter using the keyword ‘cost’. Studies were included if they

were a full economic evaluation that included long-term outcomes such as the incidence of colorectal cancer, life years or Quality Adjusted Life Years (QALYs).

Economic evaluation

We developed an independent cost-utility decision analytic model to estimate the cost-effectiveness of virtual chromoendoscopy to optically characterise diminutive polyps compared with histopathology. The model used a decision tree for patients undergoing endoscopy, combined with estimates of long-term outcomes derived from the ScHARR Bowel Cancer Screening model (SBCS). The decision tree follows a cohort of patients who receive endoscopy and who have at least one diminutive polyp identified (and no non-diminutive polyps). For the histopathology strategy, all diminutive polyps identified are resected and sent to histopathology. In the base case analysis for virtual chromoendoscopy, polyps characterised with low confidence are resected and sent to histopathology, polyps characterised with high confidence as a hyperplastic are left in situ whereas those characterised as an adenoma are resected and discarded (i.e. not sent to histopathology). The model uses the diagnostic accuracy estimates for virtual chromoendoscopy from our systematic review of diagnostic accuracy. In the long-term SBCS model, patients progress through the development of adenomas, colorectal cancer and subsequent death. Costs are included in the model for colonoscopy, histopathology, adverse events from colonoscopy (polypectomy) and the costs of treating colorectal cancer. Health outcomes are quantified in terms of incremental QALYs, including mortality and impacts on HRQoL associated with adverse effects of polypectomy and colorectal cancer. Costs and benefits are discounted at 3.5% per annum. The perspective of the analysis is that of the NHS and Personal Social Services. The model uses a lifetime horizon and reports results as costs per QALY gained.

Results

Clinical-effectiveness

From 2070 titles and abstracts screened, 125 full texts were retrieved for detailed examination. The 32 references which met the inclusion criteria described 30 separate studies. Most studies evaluated NBI (n=22) with an additional two studies also evaluating one of the other interventions of relevance (NBI and i-scan, NBI and FICE). Four further studies evaluated i-scan and two further studies evaluated FICE. We assessed the studies to be generally at a low risk of bias across the domains measured by the QUADAS.

The ability of NBI to correctly identify diminutive polyps as adenomas (i.e. the sensitivity of the test) in the whole colon ranged from 55% to 97% (17 studies) for all assessments regardless of endoscopist confidence. For high confidence characterisations, sensitivity ranged from 59% to 98% (13 studies) for

the whole colon, and from 83% to 96% (five studies) for high confidence characterisations in the rectosigmoid colon. The ability of NBI to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) was typically lower, ranging from 62% to 95% (16 studies) for all assessments in the whole colon, from 44% to 92% (11 studies) for high confidence characterisations in the whole colon and from 88% to 99% (five studies) for high confidence characterisations in the rectosigmoid colon. A bivariate meta-analysis using available data (16 of the 24 NBI studies), produced a summary value for sensitivity of 0.88 (95% CI 0.83 to 0.92) (i.e. 88%) and for specificity of 0.81 (95% CI 0.75 to 0.85) for all characterisations in the whole colon. Bivariate meta-analysis of high confidence NBI characterisations in the whole colon produced summary values for sensitivity of 0.91 (95% CI 0.85 to 0.95) and for specificity of 0.82 (95% CI 0.76 to 0.87) (11 studies), and for high confidence characterisations in the rectosigmoid colon summary values for sensitivity of 0.87 (95% CI 0.80, 0.92) and for specificity of 0.95 (95% CI 0.87, 0.98) (four studies). We found that endoscopists with prior experience of using NBI to characterise diminutive colorectal polyps achieved higher sensitivity and specificity than endoscopists with no prior experience of using NBI.

The five included studies evaluating i-scan varied in how they reported results. One reported results for all polyp assessments in the whole colon, and four reported assessments made in particular parts of the colon. Sensitivity was above 90% in four studies (range: 93% to 95%) and was 82% in a study that used a per patient (rather than per polyp) analysis. Specificity ranged from 83% to 96%. Sensitivity and specificity for high confidence assessments ranged from 94% to 98% and 90% to 96%, respectively. A bivariate meta-analysis of two studies reporting on high confidence characterisations of polyps in the whole colon produced a summary sensitivity of 0.96 (95% CI 0.92 to 0.98) and specificity of 0.91 (95% CI 0.84 to 0.95).

The three included studies evaluating FICE assessed polyps in any part of the colon and did not provide analyses by confidence level. Sensitivity and specificity ranged from 74% to 88% and 82% to 88%, respectively. A bivariate meta-analysis produced a summary value for sensitivity of 0.81 (95% CI 0.73 to 0.88) and for specificity of 0.85 (95% CI 0.79 to 0.90) (three studies).

The negative predictive value (NPV; that is, the probability that patients who are diagnosed by virtual chromoendoscopy as having a hyperplastic polyp truly do not have an adenoma) was more variable across the NBI studies than the FICE or i-scan studies. i-scan had the most consistently favourable results on this outcome, but this may have been due to a higher proportion of the i-scan studies involving endoscopists with prior experience of i-scan.

The percentage agreement between surveillance intervals allocated following NBI (13 studies) and those allocated following histopathology ranged from 84% to 99%. The agreement following i-scan (two studies) ranged from 93% to 97% and for FICE (two studies) from 97% to 100%. When only considering studies in which surveillance intervals were assigned in accordance with the two Preservation and Incorporation of Valuable endoscopic Innovation programme (PIVI) criteria (guidance on the requirements that new technologies should meet before a 'resect and discard' strategy can be applied in practice), eight of the nine NBI studies reporting this outcome achieved a level of agreement that was ≥ 90%, thus meeting the first PIVI criterion. Both the i-scan studies reporting this outcome achieved an agreement ≥ 90%. All NBI (five) and i-scan (one) studies that assessed NPV for high confidence assessments of diminutive polyps in the rectosigmoid met the second PIVI criterion of achieving an NPV ≥ 90%. There was no evidence for FICE in relation to the PIVI criteria.

None of the identified studies measured health-related quality of life (HRQoL), anxiety, number of outpatient appointments or telephone consultations, incidence of colorectal cancer or mortality. Four studies assessed adverse effects, stating there were none. Data were too limited on the number of polyps that would be left in place, resected, discarded or sent histopathology, and the time to perform the colonoscopy, for the review to draw conclusions about these outcomes.

Cost-effectiveness

We included two studies of virtual chromoendoscopy compared to histopathology in our systematic review of economic evaluations. Both compared a resect and discard strategy with current practice of submitting all polyps to histopathology. The evaluations were published in the USA and found that there were cost savings for the resect and discard group ranging between US$25 and US$174 per person.

In addition, a study by Olympus, the manufacturer of NBI systems, describes a budget impact analysis of NBI for the NHS in England. The decision tree model has a time horizon of seven years and in each year there is a cohort of patients that undergo endoscopy. The study estimated that NBI offers cost savings of £141 million over seven years.

Results of our independent economic model suggest that virtual chromoendoscopy is cost saving compared to histopathology with a mean saving of between £73 and £87 per person over their lifetime for the different VCE technologies. There is a small increase in QALYs with NBI and i-scan compared to histopathology of between 0.0005 – 0.0007 QALYs per person, while FICE is associated with 0.0001 QALYs fewer per person than histopathology. NBI and i-scan dominate histopathology, i.e. they are less

expensive and more effective. FICE is cost effective compared to histopathology, with a cost saved per QALY lost of £671,383. The model estimates that the correct surveillance interval would be given to 95% of patients with NBI, 94% of patients with FICE and 97% of patients with i-scan. Results are most sensitive to the pathology cost, the probability of perforation with polypectomy and the proportion of patients who die from perforation. Probabilistic sensitivity analyses were conducted for pairwise and incremental comparisons for histopathology with virtual chromoendoscopy technologies. The probabilistic ICERs were similar to the base case deterministic ICERs. At a willingness-to-pay threshold of £20,000 and £30,000, i-scan was most cost effective in 95% and 33% of simulations respectively.

Discussion

Evidence was limited for FICE and i-scan, and was generally limited for high confidence characterisations in the rectosigmoid colon. The heterogeneity among the NBI studies in setting, country, endoscopists' experience and training makes it difficult to determine the diagnostic accuracy of NBI. Uncertainties include the generalisability of the evidence base to the UK, how the settings of studies' may have impacted on the results (e.g. academic centres compared to community hospitals), and a lack of data on longer-term health outcomes among patients undergoing virtual chromoendoscopy for assessment of diminutive polyps. Studies providing evidence on the diagnostic accuracy of characterising polyps did not relate this to the prediction of surveillance intervals of patients, in order to predict disease progression in patients. The economic analysis includes only diminutive polyps and does not differentiate between the type of polyp such as depressed polyps or sessile serrated polyps. There were limitations in the data available for the prevalence of adenomas across risk classification, the distribution of polyps and the proportion of patients in the higher risk categories with small and large adenomas, which necessitated assumptions in the economics model. There are also limitations in the data on recurrence rates post-polypectomy. The full uncertainty around the model results have not been explored in the PSA as the long-term outcome parameters have not been varied.

Conclusions

Implications for service provision

Virtual chromoendoscopy technologies, using high definition systems without magnification, have the potential for use in practice for the real-time assessment of diminutive colorectal polyps, if endoscopists have adequate experience and training. NBI and i-scan generally meet the PIVI requirements to be used to perform a 'resect and discard' strategy, but it is unclear how the findings generalise to UK practice. Virtual chromoendoscopy was estimated to be cost saving compared to histopathology. It was associated

with a small gain in QALYs for NBI and i-scan and a small decrease in QALYs for FICE. The least costly and most effective of the technologies in terms of diagnostic accuracy was i-scan, which might be explained by the the sparseness of data on diagnostic accuracy for i-scan, and the fact that most of the studies involved experienced endoscopists working in specialist centres.

Suggested research priorities

Future research priorities include: head-to-head RCTs of all three virtual chromoendoscopy technologies; more research on the diagnostic accuracy of FICE and i-scan (when used without magnification); further studies evaluating the impact of endoscopist experience and training on outcomes; studies measuring adverse effects, HRQoL and anxiety; and, longitudinal data on colorectal cancer incidence, HRQoL and mortality.

Word count: 2,624

Plain English Summary

Colorectal polyps are growths in the large bowel. Some polyp types, called adenomas, can develop into bowel cancer if not diagnosed and removed. Specialised doctors or nurses, called 'endoscopists' can find polyps when they look at the inner lining of the large bowel (colonoscopy). If a polyp is found, it is removed and sent to a laboratory to see if it is an adenoma (this is called 'histopathology'). A new technique, called virtual chromoendoscopy can be used during a colonoscopy to help endoscopists decide if a very small polyp (5 mm or smaller) is an adenoma or not, instead of sending the polyp to a laboratory. If the endoscopist is confident the very small polyp is not an adenoma it could be left in the bowel, rather than removed. We aimed to assess the benefits and harms of three virtual chromoendoscopy technologies for diagnosing very small polyps compared to histopathology, and whether these are an effective use of NHS financial resources. We found and reviewed all the studies that had assessed these techniques, using standard methods, and created an economic model. We found virtual chromoendoscopy correctly identified polyps as adenomas most of the time, although results did vary between studies. Endoscopists experienced in virtual chromoendoscopy achieved better results than those without experience. Virtual chromoendoscopy techniques were estimated to be cost saving compared to histopathology. The model estimated that NBI and i-scan had slightly better long-term outcomes than histopathology, whilst FICE had slightly worse outcomes. (244 words)

LIST OF ABBREVIATIONS

ACPGBI	Association of Coloproctology of Great Britain and Ireland
ASGE	American Society of Gastrointestinal Endoscopy
BSG	British Society of Gastroenterology
CD	Correct diagnosis
CI	Confidence interval
CRC	Colorectal cancer
CRD	Centre for Reviews and Dissemination
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
DARE	Database of Abstracts of Reviews of Effectiveness
DISCARD	Detect, InSpect, ChAracterise, Resect and Discard
EAG	External Assessment Group
EED	Economic Evaluation Database
ESGE	European Society of Gastrointestinal Endoscopy
FAP	Familial adenomatous polyposis
FICE	Flexible Spectral Imaging Colour Enhancement
FOBT	Faecal occult blood test
FN	False negative
FP	False positive
GP	General Practitioner
HCHS	The Hospital and Community Health Services (HCHS) index
HD	High definition
HNPCC	Hereditary non-polyposis colorectal cancer
HPRC	Hyperplastic polyp(s) resected correct surveillance
HPRI	Hyperplastic polyp(s) resected incorrect surveillance
HR	High risk
HRQoL	Health-related quality of life
IBD	Inflammatory bowel disease
ICER	Incremental cost-effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
IR	Intermediate risk
JAG	Joint Advisory Group

LR	Low risk
MAC	Missed adenoma(s) correct surveillance
MAHPR	Missed adenoma(s) and hyperplastic polyp(s) resected
MAI	Missed adenoma(s) incorrect surveillance
NBI	Narrow band imaging
NHS	National Health Service
NHMRC	National Health and Medical Research Council
NICE	NBI International Colorectal Endoscopic
NIHR	National Institute for Health Research
NAC	Novel Classification System
NPV	Negative predictive value
PEDro	Physiotherapy Evidence Database
PIVI	Preservation and Incorporation of Valuable endoscopic Innovation programme
PPV	Positive predictive value
PSSRU	Personal Social Services Research Unit
QUADAS	Quality Assessment Tool for Diagnostic Accuracy Studies
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RevMan	Review Manager
ScHARR	School of Health and Related Research, The University of Sheffield
SD	Standard deviation
SBCS	ScHARR Bowel Cancer Screening
SROC	Summary Receiver Operating Curve
TN	True negative
TP	True positive
UEG	United European Gastroenterology
UK	United Kingdom
UKCTG	UK Clinical Trials Gateway
US	United States
USA	United States of America
USMSTF	US Multi-Society Task Force on Colorectal Cancer

VC	Virtual chromoendoscopy
WASP	Workgroup serrAted polypS and Polyposis classification
WHO	World Health Organisation
WLE	White light endoscopy

1 BACKGROUND

1.1 Description of the health problem

Colorectal polyps are small growths (usually less than 1cm in size) on the inner lining of the colon or rectum. They are common, affecting 15-20% of the general population and they usually occur in people who are over 60 years of age.[1] Colorectal polyps do not usually cause symptoms though some larger polyps are associated with rectal bleeding, diarrhoea, constipation, and abdominal pain.

Colorectal polyps can be described in a variety of ways, e.g. by size, according to the type of cell or tissue they arise from within the colon or rectum, according to their shape, and according to their histology.[2] Histological classification generally distinguishes between polyps that are adenomatous (known as adenomas, or less commonly, neoplastic polyps), hyperplastic, or deep submucosal invasive cancers. Adenomas may eventually become cancerous if undiagnosed and untreated. Hyperplastic polyps usually do not carry a risk of developing into cancer; however, a subgroup of hyperplastic polyps, called sessile serrated polyps (polyps that have a slightly flattened shape with a saw tooth appearance), also have the potential to develop into cancer.

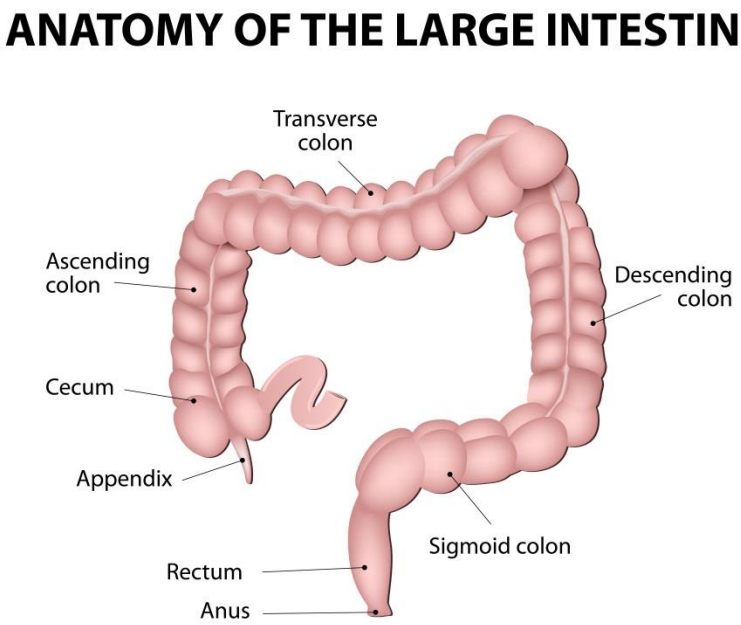
In terms of size, polyps measuring ≥10mm are referred to as large, whilst those 9mm to 6mm are considered small, and those 5mm or less are classified as diminutive. It has been estimated that 80% of polyps detected at colonoscopy are diminutive.[3] A person can have more than one colorectal polyp, and can have polyps of different sizes (e.g. diminutive polyps in addition to small polyps and large polyps). The morphology of a polyp can be described using the Paris endoscopic classification[4] (Table 1). For the prediction of malignancy the Association of Coloproctology of Great Britain and Ireland (ACPGBI)[5] recommends the use of the Paris endoscopic classification in conjunction with an estimation of the size of a polyp.

Colorectal polyps are usually detected during colonoscopy, a procedure involving examination of the rectum and the colon via a flexible tube called a colonoscope (a type of endoscope). The colonoscope is advanced inside the colon to the cecum (Figure 1) and then slowly withdrawn by the endoscopist who views images of the inner lining on a monitor. Patients might be referred for colonoscopy following an abnormal bowel screening result (see below), or following referral from primary care due to symptoms suggestive of colorectal cancer or of inflammatory bowel disease (IBD), or as part of routine colonic

surveillance [e.g. follow-up after previous polyp removal (a polypectomy), or for IBD] (see Section 1.3 for details of the care pathway).

Table 1 The Paris endoscopic classification[4]

	Type	**Features**
Protruded	Type 0-1p	Pedunculated (on a stalk)
	Type 0-1sp	Sub pedunculated
	Type 0-1s	Sessile
Superficial Elevated	Type 0-2a	Flat elevated
	Type 0-2a+2c	
	Type 0-2a+Depression	
Flat	Type 0-2b	Flat
Depressed	Type 0-2c	Slightly depressed
	Type 0-2c+2a	
Excavated (ulcer)	Type 0-3	



Designua/Shutterstock.com

Figure 1 Illustration of the large intestine

Colorectal cancer is one of the most common cancers in the UK after breast and lung cancer with approximately 41,900 new cases registered each year.[6] The prevalence of colorectal cancer increases with age, with 99% of cases occurring in people aged more than 40 years and 85% in those aged more than 60.[7] A family history of bowel cancer is a key risk factor, with the risk increasing according to greater number of first degree relatives affected. [7] Familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC) (also known as Lynch syndrome) are inherited genetic disorders that increase the risk of colorectal cancer, though are rare, accounting for only 5% of cancer cases.[7] Other factors thought to increase risk of colorectal cancer include diet (e.g. increased consumption of red and processed meat; lack of dietary fibre; lack of fruit and vegetables); obesity and lack of physical activity; consumption of alcohol and use of tobacco; and presence of longstanding IBD (e.g. Crohn's disease or ulcerative colitis).

The NHS Bowel Cancer Screening Programme offers screening every two years to men and women aged 60 to 74 years. The programme invites eligible adults to carry out a faecal occult blood test (FOBT), which detects small amounts of blood in faeces. People with an abnormal FOBT result are referred for a colonoscopy to determine risk of colorectal cancer.

Upon diagnosis of colorectal cancer patients will undergo staging and grading, with use of biopsy and imaging (e.g. computed tomography, endorectal ultrasonography or magnetic resonance imaging). The Dukes' classification is a four stage system (A-D) commonly used to determine the size and spread of the cancer. At Dukes' A the cancer is only in the innermost lining of the bowel or slightly growing into the muscle layer, whilst at Dukes' D the cancer has spread to other parts of the body such as the liver or the lungs. Treatment of the cancer will depend upon the stage, but commonly includes surgical resection, combined with chemotherapy and radiotherapy where necessary, and in some cases biological therapies.[8] Bowel cancer survival rates in England vary according to stage, with rates for Stage 1 patients (known as Dukes' A colorectal cancer) in the range 95% to 100% at five years or more after diagnosis.[6] At Stage 4 (Dukes' D) survival rates at five years or more are just 5% to 10% (though this could be as high as 40% if liver metastases can be successfully removed by surgery).[6] Generally for people with colorectal cancer in England and Wales almost 60% survive their cancer for 10 years or more following diagnosis (based on all stages).[6]

1.2 Description of the diagnostic technologies under assessment

Current clinical practice is to detect colorectal polyps using conventional white light endoscopy. This may be used in combination with dyes (chromoendoscopy) to enhance visualisation of tissues in the area being

inspected. Detected polyps are then removed and each is sent for laboratory histopathological examination to determine whether it is an adenoma (therefore at a high cancer risk) or hyperplastic (at a low cancer risk).[1] (NB. In some centres some polyps may be left in situ if endoscopists are confident on the basis of white light endoscopy that they are hyperplastic). The aim is to communicate the results to patients within a two week period. Histopathological examination is regarded as the reference standard method for characterising polyps, though it can be associated with errors of measurement and interpretation. For example, concerns have been raised about poor inter-rater reliability between gastrointestinal histopathologists.[9] Further, some diminutive polyps may be damaged during resection (or cannot be resected at all), impairing the effectiveness of histopathological analysis.[3]

Virtual chromoendoscopy refers to electronic endoscopic imaging technologies that provide detailed contrast enhancement of the mucosal surface and blood vessels in the colon and rectum. A number of virtual chromoendoscopy technologies are available. All of these technologies use an endoscopy system typically consisting of an endoscope, a light source, a video processor and a visual display monitor.[10,11] The light source produces light that is transmitted to the distal end of the endoscope to illuminate the area under inspection. The video processor captures and processes electrical signals to enable an image of the inspected area to be displayed on the monitor.[11]

The aim of virtual chromoendoscopy technologies is to provide enhanced visualisation of tissues without the need for dyes, enabling the endoscopist to differentiate between adenomatous and hyperplastic colorectal polyps in real-time during colonoscopy. Virtual chromoendoscopy technologies can be classed as optical or digital. In optical virtual chromoendoscopy, optical lenses are integrated into the endoscope's light source, which selectively filters white light, resulting in narrow band light. In digital chromoendoscopy, digital post-processing by the video processor is used to enhance the real-time image.[12]

As discussed in Section 2, there are three commercial systems of relevance to this diagnostic assessment report:

- Narrow band imaging (NBI), a type of optical chromoendoscopy
- Flexible Spectral Imaging Colour Enhancement (FICE), a type of digital chromoendoscopy
- i-scan, a type of digital chromoendoscopy

Each of these will be described in turn.

1.2.1 Narrow band imaging (NBI)

Narrow band imaging (NBI) (Olympus Medical Systems) is an optical image enhancement technology used in the Olympus endoscopic video imaging systems EVIS LUCERA ELITE,[13] EVIS EXERA III[14] (not available in the UK) and EVIS LUCERA SPECTRUM.[15] NBI is achieved by using a filter in the light source unit and a function on the video processor. The white light is filtered resulting in narrow-band light which consists of two wavelengths 415 nm blue light and 540 nm green light.[12,15] These wavelengths are strongly absorbed by haemoglobin and thus NBI enhances the contrast between blood vessels and the surrounding mucosa in comparison to illumination by standard white light. The endoscopist can switch viewing mode from standard white light to NBI and vice versa at any time. The image quality achieved varies between the different endoscopy systems due to differences in image sensors and video processors with the newer EVIS LUCERA ELITE system offering the highest quality images. Furthermore, within a class of endoscopy system there will also be differences in image quality depending on the precise model of endoscope used. For example, within the EVIS LUCERA ELITE group the EVIS LUCERA ELITE 290HQ (high definition) endoscope offers the highest image quality, followed by the EVIS LUCERA ELITE 290H endoscope. The EVIS EXERA system is considered to be comparable with the EVIS LUCERA system in terms of diagnostic performance. The Olympus endoscopy system (including processor, endoscope and annual maintenance) is estimated to cost £87,385.

1.2.2 Flexible Spectral Imaging Colour Enhancement (FICE)

FICE (Aquilant Endoscopy/FujiFilm) is a digital image processing function used in the Fuji video endoscopy systems EPX-4450HD, EPX-3500HD and EPX-4400.[16] White light illuminates the area of interest and the conventional images captured from the reflected light can be processed in real-time by software into spectral images (images based on specific light wavelengths). FICE has ten pre-set wavelength settings which can also be manually altered to achieve the best enhancement of the image.[12,16] The endoscopist can switch between viewing conventional or FICE images at any time. The image quality achieved varies between the different systems, being higher on the EPX-4450HD and EPX-3500HD systems than on the EPX-4400 system. As well as being a feature of three Fuji endoscopy systems the 500 series and 600 series endoscopes can also use FICE and it can be used in combination with magnifying endoscopes. The Aquilant Endoscopy/FujiFilm endoscopy system (including processor, endoscope and annual maintenance) is estimated to cost £59,312.

1.2.3 i-scan

i-scan (Pentax Medical) is a digital image processing technology used with Pentax endoscopy systems.[17] White light illuminates the area of interest and there are three different algorithms for real-time image processing:[12,18]

- Surface enhancement - helps to visualise the edges of anatomical structures by improving light-dark contrast.
- Contrast enhancement - helps to visualise depressed areas by digitally adding blue colour to relatively dark areas.
- Tone enhancement - modifies the colour contrast of the normal image to create an improved image with enhanced visibility of minute mucosal structures and subtle changes in colour.

The three different algorithms are then used in different combinations for three i-scan modes: (i) i-scan 1 for detection of lesions; (ii) i-scan 2 for characterisation of lesions; and (iii) i-scan 3 for demarcation of lesions. The endoscopist can switch between the conventional image and the three i-scan modes at any time. If using equipment enabled with the capability (the EPK-i7000) it is possible to display a normal white light image and an i-scan image simultaneously side by side.[18] The Pentax endoscopy system (including processor, endoscope and annual maintenance) is estimated to cost £83,616.

1.2.4 Definition and magnification

The manufacturers of the technologies recommend that high definition endoscopy systems are used to optimise the quality of the image. A high definition system would be one in which the endoscope, the video processor, the display monitor and the cabling are, collectively, capable of producing an image corresponding to 650 to 720 lines of resolution.[19] The majority of monitors currently in use would be high definition capable, though not all endoscopes would be high definition. When equipment is due for replacement they will be upgraded to high definition status.

Magnifying endoscopes (also sometimes referred to as near focus or zoom endoscopes) can be used to enhance the clarity of images by magnifying up to 150 times. A movable lens can be fitted to the tip of the endoscope to provide optical zoom. However, magnifying endoscopes are largely unavailable in routine settings as they are not considered practical for day to day use. Most standard endoscopes can provide magnification of up to 35 times at the push of a button.

1.2.5 Classification schemes

Endoscopists make a general assessment of polyps based on observation of elements such as colour, blood vessels and surface pattern. There are several different classification schemes available, with particular schemes used with specific technologies. For example, the NBI International Colorectal Endoscopic scheme was devised specifically for use with NBI.[20] The Novel Classification System (NAC) has been developed for use with FICE.[21] Examples of classification schemes are shown in Table 25.

Table 2 Examples of virtual chromoendoscopy classification schemes for colorectal polyps

Name of Scheme	**Basis for classification**	**Classification categories**		
NBI International Colorectal Endoscopic (NICE) classification[20]	Polyp histology (based on colour, vessels and surface pattern when viewed by NBI)	Type 1		Hyperplastic
		Type 2		Adenoma
		Type 3		Deep submucosal invasive cancer
Kudo classification[22]	Pit pattern (fine surface structure of the of the mucosa when viewed by magnifying chromoendoscopy)	Round pits	Type 1	Benign changes (e.g. normal, hyperplastic, inflammatory polyps)
		Stellar or papillary pits	Type II	
		Large tubular or roundish pits	Type III L	Neoplastic and malignant changes
		Small tubular or roundish pits	Type III s	
		Branch-like or gyrus-like pits	Type IV	
		Non-structural pits	Type V	

Showa classification[23]	Vascular pattern (pattern of microvessels surrounding the pit when viewed by NBI)	Normal	Characteristic of non-neoplasia
		Faint	
		Network	Seen in neoplasia
		Dense	
		Irregular	Seen in neoplasia, useful for a diagnosis of cancer
		Sparse	

A classification system for endoscopic differentiation of small and diminutive adenomas, hyperplastic polyps and sessile serrated adenomas and polyps has recently been developed (the Workgroup serrAted polypS and Polyposis (WASP) classification).[24]

1.2.6 Training in the use of virtual chromoendoscopy

Training in the use of virtual chromoendoscopy is necessary to ensure adequate endoscopist performance in characterising polyps. Training methods vary, and can involve endoscopists making ex vivo predictions based on still images previously taken using virtual chromoendoscopy, as well as in vivo predictions in real time during colonoscopy under supervision of an endoscopist more experienced in use of the technology. The duration of training may vary, with endoscopists subject to post-training key performance indicators and auditing. For example, the manufacturers of NBI estimate that a one to two day initial course would be sufficient. An online computer training App can be used as refresher training, in conjunction with audits and use of a validated classification scheme. Results of a recent study in England showed that a learning curve is observed in practice even for endoscopists experienced in in-vivo colorectal polyp characterisation.[25] A 90% threshold for diagnostic accuracy was achieved with use of high definition white light endoscopy followed by i-scan once 200 polyps (<10mm in size) had been examined. This suggests that, following initial training, endoscopists should receive regular feedback on the accuracy of their diagnostic predictions (e.g. via histopathology on small batches of polyps) until an acceptable level of accuracy has been reached. This may take up to six months depending on the volume of colonoscopies performed. Criteria for diagnostic performance of virtual chromoendoscopy have been proposed by international guidelines (see Section 1.3), which specify the need for endoscopists to be adequately trained and audited. The Joint Advisory Group (JAG) on gastrointestinal endoscopy has issued key performance indicators and quality assurance standards for colonoscopy[26] and offers accreditation for colonoscopists, though there is no accreditation specifically for virtual chromoendoscopy.

1.3 Care pathway

Figure 2 provides an illustration of the care pathway showing indications for colonoscopy and subsequent management, reproduced from the National Institute for Health and Care Excellence scope for this diagnostic assessment.[27] As mentioned in Section 1.1, patients may be referred for colonoscopy via a number of routes. For example, they may receive colonoscopy following an abnormal bowel cancer screening result, or after referral from primary care due to symptoms suggestive of colorectal cancer (e.g. rectal bleeding, pain, or altered bowel habits).

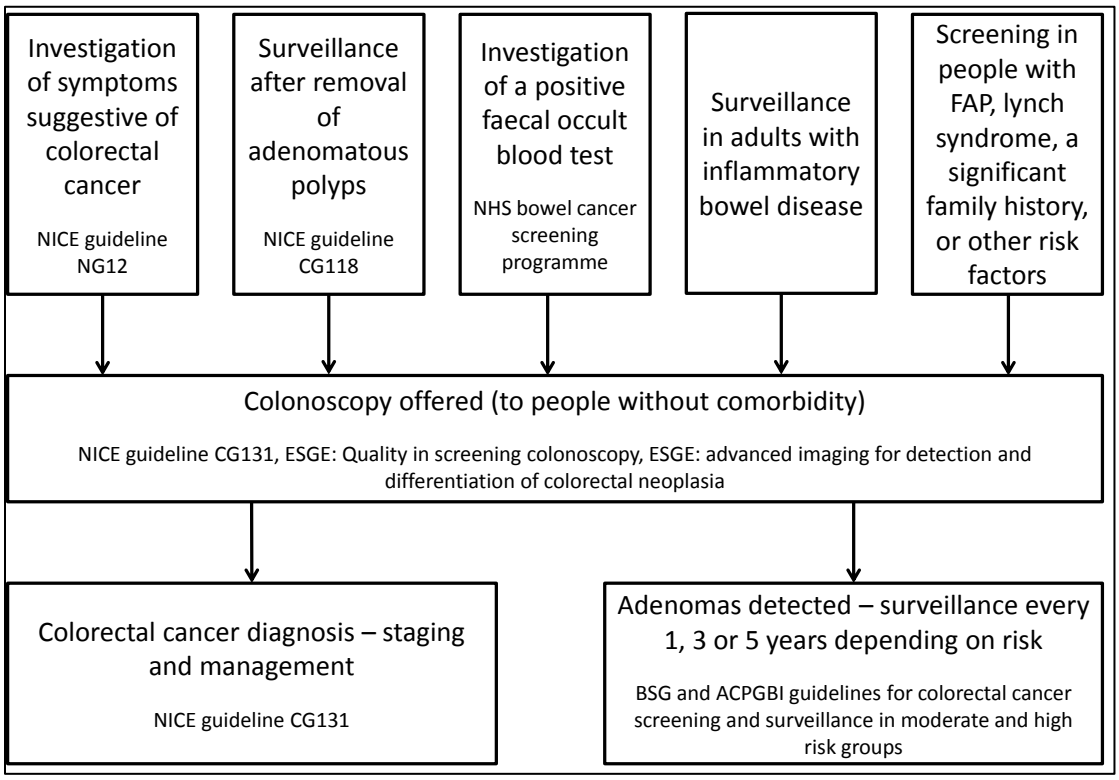


Figure reproduced with permission from the National Institute for Health and Care Excellence Scope for this appraisal[27]

Figure 2 Care pathway before and after colonoscopy

The risk of colorectal cancer varies between different patient groups. Patients with FAP and HNPCC (Lynch syndrome) have a high risk of colorectal cancer. Patients with an abnormal bowel cancer FOBT result may be at higher risk than patients undergoing surveillance for removal of adenomatous polyps.

Following the detection of colorectal adenomas by colonoscopy, a surveillance interval will be set, based on the size and number of adenomas found. The British Society of Gastroenterology (BSG) and the

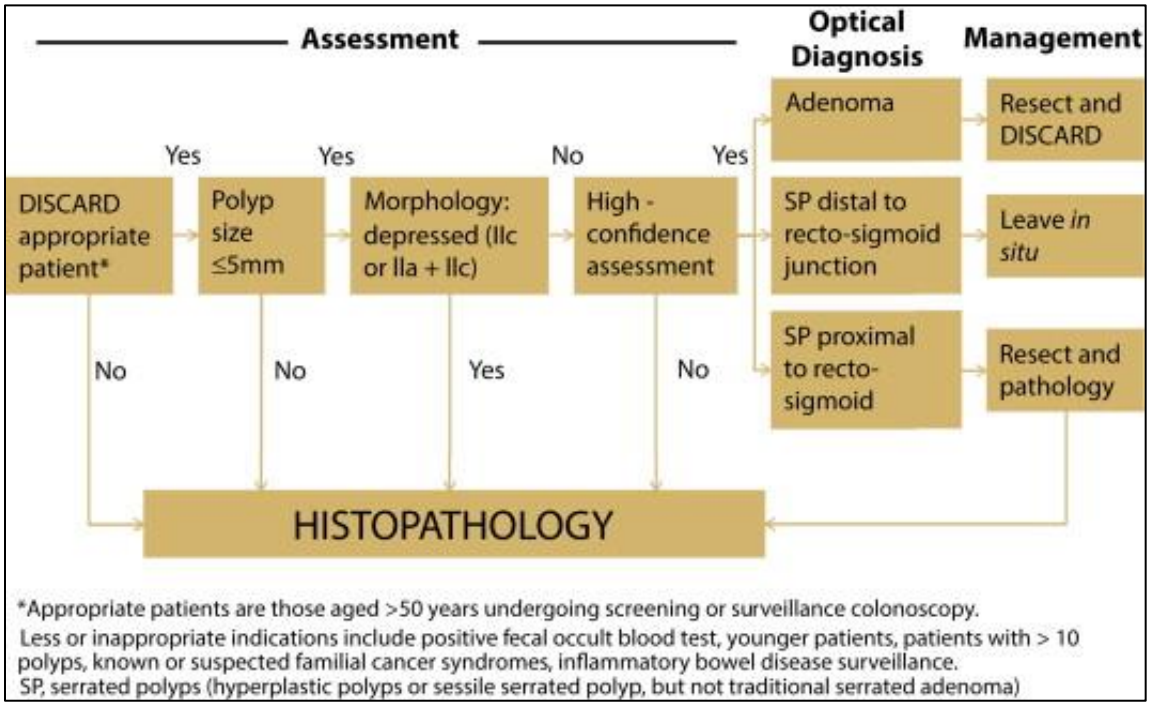
Association of Coloproctology for Great Britain and Ireland have issued guidelines for colorectal cancer screening and surveillance in moderate and high risk groups.[28] The following recommendations are made:

- people with 1 or 2 small (less than 1 cm) adenomas are at low risk, and need no colonoscopic surveillance or 5-yearly surveillance until one negative examination then cease surveillance.
- people with 3 or 4 small adenomas or at least 1 adenoma this is 1 cm or larger are at intermediate risk and should be screened 3-yearly until two consecutive examinations are negative.
- people with 5 or more adenomas, or 3 or more adenomas at least one of which is 1 cm or bigger, are at high risk and an extra examination should be undertaken at 12 months before returning to 3-yearly surveillance.

The National Institute for Health and Care Excellence clinical guideline 118 on colonoscopic surveillance in people with IBD or adenomas makes similar recommendations.[29]

Virtual chromoendoscopy takes place in secondary or tertiary care at the same point in the care pathway as current clinical practice using conventional white light endoscopy or dye-based chromoendoscopy. It is likely that virtual chromoendoscopy technologies would be used alongside conventional white light endoscopy, since all the technologies relevant to this assessment allow the endoscopist to change viewing mode from standard white light to the virtual chromoendoscopy image in real-time at the flick of a switch. For example, the endoscopist may begin examining the colon with white light endoscopy, and then (in some cases) use dye to enhance visualisation of potential adenomas. They may then switch the endoscope to use virtual chromoendoscopy to further enhance visualisation. This practice is referred to as optical assessment of colorectal polyps. The care pathways would diverge when a diminutive polyp of ≤ 5mm is detected. Under current clinical practice a diminutive polyp identified by conventional white light endoscopy would be removed and sent for histopathological examination to determine whether it is adenomatous, hyperplastic, or cancerous.[30] However, use of a virtual chromoendoscopy technology would enable the endoscopist to differentiate between adenomas and hyperplastic polyps during colonoscopy. Where the endoscopist has high confidence in the polyp characterisation, adenomas would be removed and discarded whereas hyperplastic polyps in the rectosigmoid colon would be left in situ (as these would be considered very low risk for colorectal cancer). This is referred to as the DISCARD strategy (Detect, InSpect, ChAracterise, Resect and Discard)[3] (Figure 3). Where there is low confidence in determining whether a polyp is adenomatous or hyperplastic it should be resected and sent for histopathological examination. Any flat depressed polyps, polyps with a distorted shape, and hyperplastic appearing (serrated-appearing) polyps in the proximal colon should be sent for histopathology examination, irrespective of size. The level of confidence with which polyp classifications are made is subjective and varies between endoscopists. Some endoscopists increase objectivity by referring to the relevant

classification system, e.g. a high confidence assessment made with NBI might be based on whether at least two of the NICE classification criteria apply to the particular polyp (i.e. based on polyp colour, vessels and surface pattern).



Reprinted from Gastrointestinal Endoscopy, 82/2, Wang L.M. and East J.E. Diminutive polyp cancers and the DISCARD strategy: Much ado about nothing or the end of the affair? Pages 385-8. Copyright (2015), with permission from Elsevier

Figure 3 Flow chart for low-risk application of the DISCARD strategy for diminutive colorectal polyps (from Wang and East, 2015)[3]

Advantages of the DISCARD strategy include the fact that real-time characterisation of polyps may potentially alleviate patient anxiety associated with waiting for histopathology results and reduce health service and patient costs associated additional appointments. A surveillance interval can be set on the day of the procedure, rather than at a follow-up appointment following the results of histopathology and savings may be made through reduced use of histopathology. It has been reported that histopathology accounts for up to 10% of the cost of colonoscopy,[3] and that use of colonoscopy in the NHS is increasing each year.

There may be potential disadvantages associated with the use of virtual chromoendoscopy. For example, endoscopists will need to have sufficient experience with in-vivo characterisation of polyps and adequate training in, and experience of, the particular virtual chromoendoscopy technology. This is a requirement

of European and American endoscopy guidance (see Section 1.3.1). It has been noted that performance among community-based endoscopists may not necessarily meet these requirements.[3] Furthermore, there is the risk that a diminutive polyp cancer (incidence rates of which vary from 0% to 0.6%[3]) may inadvertently be characterised as an adenoma, resected and discarded without histopathological examination, with malignant cells left behind, and subsequent potential development of undiagnosed metastatic disease and death.[3] To attempt to address these concerns, international professional associations have issued guidance on the use of virtual chromoendoscopy as part of a DISCARD strategy, discussed next.

1.3.1 Diagnostic thresholds and requirements for use of virtual chromoendoscopy

There are several different aspects to any decision to implement the new technology and European[30] and American guidance[31] has been published.

The European guidance,[30] produced by the European Society of Gastrointestinal Endoscopy (ESGE) in 2014 makes the recommendation that virtual chromoendoscopy (NBI, FICE, i-scan) and conventional chromoendoscopy can be used, under strictly controlled conditions, for real-time optical diagnosis of diminutive (≤5 mm) colorectal polyps to replace histopathological diagnosis. The optical diagnosis has to be reported using validated scales, must be adequately photo-documented, and can be performed only by experienced endoscopists who are adequately trained and audited (ESGE describe this as a weak recommendation based on high quality evidence).

The American guidance[31] on real-time endoscopic assessment of the histology of diminutive colorectal polyps is part of the Preservation and Incorporation of Valuable endoscopic Innovations (PIVI) initiative of the American Society for Gastrointestinal Endoscopy (ASGE). The PIVI statement defines two requirements, which new technologies for the real-time endoscopic assessment of the histology of diminutive colorectal polyps should meet, before a 'resect and discard' strategy could be applied:

1. In order for colorectal polyps ≤5 mm in size to be resected and discarded without pathologic assessment, endoscopic technology (when used with high confidence) used to determine histology of polyps ≤5 mm in size, when combined with the histopathologic assessment of polyps >5 mm in size, should provide a ≥90% agreement in assignment of post-polypectomy surveillance intervals when compared to decisions based on pathology assessment of all identified polyps.

2. In order for a technology to be used to guide the decision to leave suspected rectosigmoid hyperplastic polyps ≤5 mm in size in place (without resection), the technology should provide ≥90% negative predictive value (when used with high confidence) for adenomatous histology.

If it is judged that the polyp cannot be confidently assessed using an endoscopic technology then it should be resected and sent for histopathological diagnosis. The guidance also indicates that polyp images should be permanently stored and should be of sufficient resolution to support the endoscopists' assessment and clinical decisions.

1.4 Current service provision

As stated above, current practice is to detect polyps using white light endoscopy, with additional dye based chromoendoscopy used where necessary to provide additional information on polyp characteristics. All diminutive polyps detected are resected and undergo histopathological analysis to determine whether they are adenomatous or hyperplastic. A surveillance interval is then set based on the number and size of adenomas detected. The majority of existing endoscopy systems in use in NHS hospitals are thought to be capable of virtual chromoendoscopy. The technology is built into the light source and video processor and can be activated by the endoscopist by a switch at any time during colonoscopy. The lifecycle of an endoscopy system is estimated to be between five and eight years, and all new systems are now equipped with virtual chromoendoscopy technology. However, virtual chromoendoscopy and the DISCARD strategy is not thought to be routinely used as a management protocol. However in some centres diminutive polyps in the rectosigmoid colon are optically diagnosed using white light or virtual chromoendoscopy and left in place if there is high confidence the polyps are hyperplastic. Of the three technologies of relevance to this assessment, NBI is considered to be the most widely available, and it has the largest market share for electro-medical service contracts in England.

2 DEFINITION OF THE DECISION PROBLEM

Under current clinical practice all diminutive polyps (1-5 mm in size) identified by conventional white light endoscopy would be removed and sent for histopathological examination to determine whether they are adenomas or hyperplastic, and the consequent colorectal cancer risk. Once histopathology results are available a surveillance interval is set according to the number and size of adenomas detected. Use of a virtual chromoendoscopy technology would provide the endoscopist with enhanced visualisation to differentiate between adenomas which could be resected and discarded (i.e. not sent for histopathological assessment) and hyperplastic polyps in the rectosigmoid colon which could be left in situ. This can only be done when the endoscopist is highly confident in their characterisation of the polyp.

The potential benefits of virtual chromoendoscopy would be fewer resections (polypectomy) of low risk hyperplastic polyps (with a resulting reduction in complications such as bleeding or perforation of the bowel); the provision of results more quickly, thus potentially reducing patient anxiety; a reduction in health resource use through fewer histopathological examinations; and management (including surveillance) decisions could also be provided more quickly. Guidelines recommend virtual chromoendoscopy should be performed only under strictly controlled conditions by experienced endoscopists adequately trained in the use of the technology, using validated classification scales.[30]

In order for virtual chromoendoscopy technologies to be incorporated into routine clinical practice for the real-time assessment of colorectal polyps during colonoscopy, there needs to be evidence that the new technology provides an appropriate and efficient standard of care compared to existing practice. Therefore, the decision question for this assessment is does virtual chromoendoscopy for real-time assessment of diminutive colorectal polyps during colonoscopy, represent a cost-effective use of NHS resources?

2.1.1 Populations and relevant subgroups

The population of relevance to this assessment is: people referred for colonoscopy through the NHS Bowel Cancer Screening Programme because of an abnormal FOBT test result; people offered colonoscopic surveillance because they had adenomas previously removed; and people undergoing colonoscopy with diminutive colorectal polyps referred for colonoscopy by a GP because of symptoms suggestive of colorectal cancer.

At the scoping stage of this assessment it was agreed that patients with IBD, or conditions such as FAP or HNPCC would not be relevant, as these are distinct patient groups with increased risks of colorectal

cancer in whom differentiation between adenomatous and non adenomatous polyps during colonoscopy is more complicated (e.g. in patients with IBD because of factors such as increased amount of microvessels). Virtual chromoendoscopy with a DISCARD strategy would be unlikely to be used in these patients.[8] At the scoping stage it was also considered that small polyps (6-9mm in size) would not be included in the scope of the assessment.[8]

2.1.2 Index tests

Virtual chromoendoscopy is the index test, of which three technologies are considered relevant to this diagnostic assessment. These are:

- NBI
- FICE
- i-scan

Each technology should be used with high definition or high resolution monitors and endoscopes without the use of magnification.

2.1.3 Reference standard

The reference standard for virtual chromoendoscopy is histopathological assessment of diminutive polyps.

2.1.4 Outcomes

A range of outcomes are relevant to this assessment, which can be classified as diagnostic test accuracy [e.g. accuracy (i.e. proportion of correctly classifiedpolyps among all the polyps), sensitivity, specificity, accuracy, negative and positive predictive values], intermediate outcomes (e.g. recommended surveillance intervals, time taken to perform colonoscopy), patient reported outcome measures (e.g. health-related quality of life), clinical outcomes (e.g. adverse effects of polypectomy, colorectal cancer) and cost outcomes (e.g. endoscopy system costs, colonoscopy and related costs, training costs, histopathology costs).

2.2 Overall aims and objectives of assessment

The aim of this research is to assess the clinical-effectiveness and cost-effectiveness of technologies that could aid the characterisation of diminutive colorectal polyps that have the potential to become cancerous.

Specific objectives are to determine, through a systematic review and economic evaluation, the clinical-effectiveness and cost-effectiveness of the virtual chromoendoscopy technologies NBI, FICE, and i-scan in the characterisation and management of diminutive colorectal polyps.

3 METHODS

We set out the methods for the systematic reviews of clinical and cost-effectiveness a priori in a research protocol, which was published on the National Institute for Health and Care Excellence's website (https://www.nice.org.uk/guidance/GID-DG10004/documents/final-protocol). The protocol was also registered with PROSPERO, a prospective register of systematic reviews (registration ID: CRD42016037767).[32] Our expert advisory group commented on a draft of the protocol. The reviews were undertaken following the general good practice approaches recommended by the Centre of Reviews and Dissemination (CRD),[33] the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy,[34,35] and the National Institute for Health and Care Excellence Diagnostics Assessment Programme Manual.[36] Here, we outline the methods specified in the protocol and note minor modifications that were made during the review.

3.1 Identification of studies

An experienced information specialist developed and tested a comprehensive search strategy. The strategy was designed to identify studies of the diagnostic accuracy of virtual chromoendoscopy and studies providing relevant clinical outcomes (morbidity, mortality, HRQoL) associated with virtual chromoendoscopy and histopathological diagnosis. The strategy was also designed to capture relevant cost-effectiveness studies, to inform the economic evaluation (Section 5).

The following databases were searched from inception to June 2016 for published research: MEDLINE, PreMedline In-Process & Other Non-Indexed Citations, EMBASE, Web of Science, the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effectiveness (DARE), Health Technology Assessment database, and NHS Economic Evaluation Database (EED). (NB. The protocol for the systematic reviews stated that the Medion database of diagnostic studies would be searched; however, when the review commenced we found that this database had been discontinued.) Grey literature and ongoing studies were also identified, through searches of the following databases in March 2016: UK Clinical Trials Gateway (UKCTG), World Health Organization International Clinical Trials Registry Platform (WHO ICTRP), ISRCTN (controlled and other trials), clinicaltrials.gov, and PROSPERO. (NB. The protocol for the systematic reviews stated that the UK Clinical Research Network Portfolio Database and the NIHR Clinical Research Network Portfolio would be searched but these are now part of the UKCTG). All searches were limited to the English language.

We additionally searched conference proceedings and the internet pages of relevant organisations for publications, both in April 2016. Proceedings from the following conferences were searched: The Association of Coloproctology of Great Britain and Ireland Annual Meeting; the Annual Meeting of the European Society of Coloproctology; the American Society for Gastrointestinal Endoscopy (ASGE) Digestive Disease Conference; Digestive Disease Week Conference; and, United European Gastroenterology (UEG) Week Conference. We searched the following organisations' websites: the British Society of Gastroenterology, the European Society of Gastrointestinal Endoscopy, ASGE, and the American Gastrointestinal Association.

We also searched the bibliographies of the included studies and of relevant systematic reviews found during the searches to identify further references, and asked our advisory group of experts to identify additional published and unpublished studies. Information provided by the companies to the National Institute for Health and Care Excellence was also searched for additional studies that might meet the review inclusion criteria. A full list of databases searched, search dates and an example search strategy are provided in Appendix 1.

3.2 Inclusion and exclusion criteria

We screened all the publications identified from the searches against the pre-specified eligibility criteria set out here, to determine if they should be included in the reviews of clinical effectiveness and cost-effectiveness.

Study design

For the systematic review of clinical effectiveness, studies were eligible for inclusion if they were Randomised Controlled Trials (RCTs), prospective longitudinal cohort studies or cross-sectional studies. Systematic reviews were not included and were only retrieved during screening to check their reference lists for potentially relevant primary research studies. Editorials and case-reports were not included.

For the systematic review of cost-effectiveness, studies were included if they were full economic evaluations, assessing costs and consequences, of the specified virtual chromoendoscopy technologies.

Population

For both the reviews of clinical effectiveness and cost-effectiveness, studies had to include at least one of the following populations to be eligible for inclusion in the review:

- People referred for colonoscopy following an abnormal bowel cancer screening result

- People offered colonoscopic surveillance because they have had adenomas removed
- People with symptoms that may be suggestive of colorectal cancer who are referred for colonoscopy by a GP

As stated earlier (Section 2.1.1) the target population in this assessment does not include people undergoing monitoring for IBD (e.g. Crohn's disease); and people with polyposis syndromes such as HNPCC or FAP. Studies including these populations were therefore excluded.

Index test

Studies were included in both reviews if they evaluated one or more of the technologies of interest for the real-time diagnosis of colorectal polyps (as opposed to post-procedure image-based diagnosis):

- NBI - EVIS LUCERA ELITE, EVIS LUCERA SPECTRUM or EVIS EXERA (Olympus Medical Systems). The EXERA system is not available in the UK but expert advice to the External Assessment Group (EAG) was that diagnostic outcomes are similar to the EVIS LUCERA series.
- FICE (Fujinon/Aquilant Endoscopy)
- I-scan (Pentax Medical)

Studies of these technologies were only included if they used high definition or high resolution endoscopy systems, without the use of magnification (in at least one study arm, in the case of RCTs; arms not meeting this criterion were excluded). These limitations were applied, because, as explained in section 1.2.4, the majority of endoscopy equipment used in practice is (or will be in the future) high definition capable and because magnifying endoscopes are largely unavailable and not considered practical in routine care. During screening, the following decision rules were created to address uncertainty about inclusion of studies in the clinical effectiveness review when they used inbuilt or optional magnification or did not mention magnification:

- studies or study arms using inbuilt (close focus) magnification (which is a low level of magnification, e.g. ×1.5) that did not require a zoom endoscope or any additional equipment were included.
- when magnification was described as optional and no further details were provided or when magnification was not mentioned, we erred on including the study (i.e. presumed no magnification).

Additionally, if a standard definition endoscope was used with a high definition monitor in a study, we excluded the study as this type of monitor cannot compensate for lack of a high definition endoscope.

Studies or study arms using endoscopes with a push-button 'near focus' capability were excluded, as these endoscopes use magnification, unless it was clear that the 'near focus' function had not been used during polyp characterisation.

Reference test (Comparator)

Only studies using histopathological assessment of resected diminutive (≤5 mm in size) colorectal polyps as the reference test were included the reviews. Studies of larger sized polyps were eligible if outcome data were given for a sub-group of diminutive polyps.

Outcomes

Studies had to measure and report results for at least one of the following outcomes to be included in the clinical effectiveness review (none were specified as primary or secondary outcomes for the review):

- Accuracy of virtual chromoendoscopy diagnosis of polyp (e.g. adenoma, hyperplastic)
- Number of polyps designated to be left in place
- Number of polyps designated to be resected and discarded
- Number of polyps designated to be resected and sent for histopathological examination
- Recommended surveillance interval
- Length of time to perform the colonoscopy
- Number of outpatient appointments or telephone consultations
- Health-related quality of life (HRQoL), including anxiety
- Adverse effects of the removal of polyps (i.e. of polypectomy)
- Colorectal cancer
- Mortality

To be included in the cost-effectiveness review, studies needed to measure relevant outcomes including life years, incidence of colorectal cancer or Quality Adjusted Life Years (QALYs).

3.2.1 Inclusion screening process

Reviewers selected studies for inclusion through a two-stage process using the predefined and explicit criteria specified above. Two reviewers independently assessed the titles and abstracts of the publications identified through the searches for potential relevance to the review. We then obtained the full texts of agreed potentially relevant publications for full text screening. During full text screening, one reviewer assessed each publication against the eligibility criteria, using a standardised inclusion flow chart, and

another reviewer checked the first reviewer's decision and a final decision regarding inclusion was agreed. Studies had to meet all of the eligibility criteria to be included in the review. At both stages any disagreements were resolved by discussion, with involvement of a third reviewer where necessary. The inclusion flow chart is shown in Appendix 2. The first item in the flowchart that the reviewers agreed would be a reason for exclusion was recorded as the primary reason for exclusion.

During full text screening, we found that the population was unclear in some of the publications assessed (e.g. due to lack of description). In these instances, we erred on including the study in the review, unless there was evidence that it included a population not relevant to this assessment (e.g. inflammatory bowel disease, polyposis syndromes). Studies published as abstracts or conference proceedings were only included in the reviews if they were published in 2014, 2015 or 2016 and if sufficient details were presented to allow appraisal of the methodology and assessment of results to be undertaken (as pre-specified in the protocol).

3.3 Data extraction strategy

One reviewer extracted data from each included study, using a standardised and pilot-tested data extraction form, and a second reviewer checked the extracted data for accuracy. Reviewers resolved any discrepancies in the data extracted through discussion or, where necessary, arbitration by a third reviewer. Publications that reported the same primary study were data extracted together as one study, to avoid double-counting information. Reviewers extracted data, where available, on the study and population characteristics, the endoscopic equipment used (including model numbers), the study endoscopists' experience and training, the polyp classification system used, the sample size calculation, and results for all outcomes of interest in this review. Where data were available, we extracted the results of subgroup analyses of diagnostic accuracy by the endoscopists' level of expertise and experience in optical assessment of polyps, their level of confidence in their polyp assessment (i.e. high or low), and the location of the polyp. See Appendix 3 for the completed data extraction form for each study.

When we extracted the diagnostic accuracy results from each study, we used available data in the study publication(s) to populate a 2×2 contingency table showing how the index test results related to the histopathological analysis results, for each analysis or subgroup analysis of diminutive polyps. The contingency tables showed the number of true positives, false positives, true negatives and false negatives. Where these data were only partially reported in the study publications or not reported at all, reviewers imputed the data from other available results information, if possible. It was necessary to extract or impute these data, as we needed complete 2×2 tables to be able to include a study in a meta-analysis (see

Section 3.5 for further details about data synthesis). It was not always possible to impute these data (e.g. total number of diminutive polyps not reported and numbers of adenomas and hyperplastic polyps not reported). We contacted the contact study author for five studies to request the 2x2 table data. Two authors replied but neither were able to supply data. Reviewers also calculated the accuracy (proportion of correctly classified polyps among all the polyps), clinical sensitivity, clinical specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, negative likelihood ratio and diagnostic odds ratio for each diagnostic accuracy analysis and subgroup analysis reported in each study. Reviewers compared the values they calculated with the study values and noted any discrepancies. If any of these outcomes had not been reported in the studies, the reviewer's calculated values were used. We used an online calculator MedCalc (https://www.medcalc.org/calc/diagnostic_test.php) to calculate clinical sensitivity, clinical specificity, PPV, NPV, and positive and negative likelihood ratios.

3.4 Quality assessment

The quality of studies reporting diagnostic accuracy was assessed using the Cochrane Collaboration adaptation[37] of the QUADAS tool (Quality Assessment Tool for Diagnostic Accuracy Studies)[38] which can be used to assess a variety of study designs (e.g. RCT, non-RCT, prospective cohort studies). Table 3 shows the types of bias assessed by the QUADAS tool and how these we assessed whether these types of bias were present in studies in this review. One reviewer assessed the methodological quality of each study and a second reviewer checked the first reviewer's judgements, with any disagreements resolved by consensus or if necessary by arbitration by a third reviewer.

Table 3 Types of bias assessed by the QUADAS tool and their application to studies of the accuracy of virtual chromoendoscopy for the real-time assessment of colorectal polyps in vivo

QUADAS question	Type of bias	Explanation
1	Spectrum bias	The study population is not representative of those who will receive the index test (virtual chromoendoscopy i.e. NBI, i-scan or FICE) in clinical practice
2	Verification bias	The reference standard (histopathology) does not accurately distinguish between adenomas and hyperplastic polyps
3	Disease progression bias	The time interval between the index (virtual chromoendoscopy) test and reference standard (histopathology) is long enough that the two

		tests may not have measured the same disease state
4, 5[a]	Differential verification bias	Diagnosis is inaccurate because not all patients receive the same reference standard
6	Incorporation bias	The index (virtual chromoendoscopy) test is not independent of the reference standard (e.g. if it was one of several tests used as the reference standard)
7	Diagnostic review bias	The index test (virtual chromoendoscopy) result influences interpretation of the reference standard result
8	Test review bias	The reference standard result influences interpretation of the index (virtual chromoendoscopy) test result
9	Clinical review bias	The information used when interpreting the index (virtual chromoendoscopy) test does not reflect that likely to be available in clinical practice
10	Test classification bias	If index test results classified as uninterpretable, intermediate or indeterminate are incorrectly included or excluded from the analysis, this may systematically influence sensitivity or specificity
11	Attrition bias	The exclusion of patients or test results from the analysis may systematically influence sensitivity or specificity if • the reason for exclusion is linked to test performance • if criteria for permitting exclusions differ between tests This is particularly the case if the magnitude of attrition is unbalanced across the test methods

[a] Two QUADAS questions assess differential verification bias.

3.5 Method of data synthesis

The included studies were synthesised in a narrative review with tabulation of results. Meta-analysis was also conducted to provide pooled estimates of diagnostic sensitivity and specificity. The rationale for meta-analysis was to provide a more precise estimate of diagnostic accuracy than can be provided from single studies alone. In diagnostic test studies, sensitivity and specificity are often negatively correlated, sometimes because studies have used different thresholds for defining positive and negative test results. Furthermore, heterogeneity often exists between the studies, in terms of patient characteristics, settings, and tests used. These factors need to be taken into account in the choice of meta-analysis methods applicable to a given topic. A univariate meta-analysis pools sensitivity and specificity separately, failing to take into account the correlation. Hierarchical models include statistical distributions at the lower level

(within-study variability in sensitivity and specificity) and at the higher level (between-study variability) and can therefore take into account correlation and heterogeneity.[39] In this systematic review it is likely that heterogeneity exists in factors such as the endoscopist's level of experience and training in virtual chromoendoscopy, the setting in which colonoscopy is performed, and the patient's indication for colonoscopy and therefore their risk of colorectal cancer. Virtual chromoendoscopy does not require an explicit numerical threshold for a diagnostic prediction. Rather, the prediction is a binary one, of whether a polyp is an adenoma or hyperplastic. A hierarchical bivariate meta-analysis model was used in this assessment as it estimates summary sensitivity and specificity at a various thresholds (in this case the threshold is the confidence and judgement with which the endoscopist makes their polyp characterisation).[40] Previous published meta-analyses of virtual chromoendoscopy for optical diagnosis of colorectal polyps have also used a bivariate model to estimate pooled sensitivity and specificity.[41-43]

We conducted separate meta-analyses for the each of the three respective virtual chromoendoscopy technologies relevant to this report compared with histopathology. For each technology we produced individual meta-analyses according to the level of confidence with which the polyp characterisation had been made by the endoscopist in accordance with how the data were reported in the primary studies (high confidence predictions; all predictions irrespective of confidence level). High confidence predictions are of particular relevance to the DISCARD strategy and are used to inform the economic model in this assessment report (see Section 5.2). We also meta-analysed studies according to the area of the colon in which the polyps were located and thus characterised (e.g. whole colon, rectosigmoid colon), stratified according to level of endoscopist confidence in making characterisations. Again, this is relevant to the DISCARD strategy for decisions about whether hyperplastic polyps in the rectosigmoid colon can be left in situ (see Section 1.3). Where possible we explored heterogeneity by conducting sub-group analyses for factors such as the level of experience of the endoscopist in the in vivo characterisation of polyps, and in using the specific virtual chromoendoscopy (see Section 4.1.1 for a description of the studies included in the systematic review).

Consideration was given to meta-analysing NPVs from the included studies. An NPV of ≥90% is required for a high confidence decision to leave a suspected hyperplastic diminutive polyp in place, as stated in the PIVI initiative[31] (see Section 1.3). However, PPVs and NPVs vary with differences in disease prevalence, so pooling is not always advisable when it is suspected that there may be variation in prevalence between studies.[36] Because the prevalence of adenomas and hyperplastic polyps may vary between studies [e.g. due to differences in case mix (screening, surveillance and symptomatic populations) and patient characteristics (age, sex)] we chose not to pool NPV values across studies.

We used Stata software (Stata 14.0 IC, Stata Corp, Texas) to conduct the meta-analysis, using the metandi Stata package which has been specifically designed to perform bivariate meta-analyses of diagnostic studies.[44] The Stata package xtmelogit was also used where fewer than four studies were available in a meta-analysis, as metandi was not able to perform analyses on this number of studies. We used Stata programming code supplied by the Cochrane Screening and Diagnostic Tests Methods Group for bivariate meta-analysis models.[45] Four input variables were used by Stata to perform the meta-analysis: the number of true positives, false positives, false negatives and true negatives for each study (the unit of analysis is the individual polyp). These were taken from our data extraction forms for each included study and included in a spreadsheet from which Stata directly drew the data. We also used Cochrane Review Manager (RevMan)[46] to produce coupled forest plots of sensitivity and specificity and Summary Receiver Operating Curve (SROC) plots. The forest plots allow a visual interpretation of the individual study estimates, which can be informative in the assessment of heterogeneity. The SROC plots provide confidence and prediction regions around the summary estimate to enable joint inferences to be made about sensitivity and specificity. The confidence region is based on the confidence interval around the summary estimate. The prediction region indicated the area where we would expect results from a new study in the future to lie.[39]

4 ASSESSMENT OF DIAGNOSTIC STUDIES

4.1 Results

4.1.1 Quantity and quality of research available

A total of 2068 references was identified by searches (after de-duplication) and two additional references were identified through other sources (Figure 4). We screened the titles and, where available, abstracts of the 2070 references and retrieved full copies of 125 references. We excluded 63 full text references, the majority either because the intervention (n=28) or comparator (n=29) did not meet the inclusion criteria (a list of the excluded studies with reasons for exclusion is presented in Appendix 4). Twenty-four references were designated as 'Unclear', all of which were conference abstracts (seven[47-53] of these could be linked to full papers already either included or excluded and 17 appear to be ongoing or recently completed studies, see section 4.2). The remaining 32 references met the inclusion criteria of the systematic review and were included. These 32 references describe 30 separate studies.

The majority of the 30 studies that met the inclusion criteria for this systematic review evaluated NBI (n= 24) with 2 of these also evaluating one of the other interventions of interest (NBI & i-scan n=1; NBI & FICE n=1). A further 4 studies evaluated i-scan and a further 2 studies evaluated FICE. Thus the final tally of included evidence is as shown in Table 4.

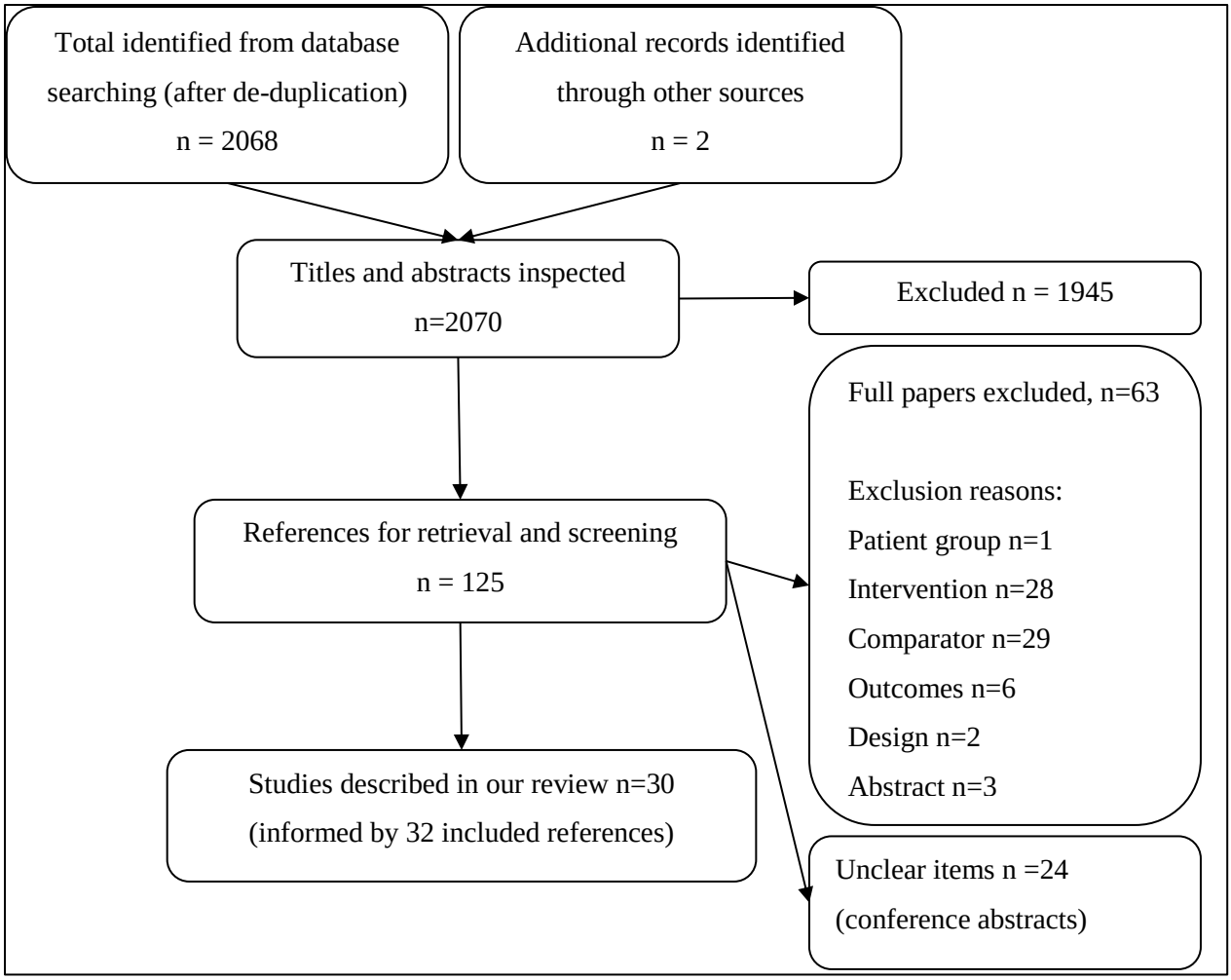


Figure 4 Flow chart for the identification of studies

Table 4 Evidence meeting the criteria for the systematic review

Interventions	**Number of studies**
NBI	22[20,54-76]
NBI & i-scan	1[77]
NBI & FICE	1[78]
i-scan	4[79-82]
FICE	2[83,84]

NBI

Twenty-four studies[20,54-78] included in the systematic review provided data on the use of NBI for virtual chromoendoscopy of colorectal polyps. From here on in the report Kaltenbach and colleagues[62,64] and Gupta and colleagues[56,69] will be identified by a single study reference to the main source of data

(Kaltenbach and colleagues[62] and Gupta and colleauges[56]). Two of these studies, a prospective cohort study by Lee and colleagues[77] and an RCT by Kang and colleagues[78] also reported on i-scan and FICE respectively and so are also included in our report in the i-scan and FICE sections.

An overview of the characteristics of the included NBI studies is presented in Table 5 (more detailed information is available in the data extraction forms presented in Appendix 3). More than half of the studies were conducted in the USA (14 studies[20,54,56-58,62,63,67,68,71-73,75,76]). Five studies were conducted in Europe (One in the UK,[59] two studies in Italy,[65,66] one in Italy and the Netherlands[70] and one in Spain[74]). The remaining five studies were conducted in Asia: two in Japan[60,61] and two in South Korea;[77,78] and Australia.[55] Seven of the studies focussed on diminutive polyps,[55,56,62,65,67,75,77] nine focussed on small polyps (<10mm in size)[20,59-61,66,70,73,74,78] and eight included polyps of any size.[54,57,58,63,68,71,72,76] The studies that included polyps larger than diminutive polyps provided at least one outcome of interest for the sub-group of diminutive polyps. One study, by Hewett and colleagues 2012a[58] restricted their study to polyps in the rectosigmoid colon.
Half of the studies enrolled participants undergoing colonoscopy either for screening, surveillance or because of symptoms [20,55,57,59,62,65,66,68,70,72,74,76] with all but two (Hewett and colleagues 2012b[20] and Patel and colleagues[67]) reporting the proportions of participants in each category. Five studies enrolled participants undergoing colonoscopy for either screening or surveillance reasons[56,58,73,75,77] but not because of symptoms, with one more study[54] including participants presenting for elective screening or follow-up colonoscopy (reasons for the follow-up colonoscopy not provided). In two studies the entire sample of participants was drawn from a screening population[60,78] In the remaining three studies the types of participants enrolled is not known because it was not reported in the publications.[61,63,71]

The male:female ratio of participants in the included studies lay between 1:1 and 2:1 in 13 studies,[54,57,58,61,65,66,68,70,72-76] and between 2:1 and 3:1 in three studies.[59,77,78] In the remaining four studies that reported the male:female ratio it was approximately 4:1,[60] 10:1,[56] 23:1[62] and the highest reported male:female ratio was 35:1.[55] The male:female ratio of participants was not reported by four studies.[20,63,67,71]

The mean age of participants, if it was reported, lay between 54 years and 67 years (16 studies[54,56,58-62,65,66,68,70,72,74,75,77,78]) or the median age lay between 60 and 69 years (four studies[55,57,73,76]). The age of participants was not reported by the remaining four studies.[20,63,67,71]

The majority of the studies were conducted in a single centre,[57-59,61,65,66,71-78] four were conducted in two centres,[55,56,60,68] and one each at three centres,[62] four centres[67] and five centres.[70] The number of centres was not reported by three studies.[20,54,63]

Study colonoscopies were undertaken by more than one endoscopist in most studies: one endoscopist in five studies,[57,58,71,73,77] two in one study,[20] three in one study,[55] four in four studies,[59,65,72,78] five in four studies,[61,62,70,74] six in three studies,[56,66,75] seven in three studies,[54,60,76] 10 in one study,[68] 12 in one study[63] and the largest number of endoscopists was 26 in one study.[67] In eight studies all the endoscopist(s) had prior experience of using NBI[55,56,58,60,65,66,70,77] and in four studies some of the endoscopists had prior experience of using NBI.[59,61,62,74] Only four studies stated that the endoscopists involved had no prior experience of using NBI to characterise colorectal polys[63,67,68,78] but there were a further eight studies where it was not clear what experience of using NBI, if any, the endoscopist(s) may have had.[20,54,57,71-73,75,76] The majority of the studies included an element of training for the endoscopist(s) in the characterisation of colorectal polyps using NBI, either training all endosopists[20,54,55,57,62,63,65-68,70-72,74-76,78] or the non-experts.[59] In the study by Gupta and colleagues, which is a re-analysis of three earlier studies, training occurred in one of the three studies.[56] There were five studies[58,60,61,73,77] that did not state any training had taken place. In three of these, the endoscopists had prior experience of NBI.[58,60,77] In the Iwatate and colleagues' study[61] the five endoscopists had mixed levels of NBI experience, and it was unclear what NBI experience the single endoscopist in the Shahid and colleagues' study had.[73]

A variety of different systems were used to classify polyps as adenomas or hyperplastic polyps (Table 5). The most commonly used systems were the NICE classification scheme or a version of this which was cited by eight studies[20,54,61-63,65,74,76] and the criteria proposed by Rex and colleagues[71] which were cited by four studies.[58,66,71,78] Two studies[55,57] cited the Sano-Emura classification system, two[72,73] based characterisations on modifications of the Kudo criteria and two[56,67] on work by Rastogi and colleagues[69,85-87] with one further study[75] also citing a Rastogi and colleagues publication[88] although it is not known in this case whether the criteria were the same. One study[59] used vascular pattern intensity[89] to classify polyps, one[68] polyp colour, vessels and mucosal pattern[90] and one[77] the author's own system. In the final two studies either criteria were reported by not attributed to any named system[70] or no criteria were reported or cited.[60]

Table 5 Overview of NBI studies

Study	Country	Centres	Patient population[a]				Patient characteristics		NBI Processor	Endoscopists			Classification
			n or n/N[b]	SCR (%)	SURV (%)	SYM (%)	Age, mean (SD) or median [range][c]	sex (M/F %)		n	NBI experience	Training	
Aihara et al.[54]	USA	NR[d]	NR/67	Y[e]	NR[e]	NR	54 (NR)	64/36	NR	7	Unclear	Yes	NICE-AS[54]
Chandran et al.[55]	Australia	2	94	27	34	28	62 [19 to 84]	97/3	EXERA	3	Yes	Yes	Sano-Emura[91]
Gupta et al.[56]	USA	2	NR/410	Y	Y	N	62 (8)[f]	90/10[f]	EXERA II	6	Yes	Yes (1/3 trials)	Author's[69,85,86]
Henry et al.[57]	USA	1	NR/52	29[f]	42[f]	27[f]	60 [34 to 84][f]	63/37[f]	EXERA II	1	Unclear	Yes	Sano-Emura[91,92]
Hewett et al. 2012a[58]	USA	1	31/255	29[f]	45[f]	NR	60 (10)[f]	52/48[f]	EXERA II	1	Yes	No	Rex et al. publication[71]
Hewett et al. 2012b[20]	USA	NR	NR/108	Y	Y	Y[g]	NR	NR	EXERA II	2	Unclear	Yes	NICE – no reference cited
Ignjatovic et al.[59]	UK	1	NR/130	25	63	12	63 (11)[f]	67/33[f]	LUCERA	4	Mixed	Of non-experts	Vascular pattern intensity

Study	Country	Centres	Patient population[a]				Patient characteristics		NBI Processor	Endoscopists			Classification
			n or n/N[b]	SCR (%)	SURV (%)	SYM (%)	Age, mean (SD) or median [range][c]	sex (M/F %)		n	NBI experience	Training	
Ikematsu et al.[60]	Japan	2	NR/37	100	No	No	67 (NR)[f]	76/24[f]	LUCERA	7	Yes	No	None stated
Iwatate et al.[61]	Japan	1	NR/124	NR	NR	NR	56 (9)[f]	58/42[f]	LUCERA	5	Mixed	No	NICE[20,93]
Kaltenbach et al.[62]	USA	3	NR/281	38[f]	44[f]	19[f]	62 (9)[f]	96/4[f]	EXERA II	5	Mixed	Yes	NICE[20]
Kang et al.[78 h]	South Korea	1	203/399	100	N	N	55 (9)	68/32	LUCERA	4	No	Yes	Polyp colour, vessels and surface pattern[71,94,95]
Ladabaum et al.[63]	USA	NR	NR	NR	NR	NR	NR	NR	EXERA II	12	No	Yes	NICE [96]
Lee et al.[77 h]	South Korea	1	70/142	Y	Y	N	58 (11)	74/26	LUCERA	1	Yes	No	Author's
Paggi et al. 2015[65]	Italy	1	NR/284	43[f]	28[f]	30[f]	61 (18)[f]	63/37[f]	EXERA	4	Yes	Yes	Based on published criteria[20]

Study	Country	Centres	Patient population[a]				Patient characteristics		NBI Processor	Endoscopists			Classification
			n or n/N[b]	SCR (%)	SURV (%)	SYM (%)	Age, mean (SD) or median [range][c]	sex (M/F %)		n	NBI experience	Training	
Paggi et al. 2012[66]	Italy	1	197/286	37[f]	26[f]	36[f]	60 (16)[f]	56/44[f]	EXERA	6	Yes	Yes	Simplified NBI criteria as proposed by Rex et al.[71]
Patel et al.[67]	USA	4	451	Y	Y	Y	NR	NR	EXERA II	26	No	Yes	Previously established NBI criteria[69,86,87]
Pohl et al.[68]	USA	2	566/607	53[i]	30[i]	9[i]	62 (8)[i]	64/36[i]	NR	10	No	Yes	Polyp colour, vessels and mucosal pattern[90]
Repici et al.[70]	Italy and The Nether-lands	5	212/278	37[f]	27[f]	36[f]	63 (10)[f]	58/42[f]	NR	5	Yes	Yes	Criteria reported, but not attributed to any named system

Study	Country	Centres	Patient population[a]				Patient characteristics		NBI Processor	Endoscopists			Classification
			n or n/N[b]	SCR (%)	SURV (%)	SYM (%)	Age, mean (SD) or median [range][c]	sex (M/F %)		n	NBI experience	Training	
Rex et al.[71]	USA	1	NR/136	NR	NR	NR	NR	NR	EXERA HD 180	1	Unclear	Yes[j]	Author's[71] [also used by Hewett et al[58]]
Rogart et al.[72]	USA	1	NR/131	55	24	15	59 (10)	65/35	EXERA II	4	Unclear (without extensive experience)	Yes	Simplified Kudo pit-pattern classification[22]
Shahid et al.[73]	USA	1	NR/65	Y	Y	N	69 [44 to 91][f]	62/38[f]	EXERA	1	Unclear	No	Kudo criteria as modified by Sano et al[97]
Sola-Vera et al.[74]	Spain	1	NR/195	38[f]	16[f]	25[f]	64 (12)[f]	56/44[f]	EXERA	5	1/5	Yes	NICE[20,93]
Vu et al.[75]	USA	1	315	48	52	N	62 (9)	51/49	EXERA II	6	Unclear	Yes	Based on Rastogi[88]
Wallace et al.[76]	USA	1	NR/264	46	43[f]	10[f]	60 [33 to 85][f]	58/42[f]	EXERA II	7	Unclear	Yes	Simplified NICE[63]

NR, not reported; SCR - Screening; SURV - Surveillance; SYM - Symptomatic.

[a] If studies reported categories that appeared to fit under the 'Screening', 'Surveillance' or 'Symptomatic' headings these were grouped together. Some studies reported categories that did not fit under the 'Screening', 'Surveillance' or 'Symptomatic' headings or were described as 'Other' and these have not been reported. Percentages were rounded to whole numbers. Consequently the sum of percentages for some studies does not sum to 100%.

[b] The number of patients (n) for studies reporting only on diminutive polyps or the number of patients with diminutive polyps over the number of patients in the study overall (n/N) for studies reporting on diminutive polyps and larger polyps.

[c] Values rounded to the nearest whole number due to space limitations in the table.

[d] Number of centres not reported, however as all authors were affiliated to the same hospital, this is likely to have been a single centre study.

[e] Participants presented for elective screening or follow-up colonoscopy (reason for follow-up colonoscopy not reported).

[f] Results based on the total population and not available for the diminutive polyp subgroup (≤5 mm diminutive polyps)

[g] Described as 'Diagnostic'

[h] Study included an arm that is included elsewhere in this report. Data reported here related only to the NBI arm of the study.

[i] Values based on 1100 participants who had a colonoscopy but at least one polyp was found in only 607 participants.

[j] This study contained an element not described as training by the study author but which the review team considered could be described as training.

The QUADAS assessments of the NBI studies indicates that the studies were at a low risk of spectrum, verification, disease progression, incorporation, test review, and clinical review biases (Table 6). Supporting information for the judgements shown in Table 6 is provided in the data extraction form for each study Appendix 3). Note that 'Yes' answers to QUADAS questions 1 to 9 (Table 3) imply a low risk of bias whereas 'Yes' answers to QUADAS questions 10 and 11 reflect adequacy of reporting and further supporting information is required to assess the risks of bias associated with these questions. For five studies[54,61,63,67,71] the risk of spectrum bias (QUADAS question 1) was unclear because the reason(s) for patients having a colonoscopy were not reported. In two studies[62,76] not all the polyps received verification by histopathology. In the Kaltenbach and colleagues' study[62] this was because when two or more non-neoplastic polyps were identified in the rectosigmoid colon in any one patient, a "representative sample" was resected for histopathological analysis. How often this circumstance arose was not reported. In the Wallace and colleagues' study[76] 10 polyps (from 321 polyps, therefore representing 3% of the total) were not assessed by histopathology (and whether one further polyp had been assessed by histopathology was unclear). Overall it is our opinion that the risk of differential verification bias in these two studies was probably very low.

In all but four studies[57,65,68,76] the risk of diagnostic review bias was judged to be low (QUADAS question 7) but was unclear in the studies by Henry and colleagues,[57] Paggi and colleagues 2015,[65] Pohl and colleagues[68] and Wallace and colleagues[76] because they did not report whether the histopathologist(s) were blinded to the NBI prediction for each polyp. The majority of studies did not report on uninterpretable/ intermediate test results probably because there were no uninterpretable/ intermediate test results due to the nature of the NBI assessments (studies typically required a decision to be made, although this could be assigned as low confidence in some studies). In the studies by Gupta and colleagues and Iwatate and colleagues there was evidence for uninterpretable or intermediate test results studies.[56,61] An optical diagnosis could not be determined for four polyps (0.3%) in the study by Gupta and colleagues[56] and Iwatate and colleagues[61] excluded two patients with 'unevaluable material'. Patel and colleagues[67] reported that polyps were excluded from the analysis if a confidence level was not assigned or if histology was missing, "other", or if the polyp could not be retrieved so it seems likely that there were also some uninterpretable or intermediate test results in this study. The outcome for QUADAS item 10 was judged unclear for the Wallace and colleagues' study because not all patients who were randomised completed the study, so it is possible that uninterpretable test results were the reason for the missing data.

For the final QUADAS item (number 11, attrition bias) the judgement was 'Yes' for the majority of studies either because no withdrawals were apparent in the study[20,54,55,57,58,60,61,65,66,71-75,77] or because withdrawals or other missing data were explained.[59,62,68,70,76,78] For two studies the judgement was 'Unclear'. In the Ladabaum and colleagues' study[63] this was because endoscopists were considered the subjects in the study and it was unclear whether any of them had dropped out of the study and because endoscopists were considered the subjects there was little reporting on those undergoing colonoscopy. In the Patel and colleagues' study [67] the authors did not report the number of participants selected to take part or the number of patients included in the data analyses so it was unclear whether there had been any withdrawals. For one study, Gupta and colleagues[56] this question was not applicable, because the included data were drawn from records of participants in three earlier trials that met the inclusion criteria for a retrospective analysis and therefore no participants were able to withdraw.

In addition to the assessment of the QUADAS items the generalisability of each study was also briefly summarised during data extraction (the summary of reviewers' comments can be seen in full in the data extraction forms in Appendix 3). The overall impression from the included NBI studies is that they enrolled participants likely to be representative of the types of participants who would receive colonoscopy in the UK for screening, surveillance or on account of symptoms experienced (in line with the inclusion criteria for this systematic review). However, only one study was conducted in the UK, and just four elsewhere in Europe where it might reasonably be assumed that populations might be most similar to those in the UK. Most studies were conducted in a single centre so inherently these results may not be transferrable to other centres. In contrast, in most studies more than one endoscopist was involved in conducting colonoscopies and characterising polyps. Across all the studies a range of endoscopists was involved, some who were less experienced in conducting colonoscopy generally and had little or no experience using NBI through to very experienced endoscopists who also had extensive experience of using NBI. Training for endoscopists (which may have been to train those with no prior experience of NBI or to ensure that all endoscopists at a centre were characterising polyps to the same standard) formed a part of the majority of studies but how representative this training may have been to current UK practice is unknown. Finally a variety of classifications systems were used to determine whether polyps were adenomas or hyperplastic. The assessment group understands that, in countries such as the UK where polyp characterisation is conducted without magnification, the NICE classification is becoming widely accepted. It is unclear how generalisable the results obtained using other polyp classifications are to UK practice.

Table 6 Overview of NBI QUADAS assessments

	QUADAS ITEM (Questions are available in table footnotes)										
Study	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**
Aihara et al.[54]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Chandran et al.[55]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Gupta et al.[56]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	n/a
Henry et al.[57]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Hewett et al. 2012a[58]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Hewett et al. 2012b[20]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Ignjatovic et al.[59]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Ikematsu et al.[60]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Iwatate et al.[61]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kaltenbach et al.[62]	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Kang et al.[78]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Ladabaum et al.[63]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Unclear
Lee et al.[77]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Paggi et al. 2015[65]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Paggi et al. 2012[66]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Patel et al.[67]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Pohl et al.[68]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Repici et al.[70]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Rex et al.[71]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Rogart et al.[72]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Shahid et al.[73]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Sola-Vera et al.[74]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Vu et al.[75]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Wallace et al.[76]	Yes	Yes	Yes	No	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes

Q1 Was the spectrum of patients representative of the patients who will receive the test in practice? Q2 Is the reference standard likely to classify the target condition correctly? Q3 Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? Q4 Did the whole sample or a random selection of the sample, receive verification using the intended

reference standard? Q5 Did patients receive the same reference standard irrespective of the index test result? Q6 Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)? Q7 Were the reference standard results interpreted without knowledge of the results of the index test? Q8 Were the index test results interpreted without knowledge of the results of the reference standard? Q9 Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? Q10 Were uninterpretable/ intermediate test results reported? Q11 Were withdrawals from the study explained?

i-scan

Five studies[77,79-82] included in the systematic review provided data on the use of i-scan for virtual chromoendoscopy of colorectal polyps. An overview of the characteristics of the included i-scan studies is presented in Table 7 (more detailed information is available in the data extraction forms presented in Appendix 3). Four of the studies were conducted in Europe (Basford & colleagues in the UK,[79] Hoffman and colleagues[80] and Rath and colleagues[82] in Germany, Pigo and colleagues[81] in Italy) and one, Lee and colleagues,[77] was conducted in South Korea. Basford & colleagues[79] and Hoffman and colleagues[80] enrolled all their participants from a screening population whereas the other three studies[77,81,82] enrolled participants receiving colonoscopy for screening or surveillance purposes with one[81] also including participants with gastrointestinal symptoms. In the three studies[77,81,82] that enrolled different types of participants the proportions of participants receiving colonoscopy for screening, surveillance or symptoms was not reported. The Pigo and colleagues' study[81] enrolled almost equal proportions of men and women whereas more men than women were enrolled in the other four studies. Four studies[77,80-82] reported the mean age of the participants which ranged from 55 years to 66 years. The two studies conducted in Germany[80,82] did not report data on polyp characterisation for the whole colon, Hoffman and colleagues only reported on polyps in the last 30cm of colon and Rath and colleagues characterised polyps in the distal colon (decending colon, the sigmoid colon or the rectum). Three of the studies (Hoffman and colleagues,[80] Lee and colleagues[77] and Rath and colleagues[82]) focussed on the characterisation of diminutive polyps whereas Basford and colleagues[79] focussed on small polyps (<10mm) and Pigo and colleagues[81] included polyps of all sizes (and their data on diminutive polyps were limited to the rectosigmoid colon). Consequently, for the three studies that focussed on the characterisation of diminutive polyps, data are drawn from the whole patient population whereas it is not clear what proportion of the patients contributed data on diminutive polyp characterisation in the Basford and colleagues[79]and Pigo and colleagues[81] studies. All the studies were conducted in single centres and, in all but one study, a single endoscopist performed the study colonoscopies and characterised polyps. In the Hoffman and colleagues' study[80] three endoscopists were involved. It was clearly reported in three of the five studies (Basford and colleagues,[79] Hoffman and colleagues[80] and Lee and colleagues[77]) that the

endoscopist(s) had prior experience using i-scan but, due to an absence of reported details, it is not clear whether study endoscopists underwent any specific training with i-scan prior to the start of the studies. Only two studies[77,82] used the same system, which was developed for the Lee and colleagues' study,[77] to classify polyps as adenomas or hyperplastic polyps (Table 7) the remainder all used different systems. One study[81] cited the NICE classification system, one[80] used surface pit pattern citing references of Kudo and colleagues among others, and Basford and colleagues[79] developed their own system for their research.

The QUADAS assessments were conducted for each study and supporting information for the judgements shown in Table 8, is provided in the data extraction form for each study (Appendix 3). Note that 'Yes' answers to QUADAS questions 1 to 9 imply a low risk of bias whereas 'Yes' answers to QUADAS questions 10 and 11 reflect adequacy of reporting and further supporting information is required to assess the risks of bias associated with these questions. The QUADAS assessments of the i-scan studies indicate that the studies were at a low risk of spectrum, verification, disease progression, differential verification, incorporation, diagnostic review, test review, clinical review and test classification biases (Table 8). An exception is that, in the Hoffman and colleagues' study,[80] it was unclear how representative the patients were of those who would receive the test in practice because few details about the participants were reported, although it is known that they fulfilled the criteria for screening colonoscopy.

None of the studies indicated that there had been any uninterpretable or intermediate test results reported. Hoffman and colleagues[80] reported results for normal mucosa in addition to adenomatous and hyperplastic polyps, but there is no indication in the paper that this was due to any difficulty in interpreting the index test.

No withdrawals (of patients or of polyps from the analysis) were apparent in the Hoffman and colleagues[80] and Lee and colleagues[77] studies. The exclusion of patients screened for inclusion but who were excluded from participation was explained by Basford and colleagues.[79] Pigo and colleagues[81] recruited 78 patients and 150 polyps were included in the analysis, but it was not clear whether the 150 polyps were from the full sample of 78 recruited participants. Rath and colleagues[82] recruited 224 patients to their study but the analysis included only 77 of these (all were described as having distal diminutive polyps). It is possible that the remaining patients in these studies had larger-sized polyps located other than in the distal colon, but this is not explicitly stated. Therefore the Pigo and colleagues[81] and the Rath and colleagues[82] studies are at possible risk of attrition bias.

In addition to the assessment of the QUADAS items the generalisability of each study was also briefly summarised during data extraction (the summary of reviewer's comments can be seen in full in the data extraction forms in Appendix 3). The overall impression from the included i-scan studies is they enrolled participants likely to be representative of the types of participants who would receive colonoscopy in the UK for screening or surveillance or on account of symptoms experienced. However, only one study was conducted in the UK, with three out of the remaining four conducted in Europe (two in Germany and one in Italy) whilst the final study was conducted in South Korea. Three of the five studies were conducted by endoscopists with prior experience of i-scan and all took place in single centres often described as academic or specialist centres. The results of these studies may therefore not be applicable to less experienced endoscopists working in more generalist or community settings. Only one study used the NICE classification system (which is becoming widely accepted for polyp characterisation without magnification) to determine whether polyps were adenomas or hyperplastic. It is unclear how generalisable the results obtained using other polyp classifications are to UK practice.

Table 7 Overview of the i-scan studies

Study	Country	Centres	Patient population				Patient		Endoscopists			Classification
			n	SCR	SURV	SYM	mean age, years (SD)	sex (% M:F)	n	i-scan experience	Training	
Basford et al.[79]	UK	1	84[a]	100%	n/a	n/a	nr[b]	65:35	1	Yes	Unclear[c]	Developed by the endoscopist for this study.
Hoffman et al.[80d]	Germany	1	69	100%	n/a	n/a	55.9	62:38	3	Yes	nr	Surface pit pattern
Lee et al.[77e]	South Korea	1	72	Yes[f]	Yes[f]	No	55.4 (11.3)	86:14	1	Yes	nr	Developed by the endoscopist for this study.
Pigo et al.[81g]	Italy	1	78[a]	Yes[h]	Yes[h]	Yes[h]	52 (9)	51:49	1	nr	nr	NICE
Rath et al.[82i]	Germany	1	77	Yes[f]	Yes[f]	No	65.5 (14.4)	64:36	1	nr[j]	nr	Used that developed by Lee 2011[77]

NICE - NBI International Colorectal Endoscopic Classification; nr - not reported; SCR - Screening; SURV - Surveillance; SYM - Symptomatic

[a] The value of n reported is for the whole study because the number of participants with diminutive polyps was not reported separately. In Basford 2014 82% of the polyps were ≤ 5mm in size, in Pigo 58.7% of the polyps were ≤ 5mm in size

[b] Although the mean age was not reported the age range for the UK Bowel Screening Programme is 60-74 years.

[c] States that the endoscopist underwent a period of familiarisation with the endoscope and imaging technology which included developing the novel classification system used for the assessment of polyps by using i-scan during the study.

[d] This study allowed the optional use of magnification (level not stated) but the proportion of polyps characterised with the use of magnification was not reported. In addition the data on polyps only relates to the last 30cm of the colon.

[e] Lee 2011 also included an NBI arm which is reported in the earlier section on NBI and Table 5.

[f] The population is described as undergoing screening or surveillance colonoscopy but the proportions in each group are not stated.

[g] For diminutive polyps, data are only reported for rectosigmoid colon.

[h] The paper reports the number of participants for each of four indications for colonoscopy but it appears likely that participants could be included in more than one category because the totals sum to 87 but only 78 participants were included in the study. The indications for colonoscopy were: positivity for fecal occult blood test (51/78; 65.4%), polypectomy follow-up (20/78; 25.6%), gastrointestinal symptoms (7/78; 9.0%), and colorectal cancer familiarity (9/78; 11.5%).

[i] The focus of the study was characterisation of polyps in the distal colon (descending colon, the sigmoid colon, or the rectum).

[j] The endoscopist is described as experienced with no further details so it is not known whether the endoscopist had prior experience of i-scan.

Table 8 Overview of i-scan QUADAS assessments

QUADAS ITEM		Basford 2014[79]	Hoffman 2010[80]	Lee 2011[a77]	Pigo 2013[81]	Rath 2015[82]
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Yes	Unclear	Yes	Yes	Yes
2	Is the reference standard likely to classify the target condition correctly?	Yes	Yes	Yes	Yes	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes	Yes	Yes	Yes	Yes
4	Did the whole sample or a random selection of the sample, receive verification using the intended reference standard?	Yes	Yes	Yes	Yes	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes	Yes	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes	Yes	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	Yes	Yes	Yes	Yes
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Yes	Yes	Yes	Yes
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Yes	Yes	Yes	Yes	Yes
10	Were uninterpretable/ intermediate test results reported?	No	No	No	No	No
11	Were withdrawals from the study explained?	Yes	Yes	Yes	Unclear	No

[a] Note that this is duplicate information because Lee 2011[77] also contained an NBI arm and thus is also represented in the QUADAS table for NBI studies (Table 6).

FICE

Three studies included in the systematic review (Kang and colleagues;[78] Longcroft-Wheaton and colleagues (2012);[83] and Longcroft-Wheaton and colleagues (2011)[84]) provided data on the use of FICE for virtual chromoendoscopy of colorectal polyps (Table 9). Two of the studies were conducted in the UK[83,84] and the other study was conducted in South Korea.[78] In all three of these studies, all the included participants were undergoing colonoscopy for screening purposes. The Longcroft-Wheaton and colleagues (2012)[83] study enrolled a slightly higher proportion of women than men, whereas the other two studies enrolled a higher proportion of men than women. All three studies reported the mean age of participants which ranged from 54 years[78] to 65 years.[84] All three studies focused on the real-time diagnosis of colorectal polyps sized <10mm, and provided sub-group analyses of diminutive polyps. All the studies were conducted in single centres. In the Kang and colleagues[78] study, four endoscopists carried out the colonoscopies, while the other two studies each involved one endoscopist. Kang and colleagues[78] reported that the study endoscopists had no prior experience with FICE, while Longcroft-Wheaton and colleagues (2012)[83] and Longcroft-Wheaton and colleagues (2011)[84] reported that the endoscopist in each of these studies had previous experience of in vivo diagnosis of polyps, although the authors did not specify endoscopists' experience with FICE. Longcroft-Wheaton and colleagues (2012)[83] stated that the study endoscopist had had prior training in real-time diagnosis. In the other studies,[78,84] the endoscopists' prior training in both real-time diagnosis and, more specifically, the use of FICE was unclear. Kang and colleagues[78] noted however, that the endoscopists received feedback every two weeks during the study about the accuracy of their endoscopic predictions compared to the histopathological diagnosis. The study by Kang and colleagues[78] (which also included an NBI arm) used a classification system for polyp characterisation based on colour, vascular density and vascular pattern.[71,94,95,98] The two studies by Longcroft-Wheaton and colleagues[83,84] both used a characterisation system based on vascular patterns which was developed by Teixeira and colleagues.[99]

Table 10 shows the quality assessments of the three FICE studies.[78,83,84] Reviewers considered all three studies to be at a low risk of bias across most of the QUADAS items assessed. None of the studies, however, reported the number of uninterpretable test results, but reviewers believed this to be zero in two studies.[78,84] Two studies explained participant withdrawals.[78,83] Longcoft-Wheaton and colleagues (2011)[84] did not state whether there were any withdrawals.

In addition to the assessment of the QUADAS items, the generalisability of each study was also briefly summarised during data extraction (the summary of reviewer's comments can be seen in full in the data extraction forms in Appendix 3). Reviewers noted that two of the studies were conducted in the UK,[83,84]

and so are likely to be representative of a UK population (although it is noted that these studies included small numbers of participants – 50 and 89 participants each). It was also noted that it is unclear how representative participants in the South Korea study[78] would be of the UK population and how similar the endoscopists' training in this study would be to endoscopists' training in the NHS in the UK. As all the studies were conducted in single centres it is unclear how the results would generalise to other centres and settings.

Table 9 Overview of the FICE studies

Study	Country	Centres	Patient population				Patient characteristics		Endoscopists			Classification system for polyp characterisation
			n	SCR	SURV	SYM	age (mean)	sex (% M/F)	n	FICE experience	Training	
Kang et al.[a78]	South Korea	1	196[b]	100%	n/a	n/a	54.3 (9.0)	76:24	4	No	Unclear[c]	Based on colour, vascular density & vascular pattern. Cites four references[71,94,95,98]
Longcroft-Wheaton et al. 2012[83]	UK	1	50[b]	100%	n/a	n/a	64 (4.2)[d]	46:54[e]	1	Unclear[f]	Unclear[f]	Based on vascular patterns using a system developed by Teixeira et al.[99]
Longcroft-Wheaton et al. 2011[84]	UK	1	89[b]	100%	n/a	n/a	65 (6.7)[g]	79:21[g]	1	Unclear[f]	Unclear[f]	System developed & validated by Teixeira et al.[99]

[a] Kang and colleagues also included an NBI arm which is reported in the earlier section on NBI and in Table 5

[b] Number is for the whole study (not just those patients with diminutive polyps). .

[c] States that the endoscopists performed a pilot study of a minimum of 50 examinations but it is not clear whether this was a minimum of 50 examinations each and whether the purpose of this study was to train the endoscopists.

[d] It is not clear whether this is the mean age for the 50 participants in this group with polyps or the total of 85 participants assigned to this group.

[e] This is the proportion of M:F for the total of 85 participants in the group. The proportion of M:F amongst the 50 participants with polyps is not reported.

[f] The endoscopist is described as trained and experienced in in vivo diagnostic methods but no further details are reported. It is not clear if FICE is the in vivo diagnostic method the endoscopist is trained and experienced in.

[g] For the total group of 89 participants (not just those with diminutive polyps)

Table 10 Overview of QUADAS assessments for the FICE studies

QUADAS ITEM		Kang 2015[a]	Longcroft-Wheaton 2012	Longcroft-Wheaton 2011
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Yes	Yes	Yes
2	Is the reference standard likely to classify the target condition correctly?	Yes	Yes	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes	Yes	Yes
4	Did the whole sample or a random selection of the sample, receive verification using the intended reference standard?	Yes	Yes	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	Yes	Yes
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Yes	Yes
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Yes	Yes	Yes
10	Were uninterpretable/ intermediate test results reported?	No	No	No
11	Were withdrawals from the study explained?	Yes	Yes	No

[a] Note that this is duplicate information because Kang 2015[78] also contained an NBI arm and thus is also represented in the QUADAS table for NBI studies (Table 6).

4.1.2 Assessment of diagnostic accuracy (sensitivity, specificity, NPV, accuracy)

NBI

Sensitivity and specificity of NBI for the characterisation of diminutive colorectal polyps

All but one of the included NBI studies reported sensitivity[72] or both sensitivity and specificity[20,54-63,65-68,70,71,73,74,76-78] of NBI for the characterisation of diminutive colorectal polyps as adenomas or hyperplastic polyps as compared to the characterisation verified by histopathological assessment of the resected polyps. Only Vu and colleagues[75] did not report on either sensitivity or specificity (this study was included in the systematic review because it reported accuracy in terms of the proportion of correctly classified polyps and data on surveillance intervals). The way in which data were reported by the studies varied and is shown in Table 11. Some studies reported on all the polyp characterisations made by study endoscopists. In other studies, the endoscopist indicated how confident they were in their NBI characterisation of the polyp as adenomatous or hyperplastic and results were reported separately for high and low confidence characterisations. Some studies reported data on all the characterisations and also the subsets of data for high and low confidence characterisations (data on low confidence characterisations is available in the data extraction forms in Appendix 3). One study, by Hewett and colleagues 2012a[58] restricted their study to the rectosigmoid colon. As can be seen in Table 11 several other studies also reported data for sub-sections of the colon as well as for the whole colon. One study, Iwatate and colleagues[61] reported a sub-group analysis by type of endoscopist (specialist or generalist).

The un-numbered sub-sections that follow Table 11 report on the:

- sensitivity and specificity of NBI for the characterisation of diminutive polyps in the whole colon (firstly data on all characterisations and then the separate subset of data on the polyp characterisations made with high confidence by the endoscopists) with accompanying meta-analyses (including a post-hoc analysis of high confidence characterisations made by endoscopists with prior experience of NBI).
- sensitivity and specificity of NBI for the characterisation of diminutive polyps in the rectosigmoid colon (again for all characterisations and separately for the subset of high confidence characterisations) with accompanying meta-analyses (including a post-hoc analysis of high confidence characterisations made by endoscopists with prior experience of NBI).
- sensitivity and specificity of NBI for the characterisation of polyps in parts of the colon other than the rectosigmoid colon (too few studies to meta-analyse).
- NPV of NBI for the characterisation of diminutive colorectal polyps; accuracy of NBI (proportion of correctly classified polyps).

Table 11 Overview of the available data on sensitivity and specificity

	Reported data on all characterisations of polyps	Reported data on characterisations made with high confidence
Whole colon	Aihara et al.[54] Chandran et al.[55] Gupta et al.[56] Henry et al.[57] Ignjatovic et al.[59] Ikematsu et al.[60] Iwatate et al.[61] Kang et al.[78] Ladabaum et al.[63] Lee et al.[77] Patel et al.[a67] Repici et al.[70] Rex et al.[71] Rogart et al.[b69] Shahid et al.[73] Sola-Vera et al.[74] Wallace et al.[76]	Hewett et al. 2012b[b,c 20] Iwatate et al.[61] Kaltenbach et al.[c 62] Ladabaum et al.[b 63] Lee et al.[77] Paggi et al. 2012[c 66] Paggi et al. 2015[c 65] Patel et al.[a 67] Pohl et al.[68] Repici et al.[70] Rex et al.[71] Sola-Vera et al.[74] Wallace et al.[76]
Whole colon by colonosopist type	Iwatate et al.[61] (Specialist and generalist colonoscopists)	
Right colon		Kaltenbach et al.[62]
Proximal to splenic flexure		Pohl et al.[68]
Left colon	Gupta et al.[56]	Kaltenbach et al.[62]
Distal colon		Pohl et al.[68]
Rectosigmoid colon	Hewett et al. 2012a [58] Ladabaum et al.[63] Patel et al.[b 67] Wallace et al.[76]	Hewett et al. 2012a [58] Patel et al.[b 67] Pohl et al.[68] Repici et al.[70] Wallace et al.[76]

Proximal to rectosigmoid colon	Ladabaum et al.[63] Patel et al.[b 68]	Patel et al.[b 67]
Rectum		Kaltenbach et al.[62]

[a] Data to populate a 2x2 table were not reported and it proved difficult to impute data that would provide outcomes to match all the outcomes (accuracy, sensitivity, specificity, PPV & NPV) reported in the paper. Data imputed should be regarded as illustrative.

[b] Published papers reported values for sensitivity &/or specificity but data to populate a 2x2 table and recalculate these values were not reported or were reported incompletely and it was not possible to impute the missing data.

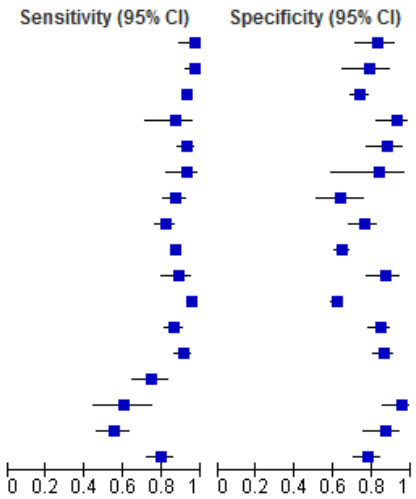
[c] Only reported outcomes for high confidence characterisations

Sensitivity and specificity of NBI for the characterisation of diminutive colorectal polyps in the whole colon

Twenty-two studies[20,54-57,59-63,65-68,70-74,76-78] reported on the characterisation of diminutive polyps within the whole colon although five of these only reported data from high confidence characterisations.[20,62,65,66,68]

The results for all characterisations of diminutive polyps in the whole colon (i.e. not separated by confidence level), where 2x2 table data were reported or calculable, are shown in Figure 22.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aihara 2016	60	10	2	49	0.97 [0.89, 1.00]	0.83 [0.71, 0.92]
Chandran 2015	105	11	3	40	0.97 [0.92, 0.99]	0.78 [0.65, 0.89]
Gupta 2012	484	97	37	266	0.93 [0.90, 0.95]	0.73 [0.68, 0.78]
Henry 2010	32	4	5	49	0.86 [0.71, 0.95]	0.92 [0.82, 0.98]
Ignjatovic 2009	144	7	11	51	0.93 [0.88, 0.96]	0.88 [0.77, 0.95]
Ikematsu 2015	50	3	4	15	0.93 [0.82, 0.98]	0.83 [0.59, 0.96]
Iwatate 2015	123	25	18	44	0.87 [0.81, 0.92]	0.64 [0.51, 0.75]
Kang 2015	190	37	42	115	0.82 [0.76, 0.87]	0.76 [0.68, 0.82]
Ladabaum 2013	995	252	155	456	0.87 [0.84, 0.88]	0.64 [0.61, 0.68]
Lee 2011	71	10	9	66	0.89 [0.80, 0.95]	0.87 [0.77, 0.94]
Patel 2016 [a]	1523	490	77	786	0.95 [0.94, 0.96]	0.62 [0.59, 0.64]
Repici 2013	203	31	32	163	0.86 [0.81, 0.90]	0.84 [0.78, 0.89]
Rex 2009	178	28	17	172	0.91 [0.86, 0.95]	0.86 [0.80, 0.90]
Rogart 2008 [b]	71		24		0.75 [0.65, 0.83]	
Shahid 2011	27	3	18	55	0.60 [0.44, 0.74]	0.95 [0.86, 0.99]
Sola-Vera 2015	85	8	70	53	0.55 [0.47, 0.63]	0.87 [0.76, 0.94]
Wallace 2014	120	35	31	124	0.79 [0.72, 0.86]	0.78 [0.71, 0.84]



[a] The data for Patel have been imputed by the reviewer but it was not possible to find a solution that agreed with all the 2x2 table outcomes reported in the paper. These imputed values (which should be regarded as illustrative) produce the reported sensitivity and specificity, but produce values for PPV and NPV that are lower than reported and an accuracy value (proportion of correctly classified polyps among all the polyps) that is higher.

[b] Rogart and colleagues did not report a value for specificity and it was not possible to complete the 2x2 table from the information reported in the published paper.

Figure 5 Accuracy of NBI for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps

The ability of NBI to correctly identify diminutive polyps as adenomas (i.e. the sensitivity of the test) ranged from 0.55 to 0.97 (i.e. 55% to 97%) across the 17 studies that reported this outcome. Sensitivity was above 90% in seven studies[54-56,59,60,67,71] (and in two of these it was 95% or higher[55,67]), lay between 80% and 90% in six other studies[57,61,63,70,77,78] and was below 80% in four studies.[72-74,76]

The ability of NBI to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) was typically lower than the sensitivity of the test, ranging from 0.62 to 0.95 (i.e. 62% to 95%) across the 16 studies that reported this outcome. Specificity was above 90% in just two studies,[57,73] lay between 80% and 90% in seven studies[54,59,60,70,71,74,77] and was below 80% in seven studies.[55,56,61,63,67,76,78]

It was possible to run a bivariate meta-analysis (using Stata/IC14 and metaandi[44]) for the 16 studies that reported both sensitivity and specificity. This produced a summary value for sensitivity of 0.88 (95% CI 0.83 to 0.92) and for specificity of 0.81 (95% CI 0.75 to 0.85). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 6 in which the individual study estimate points are scaled to the sample size of the study (i.e. larger circles represent

larger studies). The 95% confidence region around the summary point indicates where we have 95% confidence that the summary point lies. The 95% prediction region illustrates the extent of statistical heterogeneity among the studies. If the bivariate model for sensitivity and specificity is correct, we have 95% confidence that the true sensitivity and specificity of a new study in the future will lie within the 95% prediction region. As can be observed from Figure 6 the 95% prediction region is large.

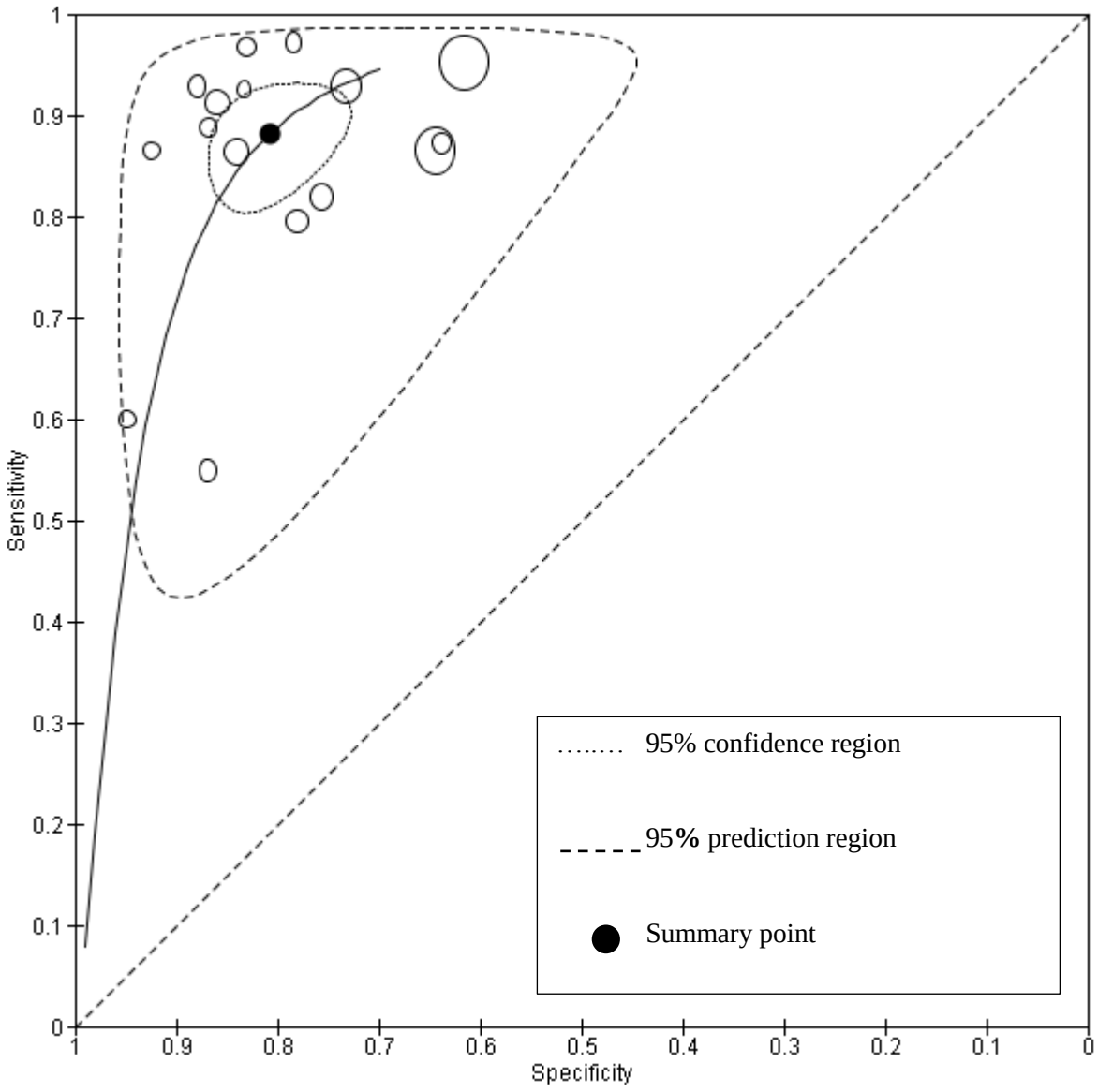


Figure 6 SROC plot from the meta-analysis of NBI for all characterisations of polyps in the whole colon.

In order to investigate the heterogeneity between studies, a covariate for endoscopist experience with NBI was added to RevMan and separate SROC curves were drawn as shown in Figure 7. Whilst caution must be taken when interpreting this figure due to the small number of studies for each subgroup, it nevertheless appears to support the hypothesis that endoscopists with prior experience of using NBI to characterise diminutive colorectal polyps achieve higher sensitivity and specificity than endoscopists who have had no prior experience of using NBI to characterise diminutive colorectal polyps (other than any training that they undertook at the start of the study).

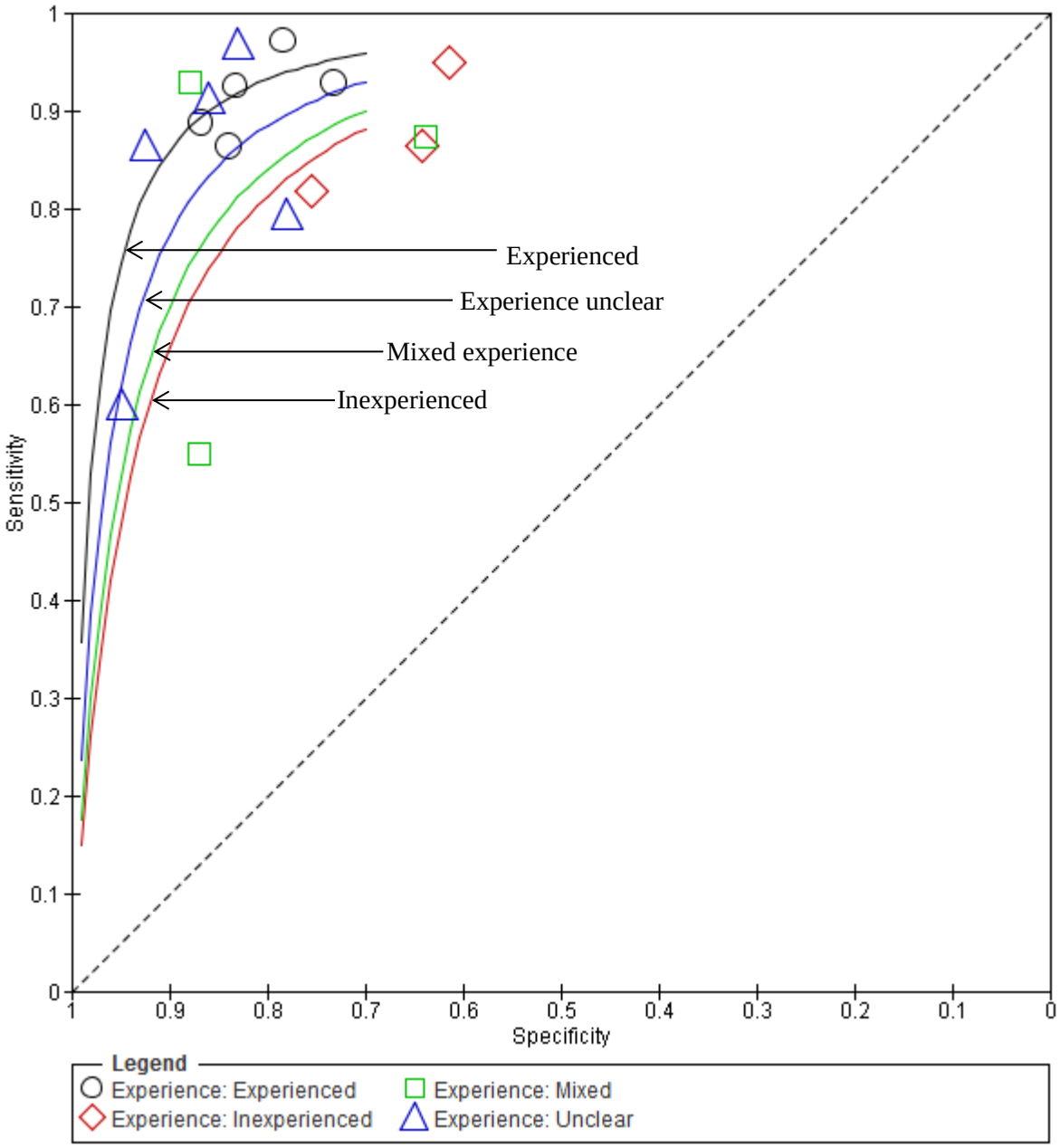
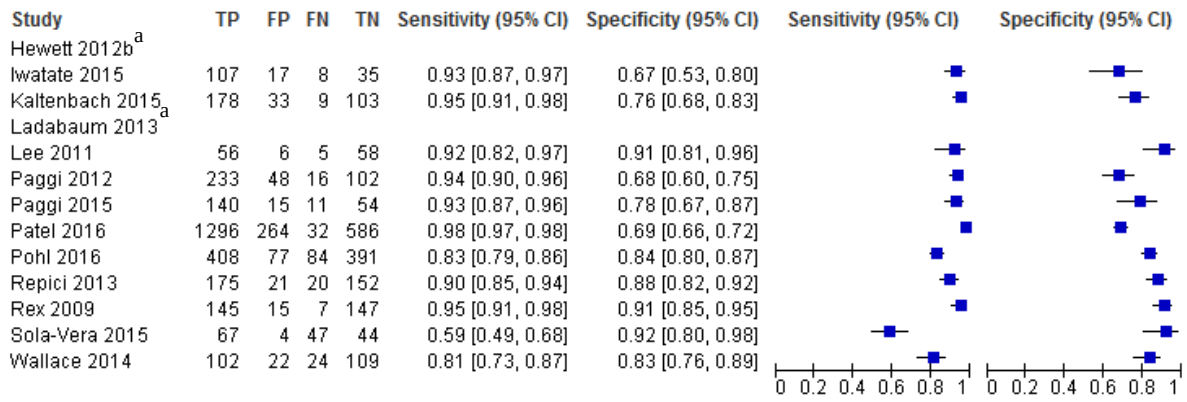


Figure 7 SROC plots for all characterisations of polyps in the whole colon by endoscopists level of experience using NBI

Results for studies that reported results from polyp characterisations using NBI that were designated as high confidence decisions, and where 2x2 table data were reported or calculable, are shown in Figure 8.



[a] It was not possible for us to impute the 2x2 table data necessary to plot these results within this figure. Hewett and colleagues' study 2012b[20] reported a value for sensitivity of 98% (no confidence interval provided and specificity not reported) and Ladabaum and colleagues'[63] reported sensitivity of 88.4% (95% CI 82.2 to 94.7) and specificity of 44.1% (26.5 to 61.6).

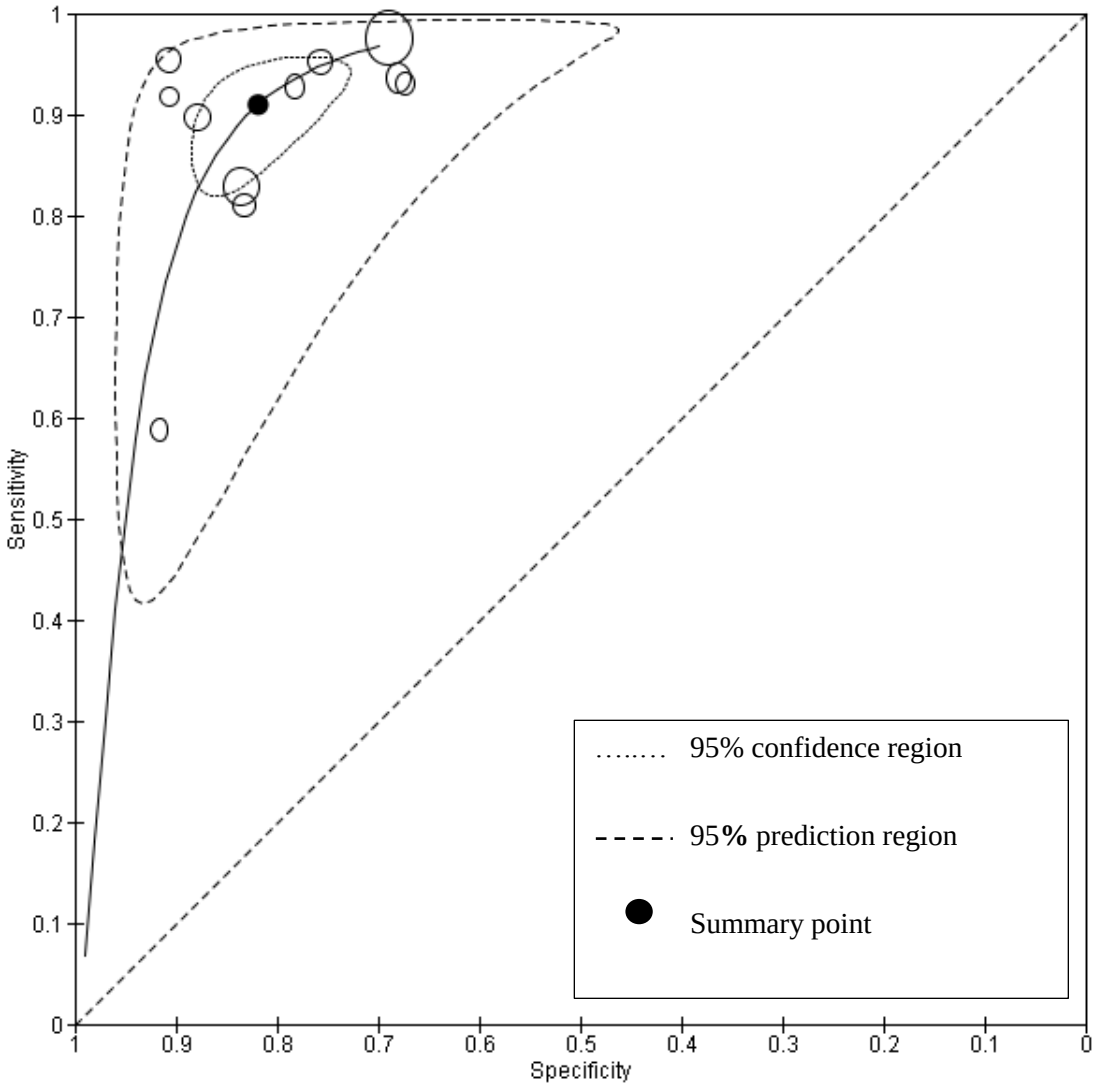
Figure 8 Accuracy of NBI high confidence decisions for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps in the whole colon

The ability of high confidence characterisations made with NBI to correctly identify diminutive polyps as adenomas (ie. the sensitivity of the test) was 0.90 or more (i.e. 90% or more) in nine of the 13 studies[20,61,62,65-67,70,71,77] (in four of these it was 95% or higher[20,62,67,71]) and it lay between 80% and 90% in three other studies.[63,68,76] The lowest sensitivity value reported was 59% by Sola-Vera and colleagues.[74] Some studies reported the sensitivity obtained from all characterisations and the sensitivity from only the high confidence characterisations. In each study where both these values were reported, the sensitivity was higher when obtained from high confidence decisions (difference ranging from an increase of 1.5% to 5.8%).

The ability of NBI to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) from high confidence polyp characterisations was just above 90% (i.e. above 0.90) in three studies[71,74,77] but did not exceed 92% in any study. In just three studies specificity lay between 80% and 90%[68,70,76] but in the majority of the studies it lay below 80%[61-63,65-67] with the lowest specificity just 44.1% reported by Ladabaum and colleagues.[63] Specificity was higher when obtained from high confidence decisions in seven of the eight studies that reported both the specificity obtained from all characterisations and the specificity from only the high confidence characterisations, with the increase ranging from 3.5% to 7.3%. The one exception was the study by Ladabaum and colleagues[63] where the specificity calculated

from high confidence characterisations was lower than that obtained from all characterisations (44.1% versus 64.4% respectively).

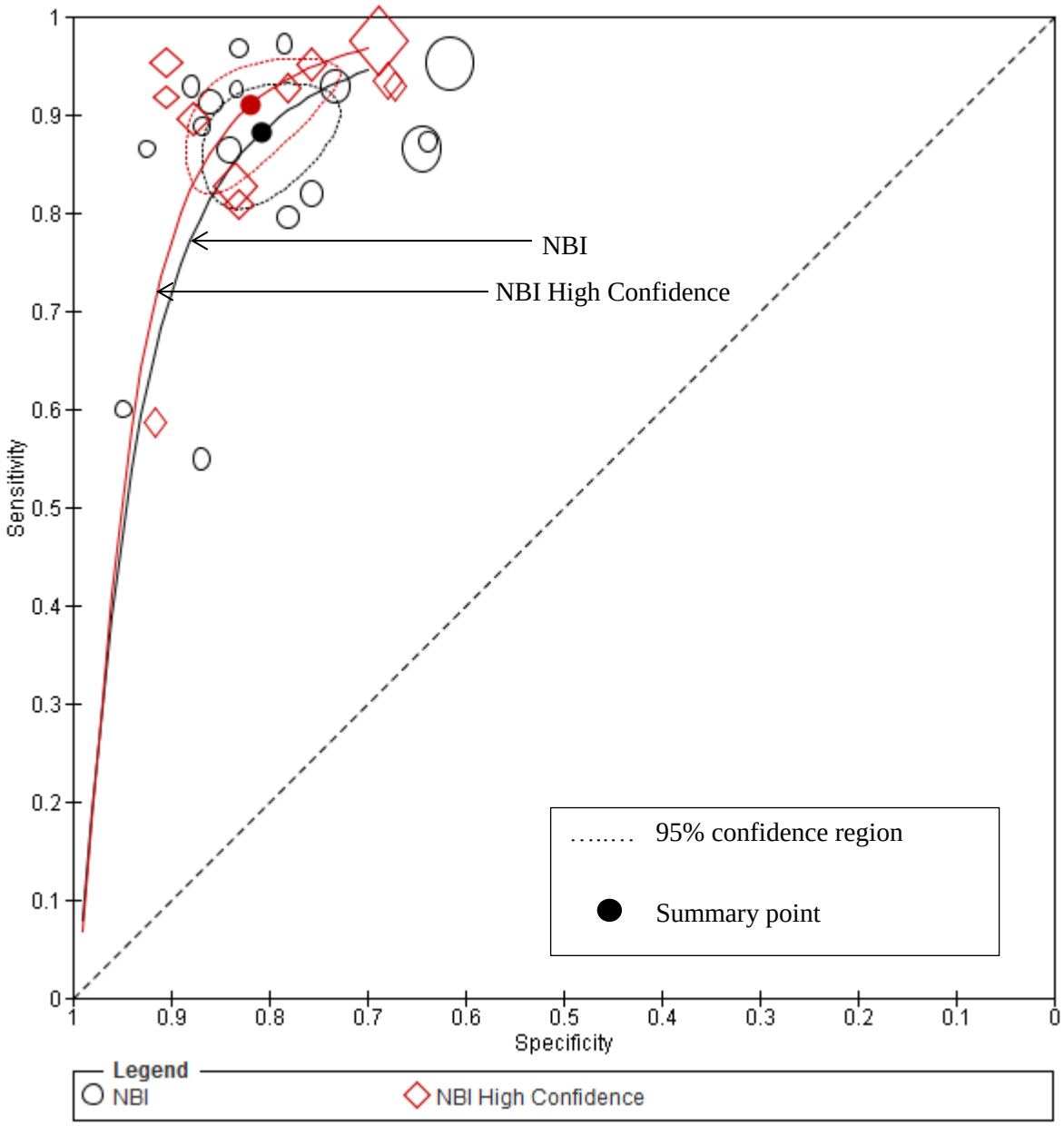
A bivariate meta-analysis (using Stata/IC14 and metaandi[44]) was run for the 11 studies that reported both sensitivity and specificity from polyp characterisations made with high confidence. This produced a summary value for sensitivity of 0.91 (95% CI 0.85 to 0.95) and for specificity of 0.82 (95% CI 0.76 to 0.87). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 9 in which the individual study estimate points are scaled to the sample size of the study (i.e. larger circles represent larger studies). The effect of reporting only on high confidence characterisations in comparison to all polyp characterisations is to move the summary estimate up (increasing sensitivity) and slightly to the left (increasing specificity).



Note that two studies were not included in the meta-analysis: Hewett and colleagues' study 2012b[20] sensitivity 98%; Ladabaum and colleagues'[63] sensitivity 88.4% (95% CI 82.2 to 94.7), specificity 44.1% (26.5 to 61.6).

Figure 9 SROC plot showing the summary point on the summary curve from the meta-analysis of NBI for high confidence characterisations of polyps in the whole colon

The impact of restricting the analysis to high confidence characterisations in comparison to including all characterisations can be observed in Figure 10 in which shows both summary curves on the same plot. As already stated the effect of reporting only on high confidence characterisations in comparison to all polyp characterisations is that the summary estimate moves up (increasing sensitivity) and slightly to the left (increasing specificity).



Note: for clarity the 95% prediction regions are not shown on this plot

Figure 10 SROC for all NBI characterisations of polyps in the whole colon and SROC for only high confidence NBI characterisations of polyps in the whole colon shown on the same plot

Seven studies[61,63,67,70,71,74,76,77] reported both sensitivity and specificity from all diminutive polyp characterisations and separately for only high confidence diminutive polyp characterisations, although for one of the these studies[63] 2x2 table data were not available for the high confidence characterisations

[which had a reported sensitivity of 88.4% (95% CI 82.2 to 94.7) and specificity of 44.1% (26.5 to 61.6)]. The pairs of results from these studies are shown in Figure 11 and forest plots in Figure 12.

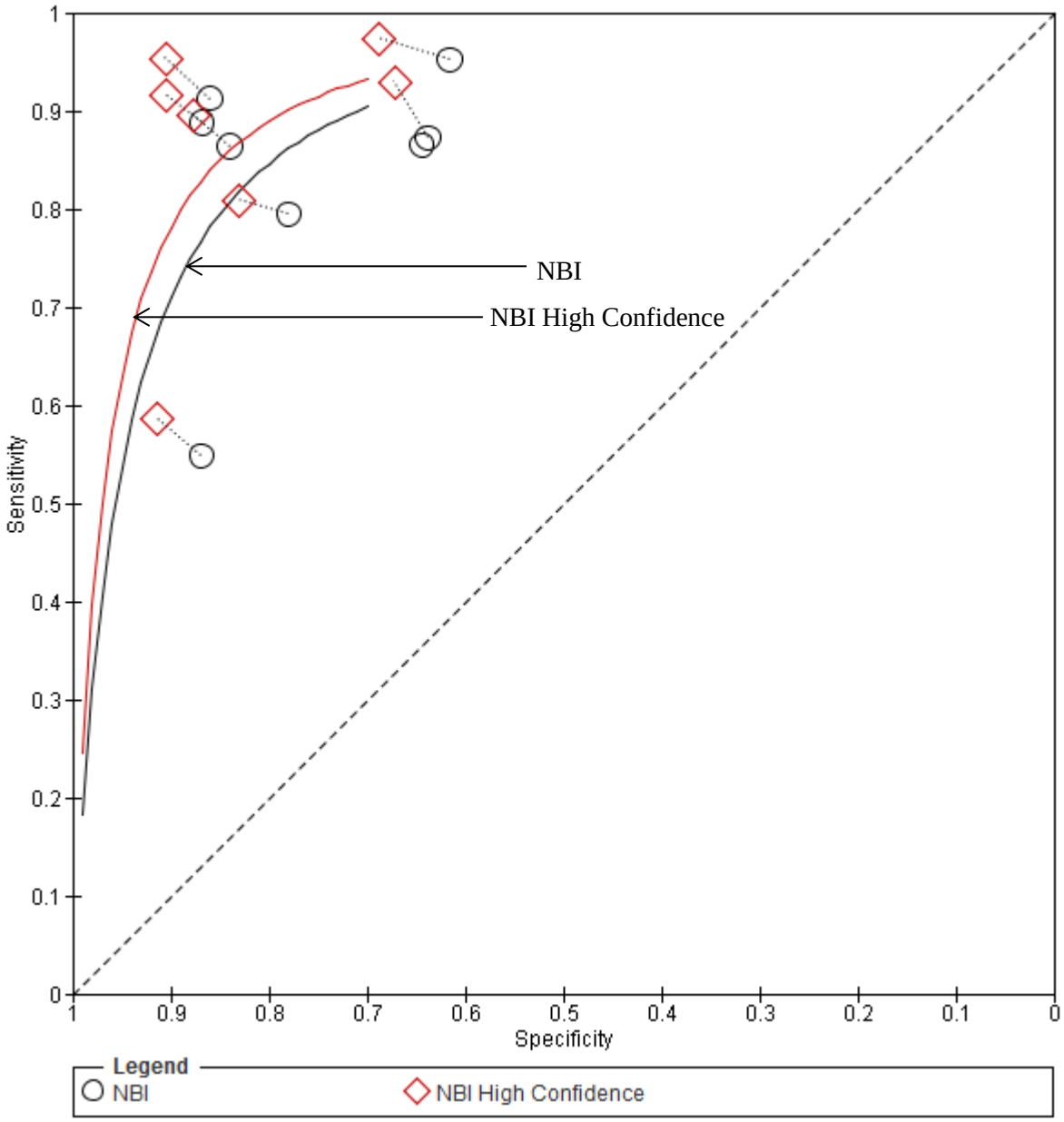
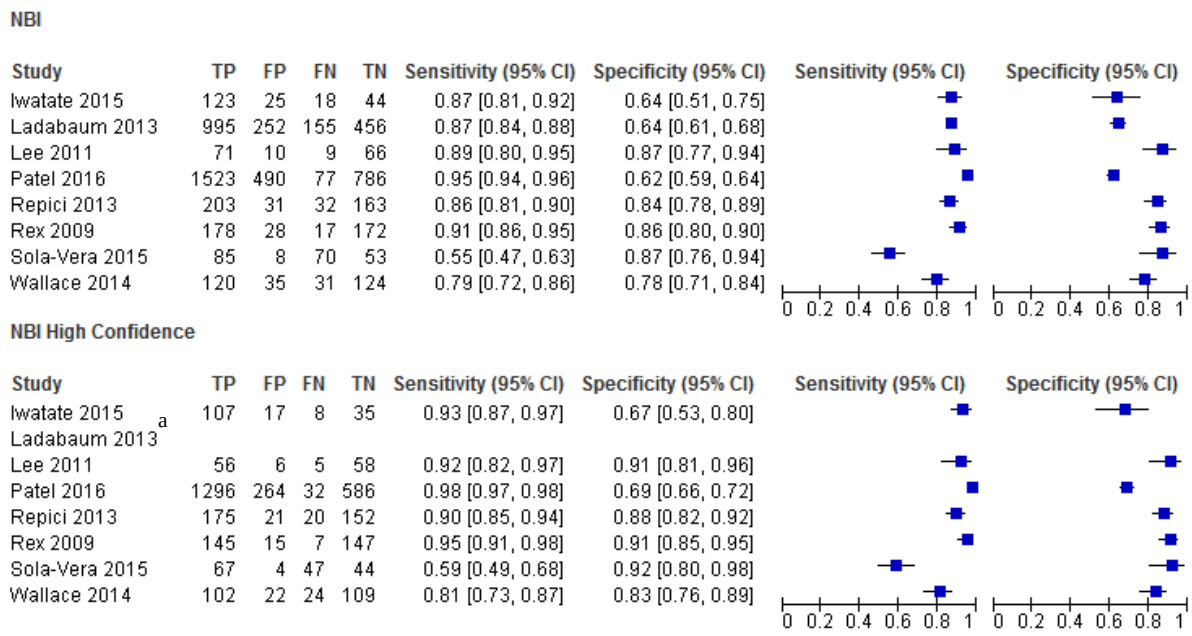


Figure 11 Plot showing paired data from the studies that reported on all diminutive polyp characterisations and separately on high confidence diminutive polyp characterisations



[a] It was not possible for us to impute the 2x2 table data necessary to plot these results within this figure [reported sensitivity of 88.4% (95% CI 82.2 to 94.7) and specificity of 44.1% (26.5 to 61.6)]

Figure 12 Accuracy of NBI in studies that reported on all diminutive polyp characterisations and separately on high confidence diminutive polyp characterisations

To obtain data for a scenario analysis within the economic model (section 5.5.2.2) a post-hoc bivariate meta-analysis (using Stata/IC14 and metaandi[44]) was run for a subgroup in which endoscopists experienced in the use of NBI characterised the polyps in the whole colon (Figure 13). There were four such studies included in this analysis.[65,66,70,77]

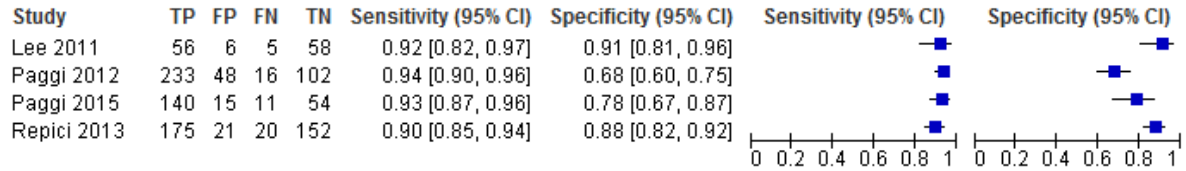


Figure 13 Accuracy of NBI high confidence decisions for characterising diminutive colorectal polyps in the whole colon as either adenomas or hyperplastic polyps when made by endoscopists experienced in the use of NBI

The meta-analysis produced a summary value for sensitivity of 0.92 (95% CI 0.89 to 0.94) and for specificity of 0.82 (95% CI 0.72 to 0.89). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 14 in which the individual study estimate points are scaled to the sample size of the study (i.e. larger circles represent larger studies). Restricting the meta-analysis from 11 studies reporting different levels of NBI experience (Experienced n=4; Mixed experience n=3; Inexperienced n=2; Unclear n=2) to the four studies that reported endoscopists experienced in the use of NBI narrowed the 95% CI for sensitivity [11 studies variety of experience: 0.91 (95% CI 0.85 to 0.95); four studies with prior NBI experience: 0.91 (95% CI 0.89 to 0.94)] and widened the 95% CI for specificity [11 studies variety of experience: 0.82 (95% CI 0.76 to 0.87); four studies with prior NBI experience: 0.82 (95% CI 0.72 to 0.89). The changes in the 95% confidence intervals are reflected in the change in the size and shape of the 95% confidence region and 95% prediction region in Figure 14 in comparison to Figure 9.

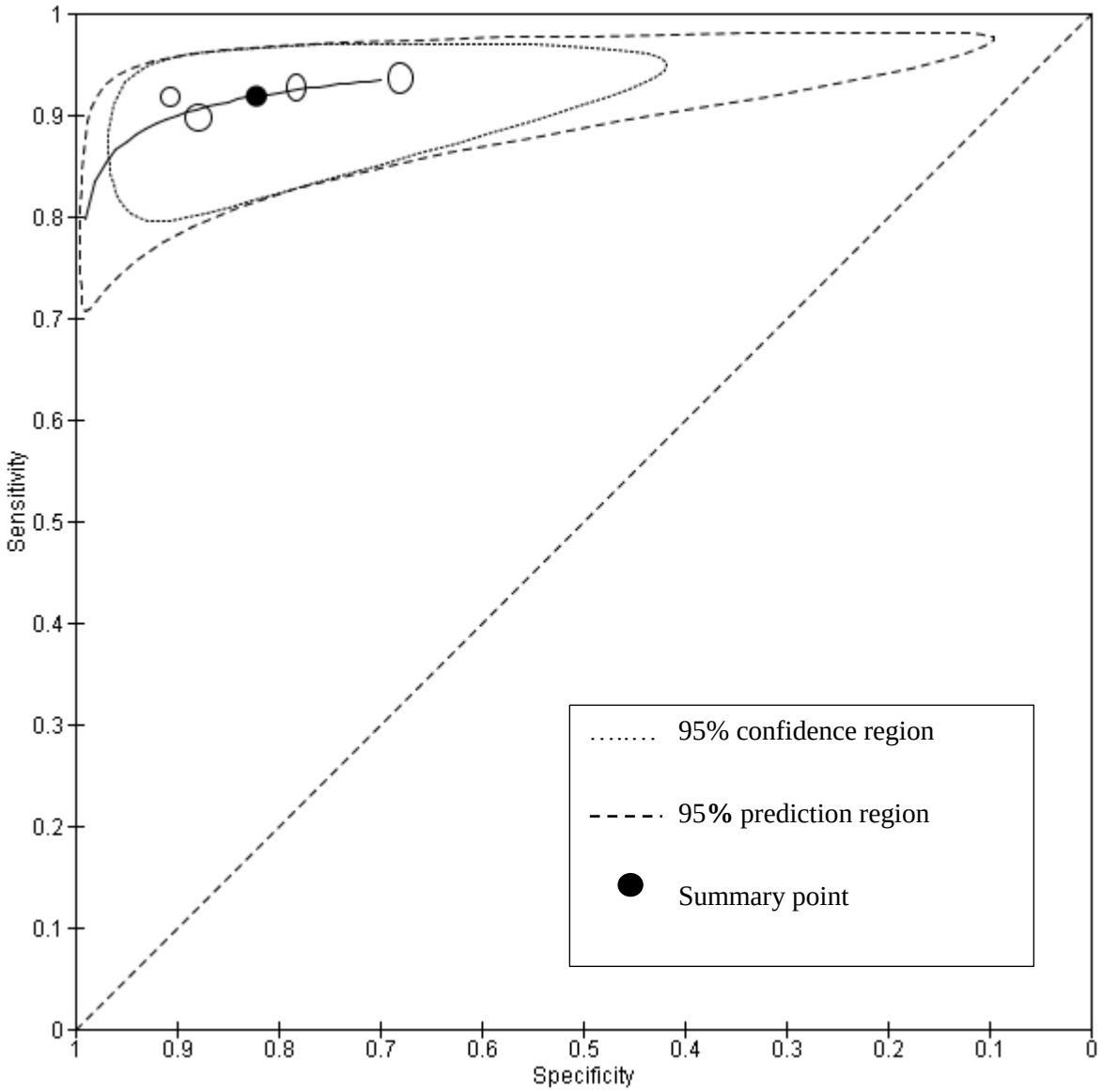


Figure 14 SROC plot showing the summary point on the summary curve from the meta-analysis of NBI for high confidence characterisations of polyps in the whole colon when made by endoscopists experienced in the use of NBI

Colonoscopies in one study, by Iwatate and colleagues[61] were conducted by five endoscopists. Two of the five endoscopists were described as specialists in colonoscopy and they had extensive experience in magnifying colonoscopy with NBI (>1000 cases). The other three endoscopists were described as general endoscopists with limited experience in magnifying colonoscopy with NBI (≤1000 cases). As shown in Table 12 the two specialist endoscopists achieved higher sensitivity and specificity than the three general endoscopists but the difference between the two was only statistically significant for specificity (p=0.007).

Table 12 Sensitivity and specificity according to experience with NBI of the endoscopists

	High confidence characterisations of polyps 1-5mm	
	Specialist endoscopists	General endoscopists
- Sensitivity 95% CI	93.5% 78.58% to 99.21% [a]	92.9% 85.10% to 97.33% [a]
- Specificity 95% CI	87.0%[1] 66.41% to 97.22% [a]	51.7%[1] 32.53% to 70.55% [a]

* calculated by reviewer

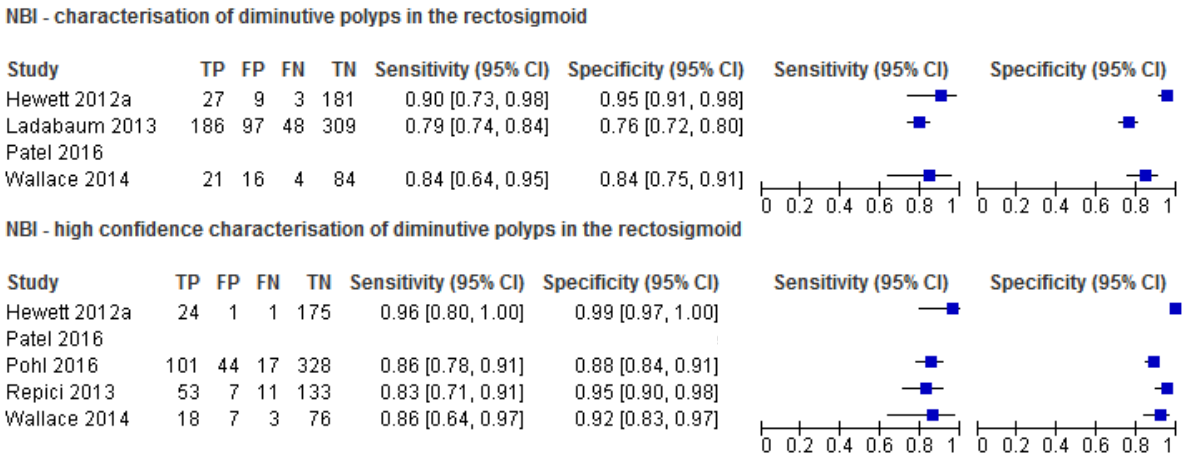
[1] The differences between the specificity rates for the SC and the GE group were significant p=0.007.

Sensitivity and specificity of NBI for the characterisation of diminutive colorectal polyps in the rectosigmoid colon.

As shown in Table 11 four studies[58,63,67,76] reported sensitivity and specificity following characterisation (any level of confidence) of diminutive polyps in the rectosigmoid colon with three of these reporting sufficient data for a 2x2 table to be constructed for entry into meta-analysis.[58,63,76]

Three of the four studies[58,67,76] that reported results for all characterisations also reported sensitivity and specificity following high confidence characterisations of polyps in the rectosigmoid colon with two further studies[68,70] only reporting high confidence characterisation data. Four of the five studies reporting on high confidence characterisations provided sufficient data for 2x2 tables to be constructed for entry into meta-analysis.[58,68,70,76]

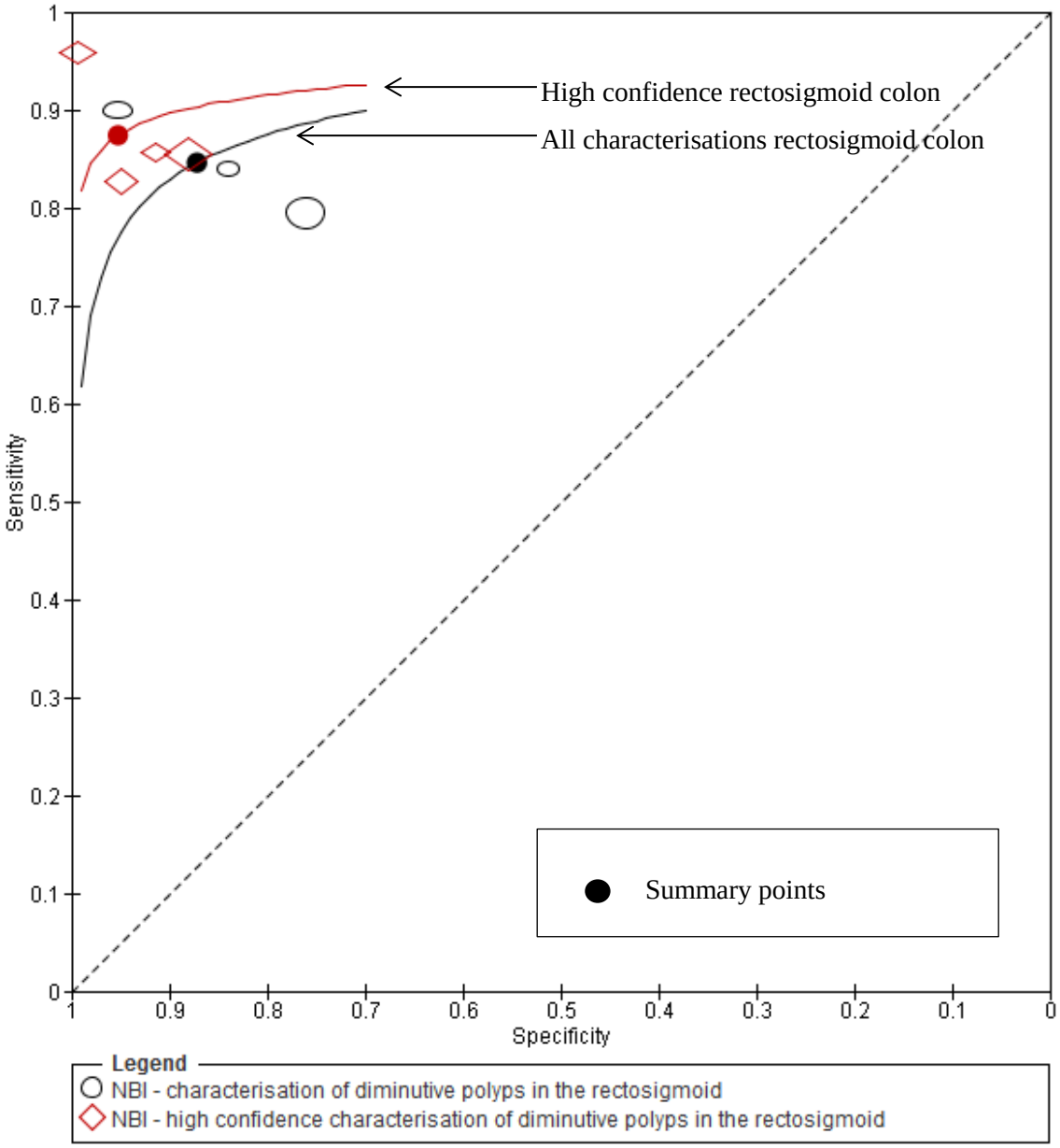
Results from the studies that used NBI to characterise polyps in the rectosigmoid colon, where 2x2 table data were reported or calculable, are shown in Figure 15. The results from Patel and colleagues[67] are not represented in Figure 15 because it was not possible to impute values into a 2x2 table that provided a solution for the reported outcomes in the paper (accuracy, sensitivity, specificity, PPV and NPV).



[a] It was not possible for us to impute the 2x2 table data necessary to plot these results within this figure. For characterisation of all diminutive polyps in the rectosigmoid colon Patel and colleagues[67] reported sensitivity of 88.4% (95% CI 84.8% to 92.0%) and specificity of 78.3% (95% CI 71.8% to 84.9%). The high confidence polyp characterisations yielded sensitivity of 90.9% (95% CI 87.4% to 94.4%) and specificity of 88.6% (95% CI 81.0% to 96.1%).

Figure 15 Accuracy of NBI for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps in the rectosigmoid colon

Bivariate meta-analyses was conducted (using Stata/IC14 and xtmelogit or using Stata/IC14 and metandi[44]) of the studies where 2x2 table data were available. For all characterisations of diminutive polyps in the rectosigmoid colon the summary value for sensitivity is 0.85 (95% CI 0.75 to 0.91) and for specificity is 0.87 (95% CI 0.74 to 0.94). For high confidence characterisations of diminutive polyps in the rectosigmoid colon the summary value for sensitivity is 0.87 (95% CI 0.80, 0.92) and for specificity is 0.95 (95% CI 0.87, 0.98). The parameter estimates for the bivariate model from these two meta-analyses were entered into RevMan to produce the SROC plot shown below in Figure 16 (individual study estimate points are scaled to the sample size of the study). As seen with the results for the whole colon, the effect of reporting only high confidence polyp characterisations in comparison to all polyp characterisations is to increase sensitivity and specificity (summary point moves up and to the left on the SROC plot).



Note that one study was not included in either meta-analysis: Patel and colleagues[67] all characterisations sensitivity 88.4% (95% CI 84.8% to 92.0%), specificity 78.3% (95% CI 71.8% to 84.9%); high confidence characterisations sensitivity 90.9% (95% CI 87.4% to 94.4%), specificity 88.6% (95% CI 81.0% to 96.1%). The large 95% confidence and a 95% prediction regions which were generated for the high confidence characterisation plot are not shown on this figure and the software used to draw the SROC plot (Review Manager 5.3) did not generate a 95% confidence region or a 95% prediction region for the other data set..

Figure 16 SROC plot showing the summary points on the summary curves from the meta-analyses of NBI for all characterisations of polyps and for only high confidence characterisations of polyps in the rectosigmoid colon

To obtain data for a scenario analysis within the economic model (section 5.5.2.2) a post-hoc bivariate meta-analysis (using Stata/IC14 and xtmelogit) was run for a sub-group of studies in which the endoscopists were experienced in the use of NBI. There were two such studies[58,70] included in the analysis (Figure 17).

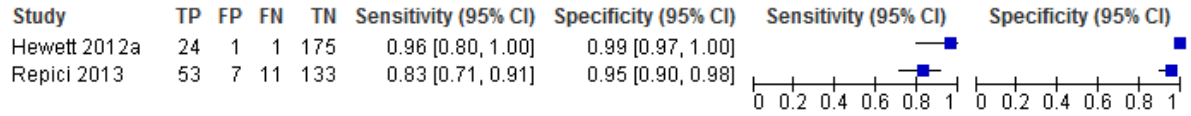
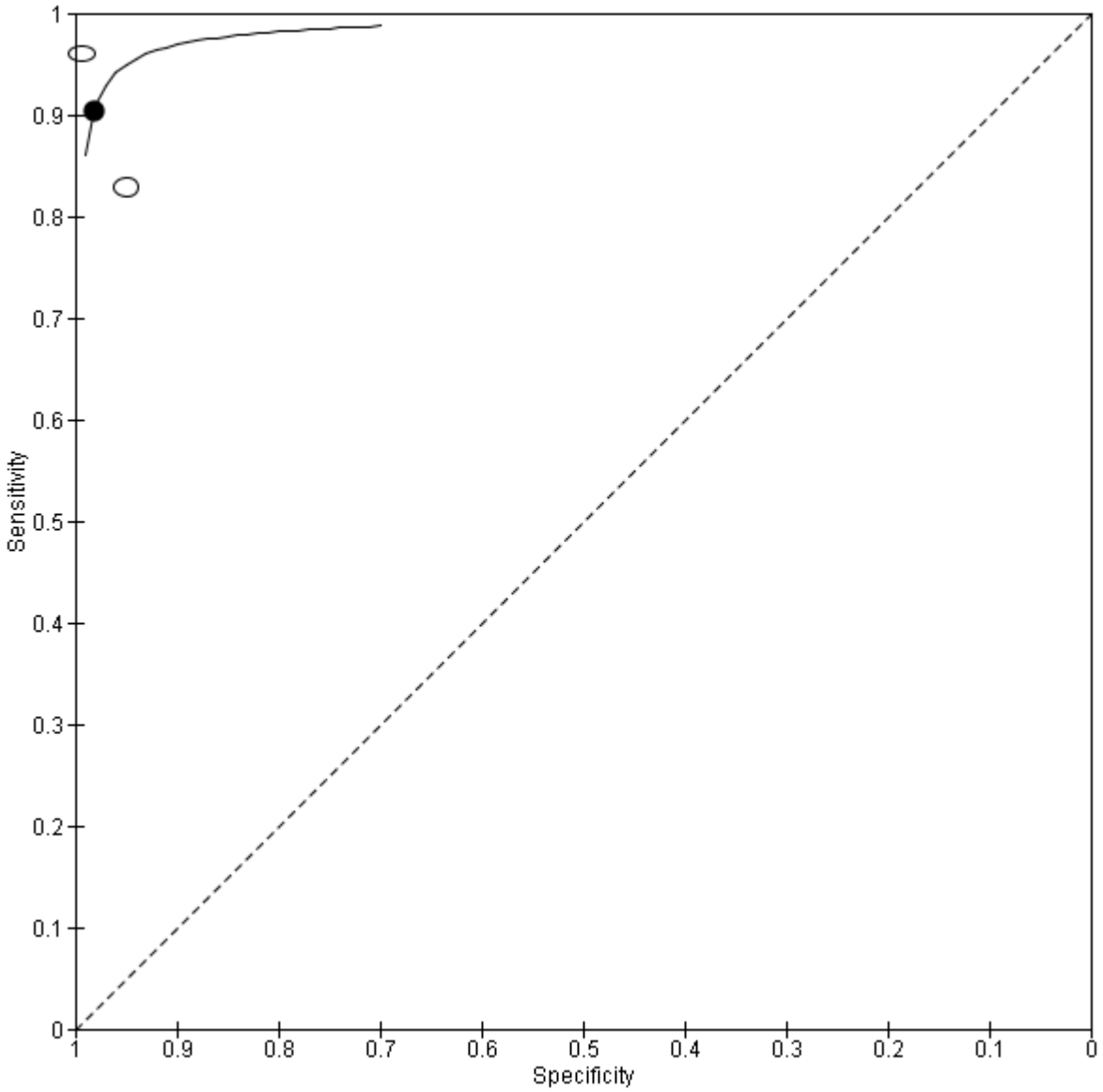


Figure 17 Accuracy of NBI high confidence decisions, made by endoscopists with prior experience of NBI, for characterising diminutive colorectal polyps in the rectosigmoid colon as either adenomas or hyperplastic polyps

The meta-analysis produced a summary value for sensitivity of 0.90 (95% CI 0.71 to 0.97) and for specificity of 0.98 (95% CI 0.91 to 1.00). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 18 in which the individual study estimate points are scaled to the sample size of the study (i.e. larger circles represent larger studies). Restricting the meta-analysis from the four studies reporting different levels of NBI experience (Experienced n=2; Inexperienced n=1; Unclear n=1) to only the two studies where endoscopists had experience in the use of NBI increased the summary value for sensitivity whilst widening the 95% CI [four studies variety of experience: 0.87 (95% CI 0.80, 0.92); two studies with prior NBI experience: 0.90 (95% CI 0.71 to 0.97)] and increased the summary value for specificity whilst narrowing the 95% CI [four studies variety of experience: 0.95 (95% CI 0.87, 0.98); two studies with prior NBI experience: 0.98 (95% CI 0.91 to 1.00).



Note that the software used to draw the SROC plot (Review Manager 5.3) did not generate a 95% confidence region or a 95% prediction region for this meta-analysis. It is presumed that this is because of the small number of studies.

Figure 18 SROC plot showing the summary point on the summary curve from the meta-analyses of NBI for high confidence characterisations of polyps in the rectosigmoid colon made by endoscopists with prior experience of NBI

Sensitivity and specificity of NBI for the characterisation of diminutive colorectal polyps in parts of the colon other than the rectosigmoid colon

Five studies[56,62,63,67,68] provided data on the characterisation of diminutive polyps in regions of the colon, other than the rectosigmoid colon (Table 11). The results reported by these studies are summarised in Table 13

Table 13 Summary of the sensitivity and specificity of NBI for the characterisation of diminutive colorectal polyps in parts of the colon other than the rectosigmoid colon

	Study	Sensitivity	95% CI	Specificity	95% CI
Right colon					
High confidence characterisations	Kaltenbach et al.[62]	96.4%	91.0% to 99.0%	61.4%	45.5% to 75.6%
Proximal to splenic flexure					
High confidence characterisations	Pohl et al.[68]	82%	77.8% to 86.4%	62%	49.8% to 73.7%
Left colon					
All characterisations of polyps	Gupta et al.[56]	91.4%	86.8% to 94.8%	78.1%	73.0% to 82.6%
High confidence characterisations	Kaltenbach et al.[62]	95.5%	87.5% to 99.1%	83.6%	71.2% to 92.2%
Distal colon					
High confidence characterisations	Pohl et al.[68]	84%	77.6% to 89.0%	87%	83.5% to 90.3%
Proximal to rectosigmoid colon					
All characterisations of polyps	Ladabaum et al.[63]	88.2	82.2% to 94.2%	49.7	34.7% to 64.6%
	Patel et al.[67]	91.0%	88.3% to 94.0%	36.9%	27.7% to 46.1%
High confidence characterisations	Patel et al.[67]	96.2%	94.1% to 98.4%	34.9%	22.1% to 47.7%
	Patel et al.[67]	73.7%	65.8% to 81.5%	44.4%	37.3% to 51.1%
Rectum					
High confidence characterisations	Kaltenbach et al.[62]	77.8%	40.0% to 97.2%	81.1%	64.8% to 92.0%

Negative predictive value of NBI for the characterisation of diminutive colorectal polyps

The negative predictive value is the probability that subjects with a negative screening test (i.e. colorectal polyp is characterised as hyperplastic) truly do not have an adenoma. However, it must be borne in mind when viewing these results that the negative predictive value is influenced by the prevalence of disease (i.e. in this case the prevalence of adenomas in the tested populations). When prevalence is increased the result is a decrease in the negative predictive value. Due to the importance of NPV within the PIVI statement (see section 1.3.1) consideration was given to meta-analysing NPVs from the included studies even though this is not advised by either the NICE Diagnostics Programme Manual[36] or the Cochrane Diagnostic Test Accuracy Handbook.[35] However, because it is clear that the prevalence of adenomas and hyperplastic polyps is likely to vary between studies [e.g. due to differences in case mix (screening, surveillance and symptomatic populations) and patient characteristics (age, sex)] we chose not to pool NPV values across studies. Instead we have provided forest plots for these outcomes and marked the 90% threshold value on each plot.

For the characterisations of diminutive polyps in the whole colon (made with any level of confidence) the NPV ranged from 43% to 96.1% (Figure 19 and Table 14). The study by Sola-Vera and colleagues[74] is noteworthy because this study reported the lowest NPV and it was far lower than in any other study. All the other studies reported NPV values over 70% with five studies reporting NPV values of 90% or more,[54,55,57,67,71] however it should be noted that the lower limit of the 95% confidence interval fell below 90% in every study except Patel and colleagues.[67]

Limiting the assessment of NPV to high confidence polyp characterisations increased the NPV which ranged from 48% to 98.3% in the studies that reported this outcome (Table 14). Again the study by Sola-Vera and colleagues had the lowest NPV of any study by a considerable margin. All the other studies reported NPV values for high confidence assessments of over 78% with five studies reporting NPV values of 90% or more.[20,62,67,71,77] Once again however, inspection of the 95% confidence intervals reveals that the lower limit of this fell below 90% in all but two studies.[67,71]

One study, by Iwatate and colleagues,[61] compared differences in NPV achieved by specialists in colonoscopy and general endoscopists. Specialists in colonoscopy achieved NPVs of over 90% (mean value 90.9%, 95% CI 70.8 to 98.9) whereas the NPVs achieved by general endoscopists were lower with a mean value of 71.4% (95% CI 47.8 to 88.8), however the difference between the groups was not statistically significant.

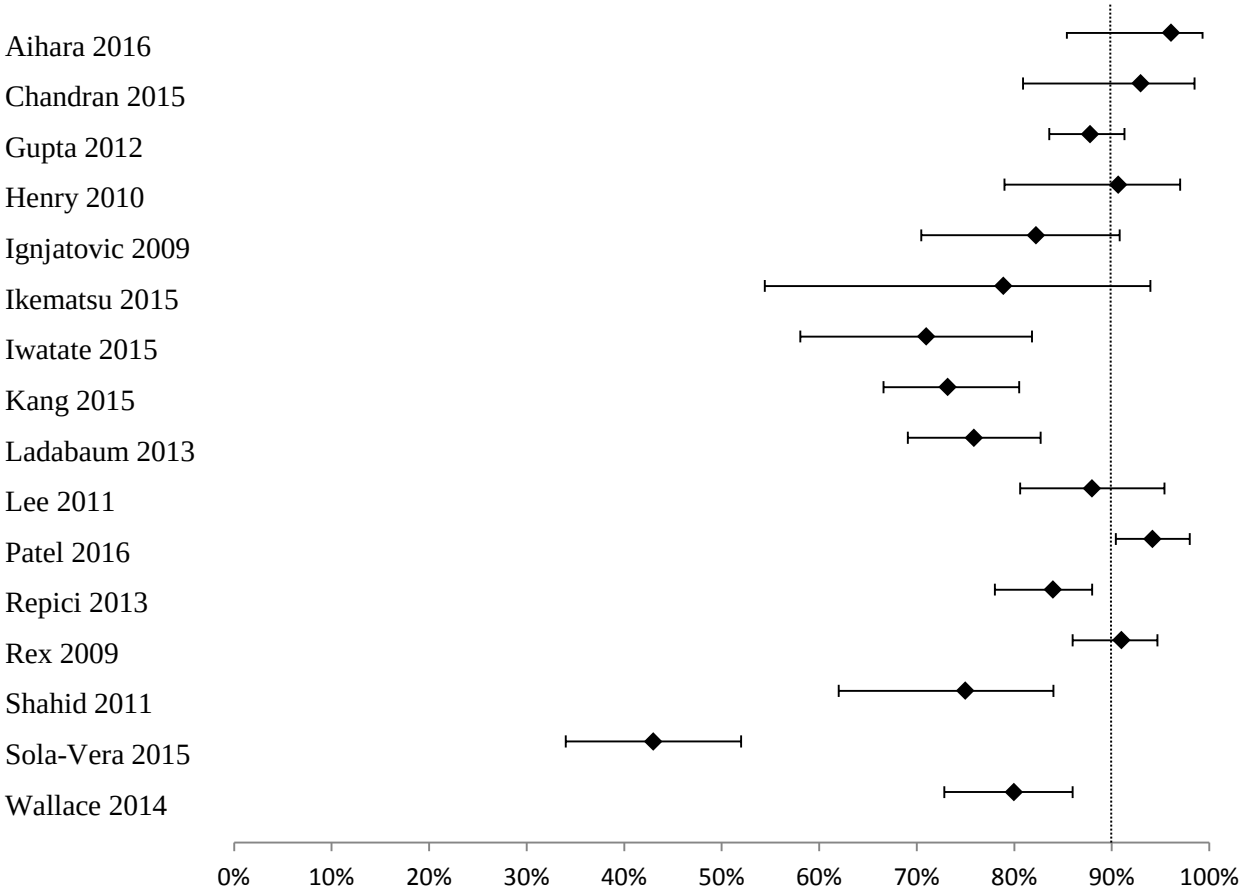


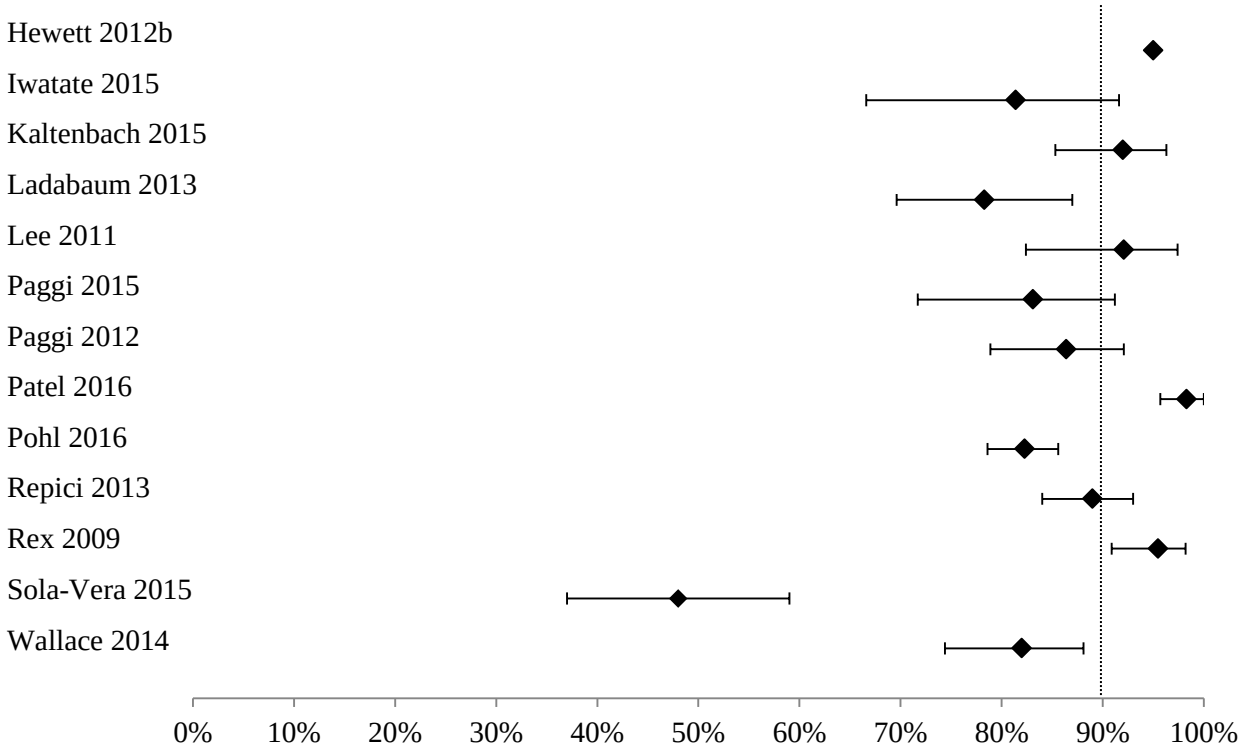
Figure 19 NPV of NBI for all characterisations of diminutive polyps in the whole colon (made with any level of confidence)

Table 14 Negative predictive value of NBI for the characterisation of diminutive polyps in the whole colon

	All characterisations		High confidence characterisations	
	Value	**95% CI**	**Value**	**95% CI**
Diminutive polyps whole colon				
Aihara et al.[54]	96.1%	85.4% to 99.3%	nr	nr
Chandran et al.[55]	93%	80.9% to 98.5%	nr	nr
Gupta et al.[56]	87.8% [a]	83.6% to 91.3% [a]	nr	nr
Henry et al.[57]	90.7%	79% to 97%	nr	nr
Hewett et al. 2012b[20]	nr	nr	95%	nr
Ignjatovic et al.[59]	82.3 % [a]	70.5% to 90.8% [a]	nr	nr
Ikematsu et al.[60]	78.9%	54.4% to 94.0% [a]	nr	nr

Iwatate et al.[61]	71.0%	58.1% to 81.8% [a]	81.4%	66.6% to 91.6% [a]
Kaltenbach et al.[62]	nr	nr	92.0%	85.3% to 96.3%
Kang et al.[78]	73.2%	66.6% to 80.5%	nr	nr
Ladabaum et al.[63]	75.9%	69.1% to 82.7%	78.3%	69.6% to 87.0%
Lee et al.[77]	88.0%	80.6% to 95.4%	92.1% [a]	82.4% to 97.4% [a]
Paggi et al. 2015[65]	nr	nr	83.1 % [a]	71.7% to 91.2% [a]
Paggi et al. 2012[66]	nr	nr	86.4% [a]	78.9% to 92.1% [a]
Patel et al.[67]	94.2%	90.4% to 98.0%	98.3%	95.7% to 100.0%
Pohl et al.[68]	nr	nr	82.3	78.6% to 85.6%
Repici et al.[70]	84%	78% - 88%	89%	84% to 93%
Rex et al.[71]	91.0 % [a]	86.0% to 94.7% [a]	95.5 % [a]	90.9% to 98.2% [a]
Rogart et al.[72]	nr	nr	nr	nr
Shahid et al.[73]	75%	62% to 84%	nr	nr
Sola-Vera et al.[74]	43%	34% to 52%	48%	37% to 59%
Vu et al.[75]	nr	nr	nr	nr
Wallace et al.[76]	80%	72.8% to 86.0% [a]	82%	74.4% to 88.1% [a]
Assessed by specialists in colonoscopy (whole colon)				
Iwatate et al.[61]	nr	nr	90.9%	70.8% to 98.9% [a]
Assessed by general endoscopists (whole colon)				
Iwatate et al.[61]	nr	nr	71.4%	47.8% to 88.7% [a]

[a] Calculated by reviewer



Note that no 95% confidence interval was reported for the Hewett 2012b study.

Figure 20 NPV of NBI for high confidence characterisations of diminutive polyps in the whole colon

Seven studies[56,58,63,67,68,70,76] reported on the NPV for the characterisation of diminutive polyps in the rectosigmoid colon (top section Table 15). Five of these studies[56,58,63,67,76] reported data for all diminutive polyp characterisations in the rectosigmoid colon and NPV ranged from 87.4% to 98.4%. In four[56,58,67,76] of these five studies NPV was over 90%. Only in the study by Ladabaum and colleagues,[63] was the 90% threshold not reached.

Data for high confidence characterisations of polyps in the rectosigmoid colon were reported by five of the seven studies.[58,67,68,70,76] In three of these five studies[58,67,76] the data on high confidence characterisations was provided in addition to data on all polyp characterisations in the rectosigmoid colon. In these studies the high confidence results led to NPVs that remained at over 90% and were slightly increased. Two studies[68,70] provided only high confidence results for the rectosigmoid colon and in both the NPV was over the 90% threshold. It is worth noting however that in two[70,76] of the five studies that report NPV for high confidence characterisations of diminutive polyps in the rectosigmoid colon, the lower limit of the 95% confidence interval falls below 90%.

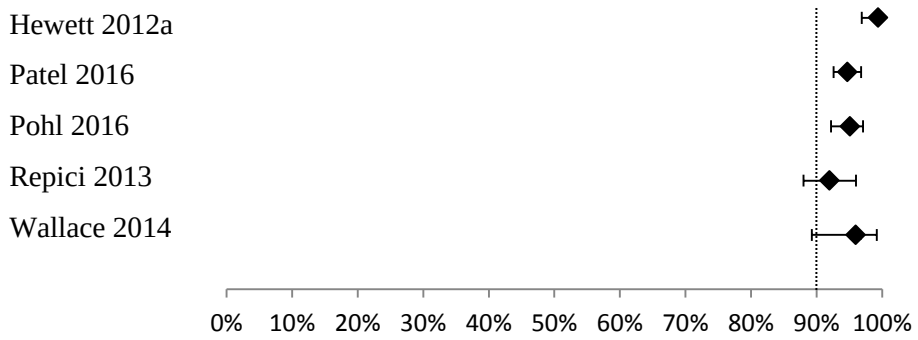


Figure 21 NPV of NBI for high confidence characterisations of diminutive polyps in the rectosigmoid colon

The NPV of NBI for characterisation of diminutive polyps in other regions of the colon (where reported by studies) is also presented in Table 15. Although the mean NPV was above the 90% threshold in some instances none of the lower limits of the 95% confidence interval lay above 90%.

One study[68] reported the NPV for characterisations of diminutive polyps in the rectosigmoid colon achieved by endoscopists with prior optical diagnosis experience in colonoscopy and by endoscopists without prior optical diagnosis experience. Endoscopists with prior optical diagnosis experience achieved an NPV of 96.6% (95% CI 92.7% to 98.7%) whereas the NPV achieved by endoscopists without prior optical diagnosis experience was lower at 93.5% (95% CI 88.7% to 96.7%).

Table 15 Negative predictive value of NBI for the characterisation of diminutive polyps in the rectosigmoid colon and other regions of the colon

	All characterisations		**High confidence characterisations**	
	Value	**95% CI**	**Value**	**95% CI**
Rectosigmoid colon diminutive polyps				
Gupta et al.[56]	95.4%	91.8 to 97.7	nr	nr
Hewett et al. 2012a[58]	98.4%	95.3% to 99.7%	99.4%	96.9% to 100%
Ladabaum et al.[63]	87.4%	82.5 to 92.4	nr	nr
Patel et al.[67]	93.7%	91.8% to 95.7%	94.7%	92.6% to 96.8%
Pohl et al.[68]	nr	nr	95.1%	92.2% to 97.1% [a]
Repici et al.[70]	nr	nr	92%	88%-96%
Wallace et al.[76]	95%	88.8% to 98.8% [a]	96%	89.3% to 99.2% [a]

Diminutive polyps located on the right side of the colon				
Kaltenbach et al.[62]	nr	nr	87.1%	70.2% to 96.4%
Diminutive polyps located proximal to the splenic flexure				
Pohl et al.6[68]	nr	nr	43.4%	33.5% to 53.8% [a]
Diminutive polyps located on the left side of the colon				
Gupta et al.[56]	93.0 % [a]	89.2% to 95.8% [a]	nr	nr
Kaltenbach et al.[62]	nr	nr	93.9%	83.1% to 98.7%
Diminutive polyps located in the distal colon				
Pohl et al.[68]	nr	nr	92.6%	89.4% to 95.0% [a]
Rectal diminutive polyps				
Kaltenbach et al.[62]	nr	nr	93.8%	79.2% to 99.2%
Diminutive polyps proximal to rectosigmoid colon				
Ladabaum et al.[63]	57.3%	38.4 to 76.2	nr	nr
Patel et al.[67]	65.6%	59.2% to 71.9%	77.1%	67.9 to 86.2%
Rectosigmoid colon diminutive polyps assessed by endoscopists with prior optical diagnosis experience in colonoscopy				
Pohl et al.[68] [b]			96.6%	92.7% to 98.7%
Rectosigmoid colon diminutive polyps assessed by endoscopists with no prior optical diagnosis experience in colonoscopy				
Pohl et al.[68] [b]			93.5%	88.7% to 96.7%

[a] Calculated by reviewer

[b] There is a discrepancy in this paper between reporting in the text (which states that NPV was for rectosigmoid diminutive adenomas) and in a table which means it is possible that the reported NPVs could relate to polyps in the distal and proximal colon rather than the rectosigmoid.

Accuracy of NBI

As well as measures such as sensitivity, specificity and NPV reported above, another global measure, diagnostic accuracy, can be calculated from the 2x2 table data. This is expressed as the proportion of correctly classified polyps (the sum of the true positive and true negative results) among all the ppolyps (true positive + true negative + false positive + false negative). Like NPV diagnostic accuracy is affected by disease prevalence such that at the same sensitivity and specificity diagnostic accuracy increases as disease prevalence decreases.

Accuracy of polyp characterisations in the whole colon was reported by, or could be calculated for, 16 studies.[54-57,59-61,63,67,70,71,73,74,76-78] Accuracy was 90% or more in five studies,[54,55,57,59,60] was between 76% and 89% in ten studies[56,61,63,67,70,71,73,76-78] and was only 63.9% in the final study.[74]

Thirteen studies[20,61-63,65-68,70,71,74,76,77] reported on the accuracy of high confidence polyp characterisations in the whole colon. Accuracy was 90% or more in two studies,[71,77] was between 81% and 90% in ten studies[20,61-63,65-68,70,76] and was only 68.5% in the final study.[74]

Accuracy of polyp characterisation was typically 3-5% higher among high confidence characterisations than all polyp characterisations in the eight studies[61,63,67,70,71,74,76,77] that reported both values.

Table 16 Accuracy (proportion of correctly classified polyps) with NBI

	Accuracy of polyp characterisations (95% CI)	**Accuracy of high confidence polyp characterisations (95% CI)**
Whole colon		
Aihara et al.[54]	90.1% (84.8 to 95.4)	nr
Chandran et al.[55]	91.2% [a]	nr
Gupta et al.[56]	84.8% (82.3 to 87.1)	nr
Henry et al.[57]	90.0% (82 to 95)	nr
Hewett et al. 2012b[20]	nr	88%
Ignjatovic et al.[59]	92%	nr
Ikematsu et al.[60]	90.3%	nr
Iwatate et al.[61]	79.5%	85.0%
Kaltenbach et al.[62]	nr	87.0% (82.8 to 90.5)
Kang et al.[78]	79.4 % (75.5 to 83.6)	nr
Ladabaum et al.[63]	78.1% (73.7 to 82.5)	81.1% (75.8 to 86.3)
Lee et al.[77]	87.8% (82.6 to 92.9)	91.2% [a]
Paggi et al. 2012[66]	nr	84.0%
Paggi et al. 2015[65]	nr	88.2% (83.9 to 92.5)
Patel et al.[67]	76.7% (75.2 to 78.3)	84.8% (82.1 to 87.5)
Pohl et al.[68]		83.2%
Repici et al.[70]	85%	89% (86 to 92)
Rex et al.[71]	88.6% [a]	93.0% [a]

Shahid et al.[73]	80% (70 to 87)	nr
Sola-Vera et al.[74]	63.9%	68.5%
Wallace et al.[76]	79%	82%
Whole colon by colonosopist type		
Iwatate et al.[61]		
- specialist colonoscopists	nr	90.7%
- generalist colonoscopists	nr	82.3%
Right colon		
Kaltenbach et al.[62]	nr	86.4% (80.0 to 91.4)
Proximal to splenic flexure		
Pohl et al.[68]		78.8%
Left colon		
Gupta et al.[56]	83.5% (80.0 to 86.6)	nr
Kaltenbach et al.[62]	nr	90.2% (83.4 to 94.8)
Distal colon		
Pohl et al.[68]		86.2%
Rectosigmoid colon		
Hewett et al. 2012a[58]	94.5% (91.5 to 97.6)	99.0% (97.6 to 100)
Ladabaum et al.[63]	77.4% (69.1 to 85.3)	nr
Patel et al.[67]	80.9% (76.7 to 85.1)	88.1% (83.2 to 92.9)
Repici et al.[70]		91% (87 to 95)
Pohl et al.[68]		87.6%
Wallace et al.[76]	84%	90%
Proximal to rectosigmoid colon		
Ladabaum et al.[63]	79.3% (74.7 to 83.9)	
Patel et al.[67]	78.8% (75.5 to 82.0	84.7% (80.7 to 88.6)
Rectum		
Kaltenbach et al.[62]		80.4% (66.1 to 90.6)

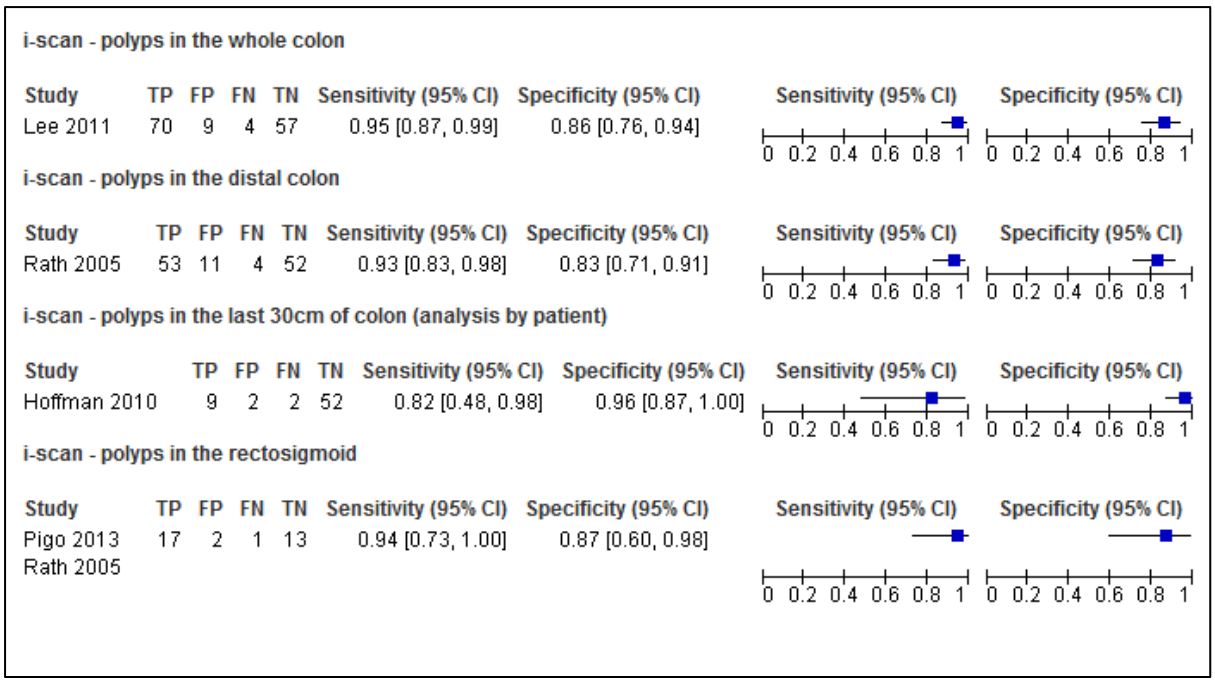
[a] calculated by reviewer

i-scan

Sensitivity and specificity of i-scan for the characterisation of diminutive colorectal polyps

Five studies[77,79-82] provided data on the characterisation of diminutive polyps as adenomas or hyperplastic polyps using i-scan with the characterisation verified by histopathological assessment of the resected polyps. The way in which data were reported by the studies varied. Two studies, Basford and colleagues[79] and Lee and colleagues,[77] reported on the characterisation of diminutive polyps within the whole colon. Basford and colleagues only presented data from the polyp characterisations where the endoscopist had high confidence they were correct whereas Lee and colleagues provided data for all characterisations and then separately for characterisations made with either high or low confidence (data for low confidence characterisations is available in Appendix 3). The other three studies presented data on the characterisation of diminutive polyps from within a part of the colon: the distal colon (Rath and colleagues[82]), the last 30cm of colon (Hoffman and colleagues[80]) where a per polyp analysis was not presented, only an analysis per patient), and the rectosigmoid colon (Pigo and colleagues[81] and Rath and colleagues[82] although it was not possible to impute 2x2 table data for this latter study). Rath and colleagues[82] also provided data separately for the polyp characterisations they had made with high confidence.

The results for all characterisations (i.e. not separated by confidence level) are shown in Figure 22. The ability of i-scan to correctly identify diminutive polyps as adenomas (ie. the sensitivity of the test) was above 90% in three of the four studies that reported results for all characterisations (Lee and colleagues,[77] Pigo and colleagues[81] and Rath and colleagues[82]) whereas sensitivity was only 82% in the per patient analysis reported by Hoffman and colleagues.[80] The ability of i-scan to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) was more variable across the studies ranging from 83% (Rath and colleagues[82] results for polyps in the distal colon) to 96% (Hoffman and colleagues[80]).

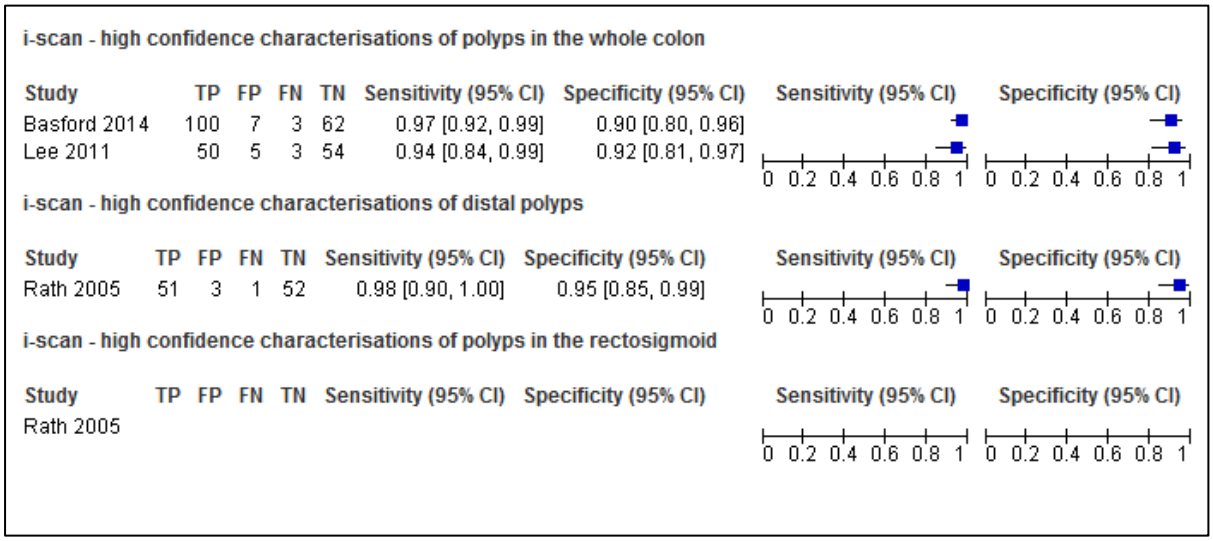


Rath 2005 presented summary data for polyps in the rectosigmoid colon but it was not possible for us to impute the 2x2 table data necessary to plot these results within this figure. The reported sensitivity was 90.3% (95% CI 73.1% to 97.5%) and specificity 87.5% (95% CI 74.1% to 94.8%).

Figure 22 Accuracy of i-scan for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps

Results for studies that reported results from polyp characterisations with i-scan that were designated as high confidence decisions are shown in Figure 23. The ability of high confidence characterisations made with i-scan to correctly identify diminutive polyps as adenomas (ie. the sensitivity of the test) in the three studies that provided data was 0.94 (i.e. 94%) (Lee and colleagues,[77]), 0.97 (Basford and colleagues[79]) and in the Rath and colleagues' study[82] 0.98 for distal polyps and 0.96 in the analysis limited to polyps in the rectosigmoid colon. For the Lee and colleagues' study[77] the sensitivity achieved from high confidence polyp characterisations was slightly lower than that obtained from all the polyp characterisations 0.94 (95%CI 0.84 to 0.99) versus 0.95 (95% CI 0.87 to 0.99) whereas the reverse was true for the Rath and colleagues' study[82] for both the data set for distal polyps and for rectosigmoid colon polyps (distal polyps: high confidence 0.98, 95% CI 0.90 to 1.00 versus overall 0.93, 95% CI 0.83 to 0.98; rectosigmoid colon: high confidence 0.96, 95% CI 0.80 to 1.0 versus overall 0.90, 95% CI 0.73 to 0.98). The ability of i-scan to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) when the characterisation was made with high confidence was over 0.90 (i.e. 90%) or more in all three studies. Furthermore, the specificity of i-scan arising from high confidence decisions was greater than the

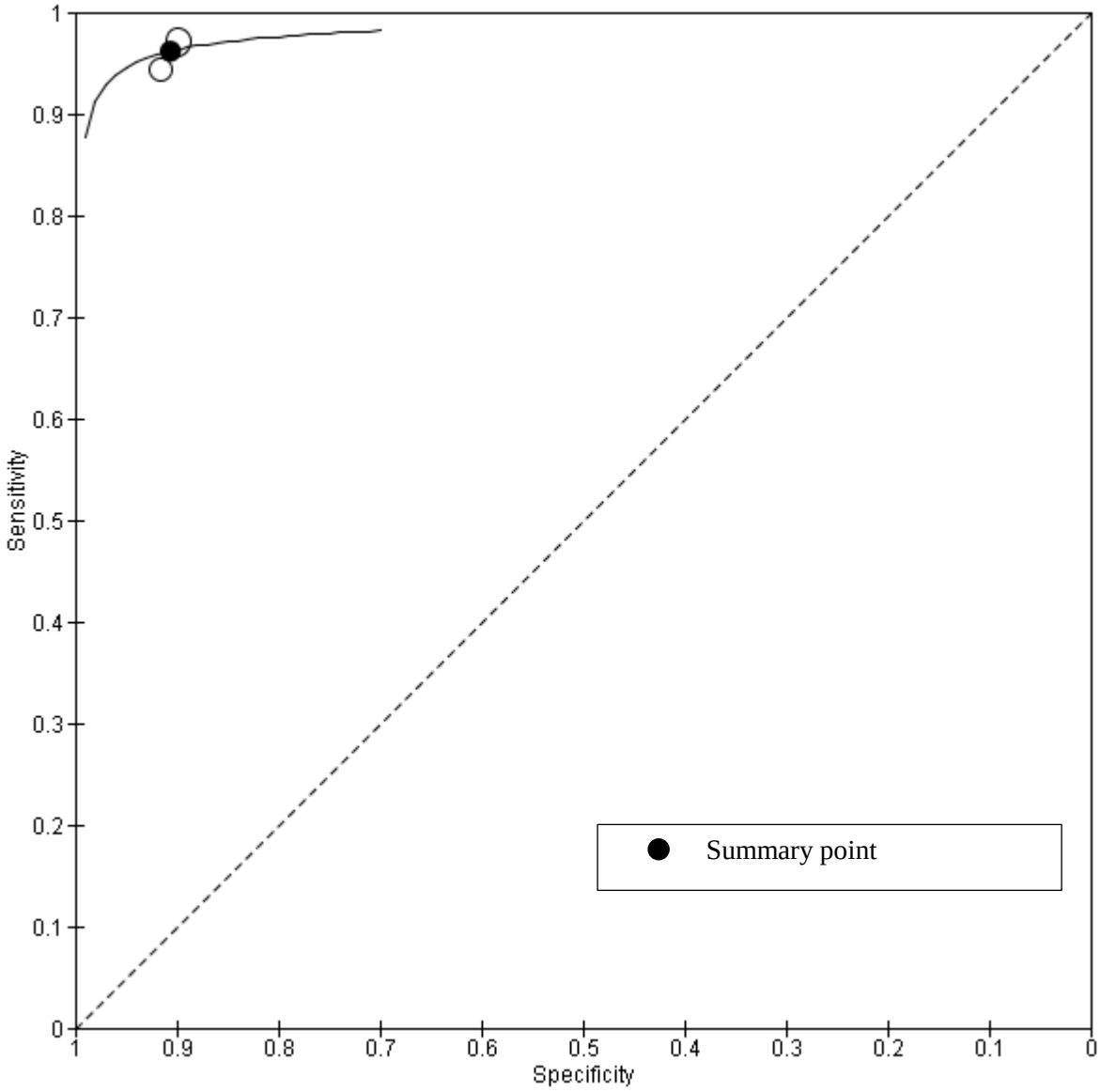
specificity observed when all the polyp characterisations were taken into account in the two studies that reported both sets of data (Lee and colleagues[77] 92% versus 86%; Rath and colleagues[82] distal polyps: 95% versus 83%; rectosigmoid colon polyps 95.5% versus 87.5%). The 2005 Rath and colleagues[82] study which was conducted in Germany among patients attending for screening or surveillance colonoscopy and which reported on characterisation of distal polyps (polyps in the descending colon, the sigmoid colon, or the rectum) achieved the best sensitivity (98%) which was coupled to the second highest value for specificity (95%). However, in common with the other studies providing data on i-scan, a single endoscopist working in what appears to be a specialist endoscopy centre achieved these results so it is not clear how transferable these results would be to less experienced endoscopists working in less specialist settings.



Rath 2005 presented summary data for high confidence characterisations of polyps in the rectosigmoid colon but it was not possible for us to impute the 2x2 table data necessary to plot these results within this figure. The reported sensitivity was 96.4% (95% CI 79.8% to 99.8%) and specificity 95.5% (95% CI 83.3% to 99.2%).

Figure 23 Accuracy of i-scan high confidence characterisations of diminutive colorectal polyps as either adenomas or hyperplastic polyps

A bivariate meta-analysis was run (using Stata/IC14 and xtmelogit) to provide a summary estimate for the two studies that reported high confidence characterisations of polyps in the whole colon, which could be used in the economic model. This produced a summary value for sensitivity of 0.96 (95% CI 0.92 to 0.98) and for specificity of 0.91 (95% CI 0.84 to 0.95). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 24 in which the individual study estimate points are scaled to the sample size of the study.



Note that the software used to draw the SROC plot (Review Manager 5.3) did not generate a 95% confidence region or a 95% prediction region for this meta-analysis. It is presumed that this is because of the small number of studies.

Figure 24 SROC plot from the meta-analysis of i-scan for high confidence characterisations of polyps in the whole colon.

Negative predictive value of i-scan for the characterisation of diminutive colorectal polyps

As previously stated the negative predictive value is the probability that subjects with a negative screening test (i.e. colorectal polyp is characterised as hyperplastic) truly do not have an adenoma. However, it must be borne in mind when viewing these results that the negative predictive value is influenced by the

prevalence of disease (i.e. in this case the prevalence of adenomas in the tested populations). When prevalence is increased the result is a decrease in the negative predictive value.

Two studies[77,80] reported NPV for the characterisations of diminutive polyps in the whole colon (made with any level of confidence) although one of these studies, Hoffman and colleagues,[80] only reported a per patient analysis. Although the mean NPV was greater 90% the lower limit of the 95% confidence interval fell below 90% in both studies (Table 17). High confidence characterisation of polyps in the whole colon was reported by two studies.[77,79] Basford and colleagues[79] reported an NPV of 100% (95% CI 93.4% to 100%) and Lee and colleagues an NPV of 94.7% (95% CI 85.4% to 98.9%).[77]

Two studies reported on the NPV for the characterisation of diminutive polyps in the distal portion of the colon[82] or the rectosigmoid colon,[81,82] with Rath and colleagues[82] also reporting on high confidence characterisations. In all cases although the point estimate for NPV lay above the 90% threshold the lower limit of the 95% confidence interval fell below this.

Table 17 Negative predictive value of i-scan for the characterisation of diminutive polyps

	All characterisations		**High confidence characterisations**	
	Value	95% CI	Value	95% CI
Whole colon				
Basford et al.[79]	nr	nr	100%	93.4% to 100%
Hoffman et al.[80] (per patient analysis)	96.3 %[a]	87.3% to 99.6%[a]	nr	nr
Lee et al.[77]	93.4%	87.2 to 99.7%	94.7%[a]	85.4% to 98.9%[a]
Distal polyps				
Rath et al.[82]	93.2%	82.7% to 97.8%	98.1%	88.4% to 99.1%
Rectosigmoid colon polyps				
Pigo et al.[81]	93%	81% to 100%	nr	nr
Rath et al.[82]	93.3 %	80.1% to 98.3%	97.7 %	86.2% to 99.9%

[a] Value calculated by reviewer. nr - not reported.

Accuracy of i-scan

Diagnostic accuracy (the proportion of correctly classified polyps among all the polyps) was reported either for all diminutive polyp characterisations,[80,81] for only high confidence polyp characterisations[79] or for both[77,82] (Table 18) with three studies providing data for the characterisations of polyps in the whole colon[77,79,80] and a single study for polyps in the rectosigmoid colon[81] or distal polyps.[82] Like NPV diagnostic accuracy is affected by disease prevalence such that at the same sensitivity and specificity diagnostic accuracy increases as disease prevalence decreases.

Accuracy was 90% or more in all the studies[77,79-82] and the accuracy of high confidence polyp characterisations was higher that among all polyp characterisations in the two studies that reported both values.[77,82]

Table 18 Accuracy (proportion of correctly classified polyps) with i-scan

	Accuracy of polyp characterisations, % (95% CI)	**Accuracy of high confidence polyp characterisations, % (95% CI)**
Whole colon		
Basford et al.[79]	nr	94.2% (92.8 to 99.2)
Hoffman et al.[80]	94% (per patient analysis)	nr
Lee et al.[77]	90.7% (85.9 to 95.5)	92.9%
Rectosigmoid colon		
Pigo et al.[81]	91% [a]	nr
Distal polyps		
Rath et al.[82]	90.1%	96.3%

[a] calculated by reviewer

FICE

Sensitivity and specificity of FICE for the characterisation of diminutive colorectal polyps

Three studies [78,83,84] provided data on the characterisation of diminutive polyps as adenomas or hyperplastic polyps using FICE compared to characterisation verified by histopathological assessment of the resected polyps. In all three studies the characterisations were made on polyps in any part of the colon,

and in all three the level of confidence with which the characterisation was made was not stated. The results of the polyp characterisations are shown in Figure 25.

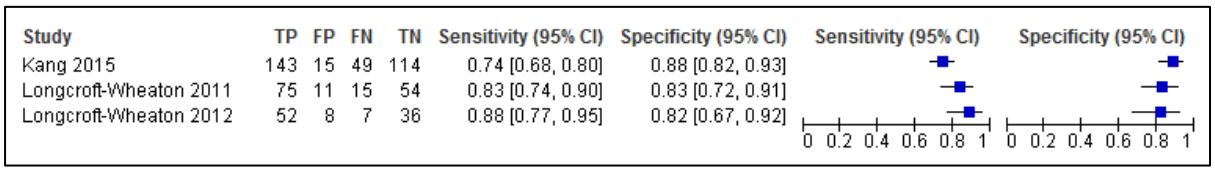
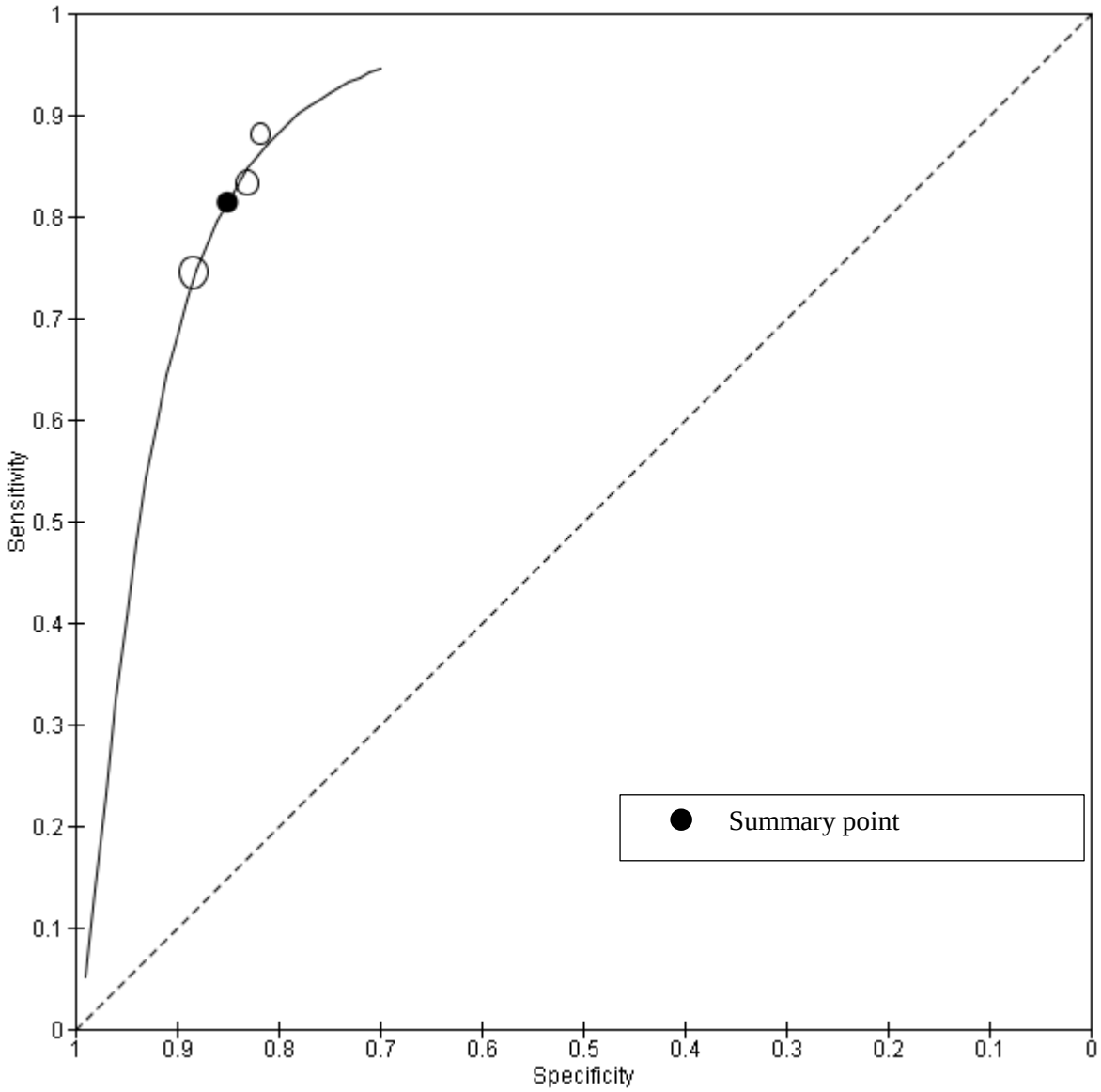


Figure 25 Accuracy of FICE for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps

The ability of FICE to correctly identify diminutive polyps as adenomas (ie. the sensitivity of the test) ranged from 74% to 88% across the studies. The ability of FICE to correctly identify diminutive polyps as hyperplastic polyps (i.e. the specificity of the test) had a narrower range across the studies, from 82% to 88%.

It was possible to run a bivariate meta-analysis (using Stata/IC14 and xtmelogit) with data from the three studies. This produced a summary value for sensitivity of 0.81 (95% CI 0.73 to 0.88) and for specificity of 0.85 (95% CI 0.79 to 0.90). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 26 in which the individual study estimate points are scaled to the sample size of the study.



Note that the software used to draw the SROC plot (Review Manager 5.3) did not generate a 95% confidence region or a 95% prediction region for this meta-analysis. It is presumed that this is because of the small number of studies.

Figure 26 SROC plot from the meta-analysis of FICE for all characterisations of polyps in the whole colon.

Negative predictive value of FICE for the characterisation of diminutive colorectal polyps

Table 19 reports the NPVs for the three FICE studies. These ranged from 70% to 84%.

Table 19 Negative predictive value of FICE for the characterisation of diminutive colorectal polyps

Study	Value	95% CI
Kang et al.[78]	70%	63% to 77%
Longcroft-Wheaton et al. 2011[83]	78%	70% to 84%
Longcroft-Wheaton et al. 2012[84]	84%[a]	69% to 93%[a]

[a] Value calculated by the reviewer

Accuracy of FICE

The three studies that reported on the use of FICE provided diagnostic accuracy (the proportion of correctly classified polyps among all the polyps) for all diminutive polyp characterisations in the whole colon (Table 20).[78,83,84] The reported diagnostic accuracy values ranged from 80% to 85%.

Table 20 Accuracy (proportion of correctly classified polyps) with FICE

	Accuracy of polyp characterisations, % (95% CI)	Accuracy of high confidence polyp characterisations, % (95% CI)
Whole colon		
Kang et al.[78]	80.1% (75.8 to 84.6)	nr
Longcroft-Wheaton et al. 2011[83]	83% (77% to 88%)	nr
Longcroft-Wheaton et al. 2012[84]	85% (76 to 91)	nr

Post-hoc pooled analysis of all virtual chromoendoscopy technologies

The appropriateness of pooling evidence from different virtual chromoendoscopy technologies together is uncertain. The technologies certainly all aim to enhance surface vessel patterns but the technologies use different methods to achieve this. We have therefore assumed that there is a 'class effect'and that they can be meaningfully pooled.

A pooled analysis of the studies included in this assessment for which 2x2 data were available was undertaken in order to inform a scenario analysis using the economic model (section 5.5.2). Data for high confidence assessments of polyps in the whole colon were available from 11 NBI studies and two i-scan studies (note that Lee and colleagues[77] contribute data on NBI and i-scan) (Figure 27). No FICE data were available to include in this analysis because the FICE studies did not report high confidence polyp characterisation separately.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Basford 2014	100	7	3	62	0.97 [0.92, 0.99]	0.90 [0.80, 0.96]
Iwatate 2015	107	17	8	35	0.93 [0.87, 0.97]	0.67 [0.53, 0.80]
Kaltenbach 2015	178	33	9	103	0.95 [0.91, 0.98]	0.76 [0.68, 0.83]
Lee 2011 (i-scan)	50	5	3	54	0.94 [0.84, 0.99]	0.92 [0.81, 0.97]
Lee 2011 (NBI)	56	6	5	58	0.92 [0.82, 0.97]	0.91 [0.81, 0.96]
Paggi 2012	233	48	16	102	0.94 [0.90, 0.96]	0.68 [0.60, 0.75]
Paggi 2015	140	15	11	54	0.93 [0.87, 0.96]	0.78 [0.67, 0.87]
Patel 2016	1296	264	32	586	0.98 [0.97, 0.98]	0.69 [0.66, 0.72]
Pohl 2016	408	77	84	391	0.83 [0.79, 0.86]	0.84 [0.80, 0.87]
Repici 2013	175	21	20	152	0.90 [0.85, 0.94]	0.88 [0.82, 0.92]
Rex 2009	145	15	7	147	0.95 [0.91, 0.98]	0.91 [0.85, 0.95]
Sola-Vera 2015	67	4	47	44	0.59 [0.49, 0.68]	0.92 [0.80, 0.98]
Wallace 2014	102	22	24	109	0.81 [0.73, 0.87]	0.83 [0.76, 0.89]

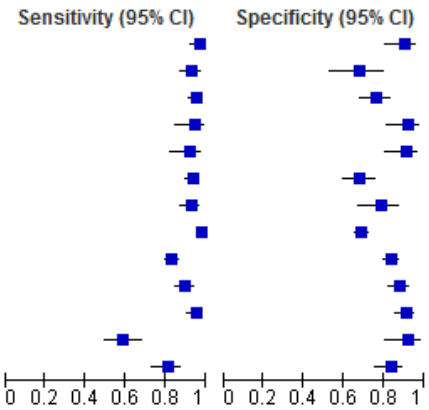


Figure 27 Accuracy of virtual chromoendoscopy high confidence decisions for characterising diminutive colorectal polyps as either adenomas or hyperplastic polyps in the whole colon

A bivariate meta-analysis (using Stata/IC14 and metandi[44]) was carried out which produced a pooled summary estimate for sensitivity of 0.92 (95% CI 0.87 to 0.95) and for specificity of 0.83 (95% CI 0.78 to 0.87). The parameter estimates for the bivariate model were entered into RevMan to produce the SROC plot shown below in Figure 28 in which the individual study estimate points are scaled to the sample size of the study (i.e. larger circles represent larger studies). The virtual chromoendoscopy pooled estimates for sensitivity and specificity do not differ greatly from the NBI pooled estimates (Figure 9) which is unsurprising given that the bulk of the evidence comes from studies of NBI.

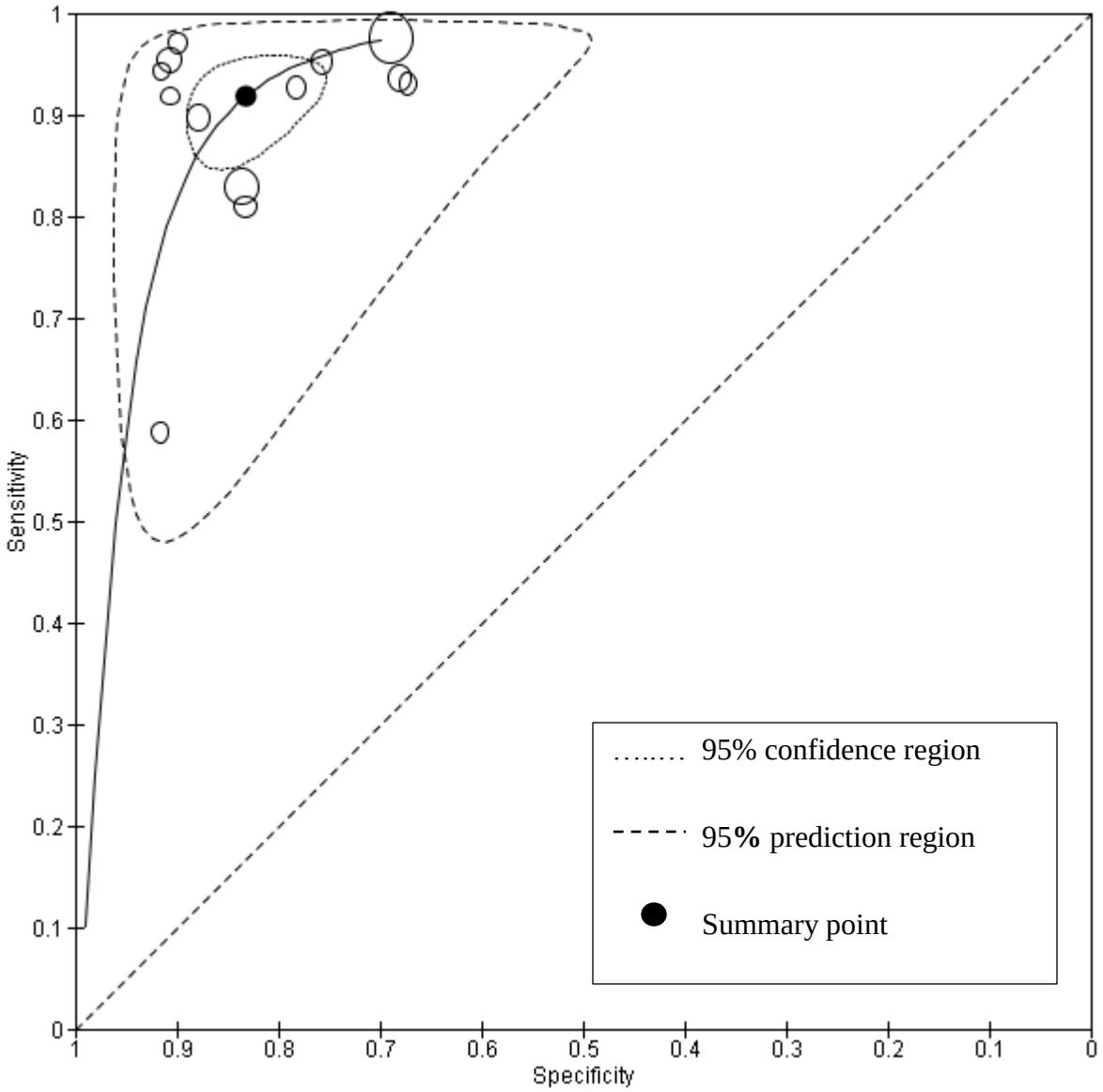


Figure 28 SROC plot showing the summary point on the summary curve from the meta-analysis of virtual chromoendoscopy high confidence decisions for characterising diminutive colorectal polyps in the whole colon

A pooled analysis of the virtual chromoendscopy studies for high confidence assessments of polyps in the rectosigmoid colon, equivalent to that above for the whole colon, has in essence already been presented earlier in this assessment. This is because the only data available for this analysis come from NBI studies and thus the results presented in Figure 15 and Figure 16 represent all the available data on high confidence assessments of polyps in the rectosigmoid colon, there are no equivalent data for i-scan or FICE.

4.1.3 Assessment of test impact on recommended surveillance intervals

NBI

Thirteen studies[55,56,59,62,63,66-68,70,71,74-76] reported results on the impact that the use of NBI would have on recommended surveillance intervals (in comparison to surveillance intervals calculated following histopathology of all polyps). The agreement between the surveillance interval allocated using an NBI based strategy and using the results of histopathology for all polyps ranged from 84%[75,76] to 99 %.[70] Eleven of the 13 studies reporting on this outcome achieved a level of agreement that was above 90%[55,56,59,62,63,67,68,70,71,74,76] although for three of these studies[56,63,76] an agreement of over 90% was only achieved by one of the tested strategies (in two studies using a modified recommendation of colonoscopy in 10 years for patients with 1-2 small adenomas instead of 5 years,[56,63] and in one study limiting the analysis to where there was a high confidence predictions for polyps ≤5mm[76]). Where there were discrepancies between the surveillance interval assigned using the NBI based strategy and the histopathology only strategy some studies reported whether the NBI strategy led to longer or shorter surveillance intervals being assigned. In the majority of studies where a discrepancy in the surveillance interval was reported, the NBI containing strategy more often led to shorter surveillance intervals being set (i.e. patients recalled for a colonoscopy earlier than would have been the case with the histopathology based surveillance interval) than longer surveillance intervals. There were however some exceptions, particularly the study by Repici and colleagues[70] where, if there was a difference between the surveillance intervals assigned, the NBI containing strategy was more likely to lead to the assignment of a longer interval (i.e. patients not recalled for repeat colonoscopy as early as they would have been with the histopathology based surveillance interval) than a shorter one.

Nine studies clearly calculated the concordance of surveillance intervals between virtual chromoendoscopy and histopathology in line with the PIVI requirements.[55,62,63,65,68,70,71,75,76] The criterion of the PIVI statement, that agreement should be ≥90%, was met by all but one study,[75] with one further study meeting the PIVI criterion in one of the two tested strategies.[63] Where the agreement was ≥90% the lower limit of the 95% confidence interval (where reported) fell below 90% in two instances.[67,70]

Table 21 Surveillance interval prediction

Study	Guideline used for determining surveillance interval (as cited by the study)	Surveillance interval correctly allocated [95% CI] (n/N)	Shorter or longer intervals set with NBI, n (% of total allocations)
Chandran et al.[55]	NHMRC, Australia 2011[100]	98% (92/94)	2 (2%) shorter
Gupta et al.[56]	Multi-Society Task Force 2008[101]: A] colonoscopy in 3 years for patients with ≥3 adenomas or ≥1 advanced adenomas, 5 years for patients with 1-2 small adenomas without advanced histology,& 10 years for patients with 0 adenomas	86.1% [95% CI 82.4 to 89.3]	
	Multi-Society Task Force 2008[101]: B] colonoscopy in 3 years for patients with ≥3 adenomas or with ≥1 advanced adenomas & 10 years for patients with 1-2 small adenomas or 0 adenomas.	94.1% [95% CI 91.4 to 96.2]	
Ignjatovic et al.[59]	BSG guidelines 2002[28] (& based on patients with no polyps >10mm)	98% (80/82)	2 (2%) shorter
Kaltenbach et al.[62]	US Multi-Society Task Force 2012[102]		
	- Overall	92.2% (259/281)	
	- High confidence NBI diagnosis+ histopathology for all other polyps	95.2% (200/210)[a]	7 (3.3%) shorter 3 (1.4%) longer
Ladabaum et al.[63]	Multi-Society Task Force 2008[101]		
	- All study colonoscopies (n=1646)	88.4% [95% CI 86.8 to 89.9]	
	- All study colonoscopies with ≥1 diminutive polyp characterised with high confidence (n=1065)	79.9% [95% CI 77.4 to 82.3][a]	136 (13%) shorter 78 (7%) longer
	Using modified recommendations 2012[56] (10 year for 1-2 small adenomas)		

Study	Guideline used for determining surveillance interval (as cited by the study)	Surveillance interval correctly allocated [95% CI] (n/N)	Shorter or longer intervals set with NBI, n (% of total allocations)
	- All study colonoscopies (n=1646)	98.4% [95% CI 97.6 to 98.9]	
	- All study colonoscopies with ≥1 diminutive polyp characterised with high confidence (n=1065)	96.8% [95% CI 95.6 to 97.8][a]	24 (2%) shorter 10 (1%) longer
Paggi et al. 2012[65]	US Multi-Society Task Force on Colorectal Cancer (USMSTF) 2006.[103]	85.3% (168/197)	22 (11%) shorter 7 (4%) longer
Patel et al.[67]	US Multi-Society Task Force 2012[102]	91.2% [95% CI 89.67 to 92.65] (1279/1403)[a]	82 (5.8%) shorter 39 (2.8%) longer
Pohl et al.[68] [b]	US multi-society taskforce guidelines[102,104]		
	- All study colonoscopies	96%	
	- All study colonoscopies with ≥1 diminutive polyp (n=566)	93%[a]	24 (4%) shorter 15 (3%) longer
Repici et al.[70]	European Guidelines 2010[105]: ≥1 polyp ≤5mm characterised with high confidence	99% [95% CI, 97%-100%][a]	3 (1%) longer
	Multi-Society Task Force 2008[101]		
	- ≥ 1 polyp ≤5mm characterised with high confidence & 5- year interval for non-advanced adenomas ≤ 2mm	92% [95% CI 88%-96%][a]	5 (2%) shorter 12 (4%) longer
	- ≥ 1 polyp ≤5mm characterised with high confidence & 10- year interval for non-advanced adenomas ≤ 2mm	99% [95% CI, 97%-100%][a]	3 (1%) longer
Rex et al.[71]	US Multi-Society Task Force on Colorectal Cancer (USMSTF) 2006.[103]		
	- Colonscopy in 5 years if 1 or 2 tubular adenomas <1 cm in size.	94% (128/136)[a]	4 (3%) shorter 4 (3%) longer
	- Colonscopy in 10 years if 1 or 2 tubular adenomas <1 cm in size.	98.5% (134/136)[a]	2 (1%) shorter 1 (0.7%) longer

Study	Guideline used for determining surveillance interval (as cited by the study)	Surveillance interval correctly allocated [95% CI] (n/N)	Shorter or longer intervals set with NBI, n (% of total allocations)
Sola-Vera et al.[74]	European Guideline 2012[106]	97.8% (46/47)	nr
	ESGE Guideline[107]	97.8% (46/47)	nr
Vu et al.[75]	Multi-Society Task Force 2008,[101] high confidence predictions	84.1%[a]	nr
Wallace et al.[76]	Based only on number and size of adenomas[108]		
	- All predictions	84% [95% CIs 79% - 88%] (221/264)	27 (10%) shorter 16 (6%) longer
	- High confidence predictions for polyps ≤5mm	95% [95% CIs 91% - 97%] (250/264)[a]	5 (2%) shorter 9 (3%) longer

NHMRC - National Health and Medical Research Council

[a] Results from analyses of surveillance interval agreement in accordance with PIVI requirements

[b] Pohl and colleagues also reported surveillance interval results by colonoscopists experience and there was no statistically significant difference between the two (Appendix 3).

i-scan

Two studies[79,82] examined the effect that use of i-scan had on recommended surveillance intervals in comparison to those that were allocated based on histopathological assessment of all polyps (Table 22). Both studies[79,82] used *in vivo* diagnosis of diminutive polyps to guide surveillance interval decisions in accordance with the PIVI requirements. Both studies[79,82] also calculated agreement in surveillance intervals between i-scan and histopathology when using two different guidelines for determining the surveillance interval. Across these two studies, a surveillance interval agreement of over 90% was achieved regardless of the guideline used, with agreement ranging from 93.2%[82] to 97.2%.[79] In the study by Basford and colleagues,[79] identical results (an agreement of 97.2%) were achieved when using both the guidelines assessed. Both studies reported whether using i-scan resulted in a longer or shorter surveillance interval being allocated than that allocated by histopathology. In the Basford and colleagues' study,[79] two patients were set a shorter interval with i-scan and one patient a longer interval. In the Rath and colleagues' study,[82] i-scan tended to results in longer intervals being allocated than with histopathology, except in one case.

Table 22 Surveillance interval prediction using i-scan

Study	Guideline used for determining surveillance interval (as cited by the study)	Surveillance interval correctly allocated % [95% CI] (n/N)	Longer or shorter intervals set with i-scan, n (% of total allocations)
i-scan surveillance intervals based on high confidence assessment of all diminutive polyps combined with histology of polyps >5mm			
Basford et al.[79]	American Society of Gastroenterology (ASGE)[102] and British Society of Gastroenterology (BSG) guidelines[28]	97.2% [not reported] (80/83)	2 (2.4%) shorter 1 (1.2%) longer
i-scan surveillance intervals based on high confidence assessment of all distal polyps			
Rath et al.[82] [a]	European guidelines[106]	94.5% [not reported] (69/73)	4 (5.5%) longer
	US guidelines[102]	93.2% [not reported] (68/73)	1 (1.4%) shorter 4 (5.5%) longer

[a] The surveillance intervals determined in this study were based on the assessment of polyps in the distal colon only. Surveillance intervals for polyps in the rectosigmoid were also reported, but are not presented here.

FICE

Two studies[83,84] reported results on the impact that the use of FICE would have on recommended surveillance intervals (in comparison to surveillance intervals calculated following histopathology of all polyps) although neither assessed this in accordance with the PIVI criteria. This analysis, in both of these studies, included polyps <10mm (i.e. neither was restricted to diminutive polyps). The agreement between the surveillance interval allocated using a FICE based strategy and using the results of histopathology was 100% in one study[83] and 97% in the other study[84] regardless of whether the BSG or ASGE guidelines were used to determine the surveillance intervals. In the single study where there was a discrepancy for two participants between the surveillance interval assigned using the FICE based strategy and the histopathology strategy it is not known whether the FICE based strategy led to a longer or a shorter surveillance interval being set (Table 23).

Table 23 Surveillance interval prediction using FICE

Study	Guideline used for determining surveillance interval (as cited by the study)	Surveillance interval correctly allocated % [95% CI] (n/N)	Longer or shorter intervals set with FICE, n (% of total allocations)
Longcroft-Wheaton et al. 2012[83] [a]	British Society of Gastroenterology (BSG)[28]	100% (38/38)	n/a
	ASGE[109]	100% (38/38)	n/a
Longcroft-Wheaton et al. 2011[84] [a]	British Society of Gastroenterology (BSG)[28]	97% [89% to 100%] (67/69)	Not reported
	ASGE[109]	97% [89% to 100%] (67/69)	Not reported

a Patients with lesions >10mm would have required histology to set the surveillance interval and so were excluded from these analyses.

4.1.4 Assessment of other outcomes

In addition to the outcomes reported above on test accuracy and the impact on recommended surveillance intervals the review also aimed to report data on the interpretability of the tests; inter-observer agreement; intra-observer agreement; test acceptability (to patients and/or clinicians); adverse events; the number of polyps designated to be left in place; the number of polyps designated to be resected and discarded; the number of polyps designated for resection and histopathological examination; the length of time to perform the colonoscopy; the number of outpatient appointments; health-related quality of life; colorectal cancer and mortality.

NBI

None of the studies reported on the interpretability of the test; test acceptability (to patients and/or clinicians), number of outpatients appointments, health-related quality of life, colorectal cancer, or mortality.

One study, Lee and colleagues[77] reported on inter-observer agreement although this was the agreement between the characterisation obtained during real-time assessment and that obtained by an independent reader who reviewed all recorded endoscopic images whilst blind to the real-time assessement and the histopathology results. The inter-observer agreement was 86.5% with a k value (95% CI) of 0.730 (0.623

to 0.837) which represents 'substantial' agreement. One other study, Rogart and colleagues[72] reported inter-observer agreement for 20 test images but as this did not include any real-time assessment these data were not extracted. Lee and colleagues[77] were also the only researchers to report on intra-observer agreement. This was the agreement between the between the characterisation obtained during real-time assessment and that obtained by the same endoscopist who reviewed all recorded endoscopic images 1-3 months after the colonoscopy. The intra-observer agreement was 89.7% with a k value (95% CI) of 0.795 (0.699 to 0.890) again representing 'substantial' agreement.

Adverse events were not reported by most studies.[20,54-61,63,65-68,70-72,74-76,78] Of the three studies that did make mention of potential adverse events[62,73,77] the reports all indicated that no events had occurred. Kaltenbach and colleagues[62] reported no postpolypectomy bleeding, coagulation syndrome, perforation, or optical misdiagnosis of advanced histology, Lee and colleagues[77] stated that participants did not experience any procedure-related complications and Shahid and colleagues[73] stated that none of the patients experienced any endoscopic complications.

Ignatovic and colleagues[59]reported on the number of diminutive polyps that would have been left in place if the management strategy was to leave diminutive hyperplastic polps in situ in the recto-sigmoid colon. The endoscopists in this study made a high confidence optical diagnosis for 323 polyps (<10mm in this study) and of these, 33 would have been left in situ. All 33 were correctly predicted to be hyperplastic polyps and all were located in the sigmoid colon or the rectum. One other study, Repici and colleagues,[70] made a statement indicating that in their study, a discard type strategy would have reduced the need for polypectomy by 48%.

Two studies reported on the number of polyps that would have been resected and discarded if a resect and discard type of management strategy had been in place. Gupta and colleagues[56] reported a hypothetical strategy in which if all the 884 diminutive polyps in their study (where the total number of polyps of any size was 1254) were resected and discarded this would represent a 70.5% reduction in histopathology. Using this strategy 13 adenomas with advanced histologic features would have been discarded. However, it must be noted that this study did not record whether characterisations were made with high or low confidence and did not report how many diminutive polyps were in the rectosigmoid colon. Ignatovic and colleagues[59] reported a high confidence optical diagnosis was made for 323 polyps (<10mm in this study) and of these 290 would have been resected and discarded. The Ignatovic and colleagues' study[59] was the only NBI study to ask endoscopists to identify polyps that they would have sent electively to histopathology even if a policy of optical diagnosis had been in place. These were polyps where the

optical diagnosis was made with low confidence or where no optical diagnosis could be made. For the sub-group of diminutive polyps in this study 8% (22 of 293 polyps) would have been sent for elective histopathology.

The length of time taken to perform the withdrawal phase of the colonoscopy was reported by three studies. Kaltenbach and colleagues[62] reported a mean withdrawal time of 10.3 minutes (SD 5.7, range 3.3 to 58 minutes). A procedure time of 12 seconds is reported but a definition of procedure time is not provided in the study publication so it is not clear what this comprises. In the Kang and colleagues[78] study the mean withdrawal time in the NBI group was 13.5 minutes (SD 7.3) whilst in the Wallace and colleagues' study[76] it was 16.1 minutes (SD 7.3). A fourth study, Shahid and colleagues,[73] reported that the average withdrawal time at their centre was typically eight to 10 minutes but it was not reported specifically for their study. However they did report that NBI inspection time was typically less than a minute.

i-scan

None of the studies reported on the interpretability of the test, test acceptability (to patients and/or clinicians), number of polyps designated to be left in place, number of polyps designated to be resected and discarded, number of polyps designated for resection and histopathological examination, number of outpatients appointments, health-related quality of life, colorectal cancer, or mortality.

One study, Lee and colleagues[77] reported on inter-observer agreement although this was the agreement between the characterisation obtained during real-time assessment and that obtained by an independent reader who reviewed all recorded endoscopic images whilst blind to the real-time assessement and the histopathology results. The inter-observer agreement was 87.9% with a k value (95% CI) of 0.751 (0.640 to 0.861) which represents 'substantial' agreement. One other study, Pigo and colleagues[81] reported inter-observer agreement but this was based on endoscopists' assessments of still images so because this did not include any real-time assessment these data were not extracted. Two studies, Lee and colleagues[77] and Rath and colleagues[82] reported on intra-observer agreement. In the Lee and colleagues' study[77] this was the agreement between the characterisation obtained during real-time assessment and that obtained by the same endoscopist who reviewed all recorded endoscopic images 1-3 months after the colonoscopy. The intra-observer agreement was 86.4% with a k value (95% CI) of 0.729 (0.616 to 0.841) again representing 'substantial' agreement. In the Rath and colleagues' study[82] it is not clear how intra-observer agreement was assessed because no details are reported in the paper. The authors state that agreement was achieved in 113 out of 121 polyps (93.4 %) with a κ coefficient of agreement of 0.867 [95 % CI: 0.799–0.967]

which indicated almost perfect agreement. In the Pigo and colleagues' study[81] intra-observer agreement was assessed based on the endoscopists' assessment of still images rather than real-time assessment and furthermore the intra-observer agreement for the evaluation of diminutive polyps was not reported so these data were not extracted.

As already stated in the NBI section, Lee and colleagues[77] stated that participants did not experience any procedure-related complications. The other four i-scan studies[79-82] made no reports of adverse events.

The length of time taken to perform the withdrawal phase of the colonoscopy was not reported in any of the studies. Basford and colleagues[79] however, commented that the in vivo assessment was performed in the time between finding a polyp and preparing for polypectomy and so did not cause a significant delay. Hoffman and colleagues,[80] who examined only the last 30cm of colon reported that with surface enhancement with i-scan the total examination time was 5 minutes.

FICE

None of the studies reported on the interpretability of test, inter-observer agreement, intra-observer agreement, test acceptability (to patients and/or clinicians), adverse events, number of polyps designated to be left in place, number of polyps designated to be resected and discarded, number of polyps designated for resection and histopathological examination, length of time to perform the colonoscopy, number of outpatient appointments, health-related quality of life, colorectal cancer or mortality.

Head-to-head comparisons

Head-to-head comparisons of NBI, i-scan and FICE were not within the scope of this assessment, nevertheless two studies met the inclusion criteria for the systematic review which did compare two technologies against each other. When NBI was compared to i-scan in a prospective cohort study of the real-time histological prediction of diminutive colonic polyps, Lee and colleagues[77] found no statistically significant differences between the two technologies (NBI vs i-scan: sensitivity, 88.8% vs 94.6%; specificity, 86.8% vs 86.4%; accuracy, 87.8% vs 90.7%, respectively; $P > 0.05$). In the RCT that compared NBI to FICE, Kang and colleagues[78] found that for polyps <5mm in size there was no statistically significant difference ($P>0.05$) in accuracy (74.9% vs 80.1%, respectively) or sensitivity (81.9% vs 74.5 %) but there was a statistically significant difference in specificity (75.7% vs. 88.4 %, $P = 0.006$). The authors concluded that better results should be achieved for both technologies before either are used for real-time optical biopsy of colorectal polyps in colorectal screening of the general population.[78] It is worth noting that in the study by Lee and colleagues[77] a single endoscopist with experienced of both NBI

and i-scan undertook the study colonoscopies whereas the four endoscopists in the Kang and colleagues' study[78] had no prior experience of either NBI or FICE.

4.1.5 Summary of diagnostic test performance evidence

- Thirty studies met the inclusion criteria for the systematic review of test accuracy. These assessed NBI (24 studies), i-scan (5 studies) and FICE (3 studies). Two of the included studies assessed two of the included interventions (NBI and i-scan; NBI and FICE). The way studies reported test accuracy outcomes (in terms of the region of the colon and the level of confidence assigned to the polyp characterisation) varied.
- Most studies enrolled participants from more than one of the populations eligible for inclusion in this review (receiving colonoscopy for screening, surveillance, or symptoms) but these studies did not report results separately for each participant type.
- The included studies were judged as likely to be at a low risk of bias.

NBI

- 23 studies reported either sensitivity (1 study) or both sensitivity and specificity (22 studies).
- In the whole colon, characterisations of diminutive polyps made with any level of confidence had a sensitivity ranging from 0.55 to 0.97 (17 studies) and a specificity ranging from 0.62 to 0.95 (16 studies). A bivariate meta-analysis (15 studies) produced a summary sensitivity value of 0.87 (95% CI 0.82 to 0.91) and specificity of 0.81 (95% CI 0.75 to 0.85). For characterisations in the whole colon made with high confidence summary sensitivity and specificity (11 studies) was slightly higher: sensitivity 0.91 (95% CI 0.85 to 0.95) and for specificity of 0.82 (95% CI 0.76 to 0.87) and limiting this analysis to studies where the endoscopists were experienced in the use of NBI (4 studies) did not greatly alter these results [sensitivity 0.92 (95% CI 0.89 to 0.94); specificity 0.82 (95% CI 0.72 to 0.89)].
- In the rectosigmoid colon, characterisations of diminutive polyps made with any level of confidence (four studies) had a sensitivity ranging from 0.84 to 0.90 and a specificity ranging from 0.76 to 0.95. A bivariate meta-analysis (3 studies) produced a summary estimate for sensitivity of 0.85 (95% CI 0.75 to 0.91) and for specificity of 0.87 (95% CI 0.74 to 0.94). For characterisations in the rectosigmoid colon made with high confidence (5 studies) sensitivity ranged from 0.83 to 0.96 and specificity from 88% to 99%. A bivariate meta-analysis (4 studies) produced a summary estimate for sensitivity of 0.87 (95% CI 0.80, 0.92) and for specificity of 0.95 (95% CI 0.87, 0.98). Limiting the analysis of high confidence characterisations in the rectosigmoid colon to the two studies where the endoscopists were experienced in the use of NBI

increased the summary values for sensitivity and specificity [sensitivity: 0.90 (95% CI 0.71 to 0.97); specificity 0.98 (95% CI 0.91 to 1.00)].

- Some studies that reported sensitivity and specificity were not included in meta-analysis because it was not possible to impute the required 2x2 table data. In two of three instances where this occurred, the sensitivity and specificity reported by the absent study lay within the 95% CI of the summary estimates of the meta-analysis. In one case (the meta-analysis of high confidence polyp characterisations in the whole colon) the missing study, Ladabaum and colleagues,[63] reported a sensitivity that lay within the 95% CI of the summary estimate but a specificity that lay outside the 95% CI of the summary estimate.
- The NPV of NBI for the characterisation of diminutive polyps in the whole colon (made with any level of confidence) ranges from 43% to 96% (16 studies). Five studies reported NPV values of 90% or more but the lower limit of the 95% confidence interval fell below 90% in every study except one. When limited to high confidence characterisations NPV ranged from 48% to 98% (13 studies) with five studies reporting NPV values of 90% or more. However, the lower limit of the 95% CI remained above 90% in only two studies.
- The NPV of NBI for the characterisation of diminutive polyps in the rectosigmoid colon (made with any level of confidence) ranged from 87% to 98% and was over 90% in four out of the five studies that reported this outcome (but the lower limit of the 95% CI remained above 90% in only two studies). When limited to high confidence characterisations in the rectosigmoid colon (five studies), NPV ranged from 92% to 99% but the lower limit of the 95% CI fell below 90% in 2 studies.
- Accuracy (the proportion of correctly classified polyps) of polyp characterisations in the whole colon was 90% or more in five studies and lay between 76% and 89% in 10 studies (16 studies reported this outcome). High confidence characterisations typically increased accuracy by 3-5% in studies reporting both overall and high confidence data (8 studies).
- Agreement between the surveillance interval allocated using an NBI based strategy and using the results of histopathology was above 90% in eleven of the 13 studies that reported this outcome. When there was a discrepancy in surveillance intervals, the NBI containing strategy more often led to an earlier recall for colonoscopy than would have occurred with the histopathology based surveillance interval.
- No outcome data were reported (Interpretability of the test; test acceptability, number of outpatients appointments, health-related quality of life, colorectal cancer, or mortality) or sparse outcome data (inter-observer agreement, adverse events, polyps designated as ‘left in place’,

polyps designated ‘resect and discard’, time taken to perform colonoscopy) were reported for other outcomes of interest to this review

i-scan

- Five studies provided sensitivity and specificity outcomes for the characterisation of diminutive polyps as adenomas or hyperplastic polyps using i-scan. Often only a single study provided data for a particular combination of the region of the colon and the level of confidence assigned to the polyp characterisation.
- In the whole colon or in regions of the colon characterisations of diminutive polyps made with any level of confidence ranged in sensitivity from 0.82 to 0.95 and in specificity from 0.83 to 0.96. It was not possible to meta-analyse any of these results. For high confidence characterisations in the whole colon or in regions of the colon sensitivity ranged from 94% to 98% and specificity from 90% to 95%. The only meta-analysis possible, which was conducted to inform the economic model, was for high confidence characterisations of diminutive polyps in the whole colon. The summary value for sensitivity was 0.96 (95% CI 0.92 to 0.98) and for specificity was 0.91 (95% CI 0.84 to 0.95).
- NPV values were above 90% (all 5 studies) however the lower limit 95% confidence interval was above 90% in only one study.
- Accuracy was 90% or more (all 5 studies) and higher for high confidence polyp characterisations in the two studies that also reported accuracy for all polyp characterisations.
- Surveillance interval agreement (2 studies) determined by i-scan and histopathology was over 90%. Where surveillance intervals differed, longer intervals were more likely to be set with i-scan than histopathology.
- No outcome data were reported (Interpretability of the test; test acceptability, polyps designated as ‘left in place’, polyps designated ‘resect and discard’, number of outpatients appointments, health-related quality of life, colorectal cancer, or mortality) or sparse outcome data (inter-observer agreement, adverse events, time taken to perform colonoscopy) were reported for other outcomes of interest to this review

FICE

- Three studies provided sensitivity and specificity with all reporting on characterisations of diminutive polyps made with any level of confidence in the whole colon. Reported values for sensitivity range from 74% to 88% and for specificity from 82% to 88%.
- None of the studies provided evidence on the high confidence characterisation of diminutive polyps or restricted their analysis to a part of the colon e.g. the rectosigmoid colon.

- It was possible to run a bivariate meta-analysis that produced a summary estimate for sensitivity of 0.81 (95% CI 0.73 to 0.88) and for specificity of 0.85 (95% CI 0.79 to 0.90).
- The NPV of FICE (3 studies) ranged from 70% to 84%.
- The accuracy of FICE (3 studies) ranged from 80% to 85%
- Surveillance interval agreement between FICE and histopathology was 100% (1 study) or 97% (1 study). When there was disagreement it was not reported whether the FICE based strategy led to a longer or a shorter surveillance interval being set.
- None of the other outcomes of interest to this review were reported.

Pooled analysis of virtual chromoendoscopy technologies

- A pooled analysis of high confidence decisions characterising diminutive polyps in the whole colon (11 NBI, 2 i-scan studies) was undertaken to inform a scenario analysis using the economic model. This produced a pooled summary estimate for sensitivity of 0.92 (95% CI 0.87 to 0.95) and for specificity of 0.83 (95% CI 0.78 to 0.87).

Head-to-head comparisons

- Head-to-head comparisons of the technologies were not within the scope for this assessment, but two included studies compare two technologies against each other. For the real-time histological prediction of diminutive colonic polyps no statistically significant differences were found when a single endoscopist with experience of NBI and i-scan compared these technologies in a prospective cohort study. An RCT conducted by endoscopists without experience of either NBI to FICE found no statistically significant difference in accuracy or sensitivity but a statistically significant difference in specificity.

Table 24 provides a summary of the pooled sensitivity and specificity values from our bivariate meta-analysis, where available.

Table 24 Summary of bivariate meta-analysis results

Type of characterisation	**Diminutive polyp location**	**NBI**		**i-scan**		**FICE**	
		Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
All characterisations [a]	whole colon	0.88 (0.83 to 0.92) 16 studies	0.81 (0.75 to 0.85) 16 studies	*0.95 (0.87 to 0.99) Single study*	*0.86 (0.76 to 0.94) Single study*	0.81 (0.73 to 0.88) 3 studies	0.85 (0.79 to 0.90) 3 studies
High confidence characterisations	whole colon	0.91 (0.85 to 0.95) 11 studies	0.82 (0.76 to 0.87) 11 studies	0.96 (0.92 to 0.98)[b] 2 studies	0.91 (0.84 to 0.95)[b] 2 studies	No evidence	No evidence
High confidence characterisations by endoscopists with prior experience of the technology [c]	whole colon	0.92 (0.89 to 0.94) 4 studies	0.82 (0.72 to 0.89) 4 studies	0.96 (0.92 to 0.98)[b] 2 studies	0.91 (0.84 to 0.95)[b] 2 studies	No evidence	No evidence
All characterisations [a]	rectosigmoid colon	0.85 (0.75 to 0.91) 3 studies	0.87 (0.74 to 0.94) 3 studies	Meta-analysis not possible 2 studies	Meta-analysis not possible 2 studies	No evidence	No evidence
High confidence characterisations	rectosigmoid colon	0.87 (0.80 to 0.92) 4 studies	0.95 (0.87 to 0.98) 4 studies	*0.96 (0.80 to 1.00) Single study*	*0.96 (0.83 to 0.99) Single study*	No evidence	No evidence

High confidence characterisations by endoscopists with prior experience of the technology [c]	rectosigmoid colon	0.90 (0.71 to 0.97) 2 studies	0.98 (0.91 to 1.00) 2 studies	No evidence	No evidence	No evidence	No evidence
		Pooled analysis of virtual chromoendoscopy technologies					
		Sensitivity (95% CI)			Specificity (95% CI)		
High confidence characterisations [c]	whole colon	0.92 (0.87 to 0.95) 11 NBI studies, 2 i-scan studies			0.83 (0.78 to 0.87) 11 NBI studies, 2 i-scan studies		

[a]All characterisations means not separated by endoscopist confidence level.

[b] The 'High confidence characterisations' result and the 'High confidence characterisations by endoscopists with prior experience of the technology' result are identical because the two studies contributing data to the high confidence meta-analysis were both undertaken by endoscopists with prior experience in using NBI.

[c] Post-hoc analysis

4.2 Ongoing studies

We identified 19 potentially relevant ongoing studies on the use of NBI, i-scan or FICE to characterise diminutive colorectal polyps. Two were identified from searches of clinical trials databases (see Section 3.1 for details of these searches) and 17 were identified from conference abstracts found by the clinical effectiveness searches. Until further details are available it is not clear whether all would meet the eligibility criteria for this review but they have the potential to do so. These studies are listed in Appendix 5.

5 ECONOMIC ANALYSIS

This section consists of a systematic review of published cost-effectiveness analyses of virtual chromoendoscopy compared to histopathology and a de novo economic evaluation.

5.1 Systematic review of existing cost-effectiveness evidence

This section describes the systematic review of published cost-effectiveness analyses of virtual chromoendoscopy. The aim of the systematic review was to inform the development of the independent economic evaluation. The same search strategy that was used to identify diagnostic test studies was used to identify cost-effectiveness studies, as described in Section 3. Once the results of this search had been downloaded into our Endnote (X7.0.2, Thomson Reuters) bibliographic database we searched for a subset of relevant cost-effectiveness studies using the keyword 'cost' in any field (NB. The search strategy for our systematic review of diagnostic accuracy did use a study design filter, therefore it would not have excluded any relevant health economic studies). Titles and abstracts were then screened by two health economists for relevance according to the inclusion criteria. The inclusion criteria were for a full economic evaluation (cost-effectiveness, cost-utility, cost benefit or cost consequence analysis) that compared virtual chromoendoscopy with conventional (white light) colonoscopy for adults undergoing a colonoscopy for detection of colorectal polyps, that included long-term outcomes (such as life years, incidence of colorectal cancer or QALYs). Full texts of references deemed relevant were then retrieved for further screening. The full texts of retrieved references were screened to identify those that met the inclusion criteria. Data from the included studies were extracted and evaluated for their quality and generalisability to the UK, based upon criteria developed by Drummond and colleagues.[110] The studies identified are described in more detail in the following section.

A total of 236 potentially relevant references from our database underwent title and abstract screening. Of these, the full text versions (where available) of ten references were retrieved for screening, and two of these met the inclusion criteria.[111,112] The characteritics of these studies are given in Table 25. Of the eight texts not included, four were abstracts with insufficient detail[113-115 51] and four did not include long-term outcomes in their analysis[55,59,84,116] (Appendix 6). The full data extraction forms for both of the included studies are shown in Appendix 7.

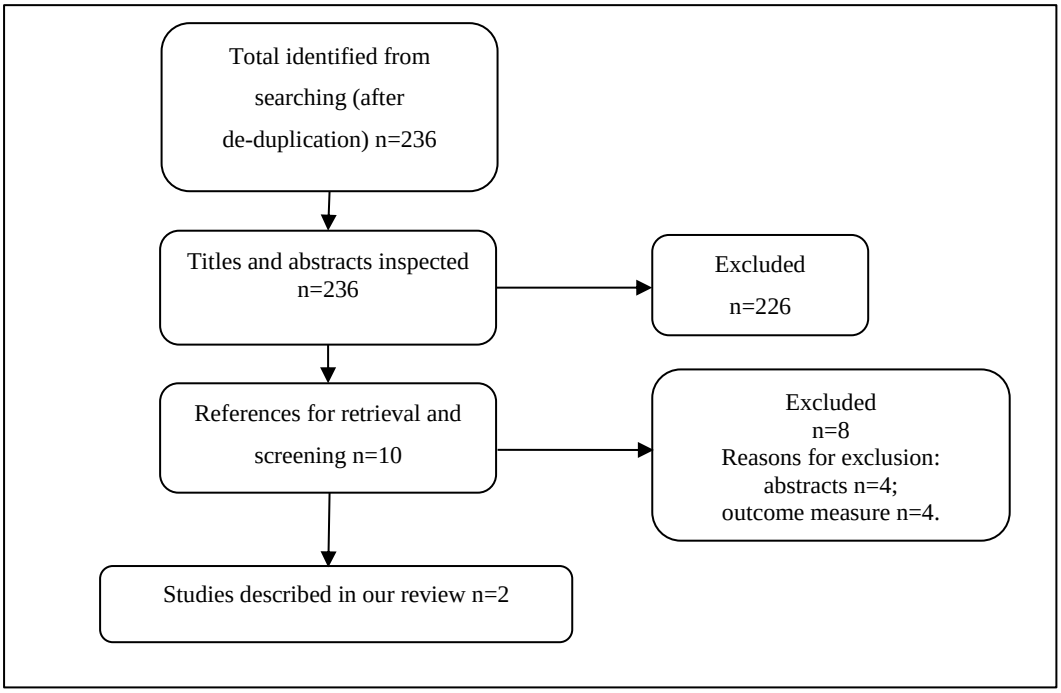


Figure 29 Flow chart of identification of studies for inclusion in the review of cost-effectiveness

Table 25 Characteristics of included economic evaluations

Author	**Hassan et al.**[111]	**Kessler et al.**[112]
Publication Year	2010	2011
Country	USA	USA
Funding source	Funding source not reported.	National Institutes of Health grant
Study type	Cost-effectiveness analysis	Cost-effectiveness analysis
Perspective	Societal	Not stated (assumed to be payer)
Study population	Hypothetical cohort of 100,000 50 year old persons in United States who underwent a colonoscopy for CRC screening.	Patients receiving a colonoscopy at a single-institution tertiary centre who had at least one polyp removed during colonoscopy, irrespective of indication. Population characteristics taken from a database of 10,060 consecutive colonoscopies from 1999 to 2004
Intervention(s)	NBI versus colonoscopy versus no screening	No pathological examination of diminutive polyps (resect and discard) vs. submitting all polyps for pathological examination (submit all)

Intervention effect	Feasibility of 84% for rate of high confidence in differentiating between hyperplastic and adenomatous diminutive polys by using NBI without magnification. Sensitivity was 94% and specificity was 89%.	Endoscopic sensitivity for non-adenoma 90%; Endoscopic sensitivity for adenoma 90%; Proportion of diminutive polyps with advanced histology 0.6%; Pathology sensitivity for large adenoma 100%; Pathology sensitivity for diminutive and small adenoma 95%; Pathology sensitivity for non-adenoma 100%.
Currency base	US dollars	US dollars
Model type, health states	Markov model with health states for: no colorectal neoplasia, diminutive (<= 5mm), small (6-9mm) or large (>=10 mm) adenomatous polyps; localised, regional, or distant CRC; and CRC related death.	Decision tree model
Time horizon	Lifetime horizon	Lifetime horizon
Base case results	Compared to standard colonoscopy, colonoscopy with NBI was $25 cheaper per person with no difference in life expectancy.	The net cost savings from forgoing histopathologic assessment was US$174.01. The expected increased benefit of the 'submit all' strategy was 0.17 days of life and the cost-effectiveness of the 'submit all' strategy compared to the 'resect and discard' strategy was US$377,460 per life year gained. The number needed to harm because of perforation, major bleed or missed cancer was 7979, i.e., an absolute risk of 0.0125%.

CRC - colorectal cancer

Critical appraisal of the studies

The assessment group critical appraisal of the identified studies by Hassan and colleagues[111] and Kessler and colleagues[112] are summarised in Table 26. Both studies report their methodology, assumptions and parameters clearly. Neither study included QALYs in their analysis and Kessler and colleagues did not include discounting. Hassan and colleagues did not present an incremental analysis, although it is possible to calculate this with the information provided.

Table 26 Critical appraisal checklist for economic evaluations (based on Drummond et al[110])

Item	**Hassan et al.**[111]	**Kessler et al.**[112]
1. Is the decision problem (including interventions compared and patient group) relevant to the UK?	Yes	Yes
2. Is the setting comparable to the UK?	Yes	Yes
3. Is the analytical and modelling methodology appropriate?	Yes	Yes
4. Are all the relevant costs and consequences for each alternative identified?	Yes	Yes
5. Are the data inputs for the model described and justified?	Yes	Yes
6. Are health outcomes measured in QALYs?	No	No
7. Is the time horizon considered appropriate?	Yes	Yes
8. Are costs and outcomes discounted?	Yes[a]	No
9. Is an incremental analysis performed?	?[b]	Yes
10. Is uncertainty assessed?	Yes	Yes
Comments [a] Discounted at 3% per annum, which differs from the National Institute for Health and Care Excellence reference case. [b] Both colonoscopy and resect and discard appear to have been compared to no screening but no ICERs were calculated		

Hassan and colleagues

Hassan and colleagues[111] developed a cost-effectiveness model to calculate the potential savings and drawbacks of a 'resect and discard' approach using NBI in a simulated colorectal cancer screening cohort. In the resect and discard approach, diminutive colorectal lesions (≤ 5mm), classified by endoscopy with high confidence, were not analysed by a pathologist. A Markov model was constructed with health states for no colorectal neoplasia, diminutive (<= 5mm), small (6-9mm) or large (>=10 mm) adenomatous polyps; localised, regional, or distant colorectal cancer; and colorectal

cancer-related death. The resect and discard policy was instituted for all the cases in which a high confidence diagnosis was achieved by NBI. All diminutive polyps in which a high confidence diagnosis was not possible were removed and sent for formal histologic evaluation. The model assumed a screening strategy of colonoscopy every 10 years. After colonoscopy, patients received follow-up surveillance based upon the size and classification of the polyp(s).

Feasibility and accuracy of NBI without optical magnification in differentiating between diminutive adenomas and hyperplastic polyps were derived from three published series.[59,69,71] Feasibility was defined as the rate of high confidence in differentiating between polyps. An 84% feasibility was assumed. The sensitivity and specificity for adenomas was 94% and 89%, respectively.

Costs were derived from Medicare reimbursement rates. No incremental cost for NBI was included because it was stated to be a standard feature in current generation colonoscopes. The cost of colonoscopy was $630, the cost of colonoscopy with polypectomy was $925, and pathologic examination was $102. Costs were also included for colorectal cancer treatment and adverse event costs, such as perforation and bleeding. Costs and life years were discounted at 3% per annum.

The discounted costs for the no screening strategy were $3390 per person over their lifetime (Table 27). The colonoscopy screening strategy reduced costs by $168 per person and the colonoscopy with resect and discard strategy reduced costs by a further $25 per person. Colonoscopy with or without resect and discard improved life expectancy by an average of 51 days per person compared with no screening. The study also extrapolated the results to the US population.

Table 27 Cost and efficacy for the screening strategies of Hassan and colleagues

	No screening	**Colonoscopy**	**Colonoscopy with resect and discard**
Cost/person	$3390	$3222	$3197
Relative efficacy	-	51 days / person	51 days / person

Kessler and colleagues

Kessler and colleagues[112] developed a decision tree model to quantify the expected costs and outcomes of removing diminutive polyps with or without subsequent pathologic assessment. They compared two strategies: 'submit all' diminutive polyps (≤ 5 mm in size) to pathological examination compared to no pathological examination of diminutive polyps ('resect and discard'). All other polyps were submitted for pathological examination for both strategies.

The decision model was populated with polyp frequencies based on a database of 10,060 consecutive patients who underwent colonoscopy for screening, surveillance or diagnostic indications. The decision model evaluated the frequency with which the surveillance follow-up (based on the most advanced polyp) matched that of the actual follow-up interval for the two strategies. Patients in the endoscopy database were distributed amongst four groups based on the characteristics that form the basis for follow-up. Group one consisted of people who had only one diminutive polyp. Group two had people who had two polyps, at least one of which was diminutive and the other not a large adenoma (⩾10mm). In group three, people had at least three polyps at least one of which was diminutive and the others were not large adenomas. In group four people had at least one diminutive polyp, as well as one or more large adenoma(s) and could have any number of additional polyps. For each of the four groups, each patient's most advanced polyp was either an advanced adenoma, a non-advanced adenoma or a non-adenoma.

The sensitivity and specificity of endoscopic and pathology assessment were based on the published literature. Costs were included for pathology, colonoscopy and colorectal cancer treatment. The cost of sending a polyp to pathology was US$103.87. Costs of colonoscopy, colonoscopic perforation and cancer treatment were obtained from the literature. The colonoscopy costs were US$1329 for diagnostic and US$2038 for therapeutic colonoscopies. The downstream costs and outcomes after the colonoscopy were obtained from a published discrete event simulation model of colorectal cancer screening and surveillance intervals.[117] Discounting was not included in the model.

The submit-all strategy resulted in an incorrect surveillance interval 1.9% of the time, while the resect and discard strategy did so 11.8% of the time, with over half of the patients having only non-adenomatous polyps but scheduled for a 5 year, rather than a 10 year surveillance examination. The cost savings from forgoing pathologic assessment were US$210 per colonoscopy when diminutive polyps were removed, while the additional cost due to the incorrect surveillance interval was US$35.92. The net saving was US$174.01. The number needed to harm because of perforation, major bleed or missed cancer was 7979, i.e., an absolute risk of 0.0125%.

The expected additional benefit of the submit-all strategy was 0.17 days of life over the lifetime horizon and the incremental cost-effectiveness ratio (ICER) of the submit-all strategy compared to the resect and discard strategy was US$377,460 per life year gained.

Deterministic sensitivity analyses were conducted for the accuracy of the colonoscopy to detect adenomas and the proportion of diminutive polyps with advanced histology. The sensitivity analyses performed indicate that the error rate in assigning post-polypectomy surveillance intervals was most

sensitive to the accuracy of endoscopic assessment of histology and to the proportion of diminutive polyps with advanced histology.

The authors concluded that endoscopic diagnosis of polyp histology during colonoscopy and forgoing pathologic examination would result in substantial upfront cost savings. Downstream consequences of the resulting incorrect surveillance intervals appear to be negligible.

Summary of published economic evaluations

The cost-effectiveness review of published economic evaluation for virtual chromendoscopy technologies found two relevant studies that were both published in the USA.[111,112] The patient population differed between the two studies, Hassan and colleagues simulated a screening population (i.e. included patients who had no polyps identified by the colonoscopy) and Kessler and colleagues' population had at least one diminutive polyp identified. Both studies compared a 'resect and discard' strategy to a 'submit all' (to histopathology) strategy to the whole colon, although Kessler and colleagues[112] assumed that the resect and discard strategy would be used for all polyps, whilst Hassan and colleagues[112] assumed that for some polyps it would not be feasible to use resect and discard (i.e. those characterised with low confidence). Neither studies used surveillance intervals for follow-up screening that correspond to those used in the UK.

The model structure differed between the two studies, Hassan and colleagues[112] used a Markov model and Kessler and colleagues[112] used a decision tree model. We consider that both approaches are appropriate. The cost saved per person varied between US$25[111] and US$174 over the patient lifetime.[112] The expected benefit of histopathology was 0.17 days of life in Kessler whilst Hassan assumed there was no difference in life expectancy between groups over the patient lifetime. The cost-effectiveness of the submit all strategy compared to resect and discard varied was US$377,460 per life year gained for Kessler and colleagues, whilst Hassan and colleagues were not able to calculate a value as there was no difference in the life expectancy between the submit all and resect and discard strategy. It is unclear how generalisable these results are to UK NHS as they have used non-UK resource costs and have not included QALYs.

Review of information provided by Olympus to the National Institute for Health and Care Excellence: economic evaluation

A budget impact model was supplied as part of the information provided by Olympus to the National Institute for Health and Care Excellence and the assessment group. The model has also recently been published by Solon and colleagues.[116] This study did not meet our inclusion criteria for cost-effectiveness models of virtual chromoendoscopy because it did not include long-term health

outcomes. However, we have provided a critical review of it as a supplement to our systematic review of cost-effectiveness studies as it has some relevance to the decision problem in this assessment.

Modelling approach

The analysis is a cost consequence and budget impact model that follows cohorts of UK patients who attend colorectal cancer screening. The population includes patients identified through the national screening programme as well as those attending for colonoscopic surveillance. The analysis is conducted from the perspective of the NHS in England. The model has a time horizon of seven years and in each year there is a new incident cohort of patients that undergo an endoscopy. The model includes a discount rate of 3.5% per year for costs and health outcomes. The model has a starting population of 550,925 attending an endoscopy test per year, to reflect the number of procedures conducted in 2012, and assumes an annual increase of 20% in the population expected to attend endoscopy each year. It was assumed that 82% of the installed endoscopy systems in England were manufactured by Olympus.

After undergoing endoscopy, patients are classified in three outcomes according to the number and size of polyps identified (no polyps; one of more polyps ≤5 mm but no polyps >5 mm; one or more polyps ≥ 5mm). For white light endoscopy (WLE), all polyps are resected and sent for histopathological examination. With NBI, for polyps ≤5 mm, the diagnosis of a proportion of polyps is assumed to be predicted with low confidence and they are sent for histological examination, whilst polyps will be left in situ if there is high confidence that they are non-neoplastic, otherwise they will be resected and discarded. Where polyps are resected, there is a risk of adverse events of bleeding and bowel perforation. The model calculates the number of true negatives, false negatives, true positives and false positives and the number of histological examinations, resects and adverse events for each cohort in each year.

Critical appraisal of the model

The assessment group critical appraisal of the Olympus economic model is summarised in Table 28. In general, the model is well reported although some aspects were reported in the economic model provided by Olympus (Appendix 8), rather than in Solon and colleagues.[116] The time horizon is seven years but consists of seven yearly cohorts and no longer-term outcomes, such as QALYs, were modelled.

Table 28 Critical appraisal checklist of economic evaluation (Questions in this checklist based on Drummond et al.[110] and the National Institute for Health and Care Excellence reference case[118]

	Item	
1	Is there a clear statement of the decision problem?	Yes
2	Is the comparator routinely used in UK NHS?	Yes
3	Is the patient group in the study similar to those of interest in UK NHS?	Yes
4	Is the health care system comparable to UK?	Yes
5	Is the setting comparable to the UK?	Yes
6	Is the perspective of the model clearly stated?	Yes
7	Is the study type appropriate?	Yes
8	Is the modelling methodology appropriate?	Unclear
9	Is the model structure described and does it reflect the disease process?	Yes
10	Are assumptions about model structure listed and justified?	Yes
11	Are the data inputs for the model described and justified?	Yes
12	Is the effectiveness of the intervention established based on a systematic review?	Yes
13	Are health benefits measured in QALYs?	No
14	Are health benefits measured using a standardised and validated generic instrument?	No
15	Are the resource costs described and justified?	Yes
16	Have the costs and outcomes been discounted?	Yes
17	Has uncertainty been assessed?	Yes
18	Has the model been validated?	No

Clinical effectiveness

The model parameters for the diagnostic accuracy of NBI, the feasibility of diagnosing diminutive polyps and adverse events were derived from a systematic literature review and are shown in Table 29.

Table 29 Effectiveness parameters used in the Olympus economic model

Parameter	Value	Source
Patients with no polyps, %	44%	Rastogi et al. (2012)[119]
Patients with polyps ≤ 5mm, %	38%	Rastogi et al. (2012)[119]
Patients with polyps > 5mm, %	18%	Rastogi et al. (2012)[119]
Polyps that are adenomatous ≤ 5mm, %	17%	Butterly et al. (2006)[120]
Polyps that are adenomatous > 5mm, %	10.1%	Butterly et al. (2006)[120]
Diminutive polyp optical diagnosis feasibility rate	75%	Kaltenbach et al. (2014)[30]
Optical diagnosis sensitivity NBI	93%	McGill et al.(2013)[42]
Optical diagnosis specificity NBI	83%	McGill et al.(2013)[42]
Probability of hospitalisation for bleeding with polypectomy	0.43%	Whyte et al. (2012)[121]
Probability of perforation with polypectomy	0.28%	Whyte et al. (2012)[121]

Estimation of costs

The model includes the costs incurred by the NHS, including equipment, maintenance, training, consumables, staff, endoscopy and histological examination costs and hospital costs for managing adverse events. Unit costs of resources were taken from a variety of sources including NHS Reference costs,[122] PSSRU,[123] and the company's own prices. The costs used in the model are shown in Table 30.

The company's list price for NBI systems is £40,395. The model assumes that at the start of the first year 82% of hospitals currently use Olympus systems, of which 95% are capable of NBI, i.e. 78% of hospitals use NBI. Of those hospitals with Olympus equipment, 50% of hospitals that do not have NBI capable systems will upgrade in year one and a similar number in each subsequent year. Of those hospitals with Olympus equipment, 50% have NBI-capable endoscopes in place in the first year. Of those hospitals with Olympus equipment, that do not have NBI-capable endoscopes, 14% will upgrade in year one and a similar number will upgrade in each subsequent year. For NBI, two training days per endoscopist per year are required, while no additional training is required for WLE.

Staff costs for colonoscopy include costs for administration, nurse and consultant contact time and are based upon a micro-costing study of a Canadian hospital.[124] The consumables for biopsy are snares and forceps. The assessment group notes that consumables and staff costs would normally be included within the NHS Reference costs and do not therefore need to be included separately.

Table 30 Cost parameters used in the Olympus economic model

Input parameter	Value	Source
Unit cost per system NBI	£40,395	Olympus list price
Unit cost per scope NBI	£38,660	Olympus list price
Training cost per year NBI	£2,272	Olympus list price
Maintenance cost NBI system	£3,525	Olympus list price
Maintenance cost NBI scopes	£4,805	Olympus list price
NHS Tariff for colonoscopy - with biopsy	£522	Monitor 2014 - HRG tariff FZ51Z
NHS Tariff for colonoscopy - without biopsy	£437	Monitor 2014 - HRG tariff FZ52Z
Cost per biopsy	£82	Unpublished data obtained from University College London Hospitals, Plymouth Hospital NHS Trust and South Devon Healthcare NHS Foundation Trust
Number of biopsies per exam	1.35	Assumption based on data reported in Lee et al, 2012
Cost per hospital bleed	£318	Monitor 2015-6 - HRG tariff FZ38F
Cost per perforation event	£2,211	Monitor 2015-6 - HRG tariff GB01B
Unit cost per hour for administration & support	£23	PSSRU 2014
Hours per test for administration & support	0.30	Modified from assumptions reported in Sharara et al. 2008[124]
Unit cost per hour nurse non-contact time	£41	PSSRU 2014
Hours per test for nurse non-contact time	0.42	Modified from assumptions reported in Sharara et al. 2008[124]
Unit cost per hour of consultant time	£142	PSSRU 2014
Hours with consultant, excluding procedure	0.50	Modified from assumptions reported in Sharara et al. 2008[124]
Length of procedure time in hours with NBI	0.30	Bisschops et al. 2012[125]
Length of procedure time in hours with comparator	0.30	This input varies where options are selected
Unit cost per hour nurse contact time	£100	PSSRU 2014
Snares - cost per pack	£240	Olympus list price

Input parameter	Value	Source
Snares - number per pack	20	Market data provided by Olympus
Forceps - cost per pack	£210	Olympus list price
Forceps - number per pack	10	Market data provided by Olympus

Results

The results for the outcomes from the model are shown in Table 31. Over seven years NBI reduced the incidence of colonoscopy-related adverse events by 32% and the frequency of histopathological examination by 39%.

Table 31 Outcomes from the Olympus economic model

Outcome	NBI	WLE	% Change
True negatives	5,713,178	5,933,416	-3.71%
False negatives	1,596	-	N/A
True positives	148,296	149,893	-1.07%
False positives	220,238	-	N/A
Histopathology exams	2,065,058	3,406,653	-39.38%
Adverse events	16,376	24,187	-32.29%

The cost over seven years for NBI is predicted to be £3,112 million and for WLE is £3,253 million, i.e. a saving of £141 million.

Deterministic sensitivity analyses were included in the model by varying the model parameters by +/- 10%. The sensitivity analysis shows the effect of the parameters on the total difference in costs between NBI and WLE. The costs of colonoscopy with and without biopsy have the greatest impact on model results. The study also conducted an analysis reducing the cost of biopsy, which showed there was still a net cost saving with NBI even when the biopsy cost was reduced to zero.

5.2 Independent economic evaluation

As described in Section 2, the decision problem for this diagnostic assessment is to assess the cost-effectiveness of real-time optical assessment of diminutive colorectal polyps in the English NHS. We therefore conducted an economic evaluation to evaluate costs and outcomes of virtual chromoendoscopy. The economic evaluation takes the form of a cost-utility model informed by the systematic review of cost-effectiveness studies, the economic evaluation by Olympus, targeted

literature searches and, where necessary, expert opinion. The economic evaluation uses the diagnostic accuracy for virtual chromoendoscopy from the meta-analyses reported in section 4.

5.3 Methods for economic analysis

5.3.1 The decision problem

The patient population in our base case analysis is people referred to colonoscopy after participating in a bowel cancer screening programme (referred to as the **screening population**). We included in scenario analyses two other patient populations of relevance to the decision problem for this assessment: people offered colonoscopic surveillance because they had previously had adenomas removed (**surveillance population**); and people referred for colonoscopy by a GP because of symptoms suggestive of colorectal cancer (**symptomatic population**).

For the purposes of the economic analysis, we only include patients with at least one diminutive polyp and exclude patients with one or more non-diminutive polyp. The scope for this assessment excludes use of virtual chromoendoscopy for real-time assessment of non-diminutive polyps (>5mm), though VCE might be considered for use in the assessment of diminutive polyps in patients who also have non-diminutive polyps. In practice, patients do have a mixture of polyps of different sizes. Although most polyps are diminituive, patients are assigned to surveillance intervals according to their most advanced polyp. However, we could not identify data on the mix of different sized polyps in patients or how they affect the allocation to surveillance interval. Additionally, all data in the model on adenoma and cancer risk is based on data that averages risk across adenoma sizes.

Further, the model does not differentiate between the type of polyp, such as depressed polyps or sessile serrated polyps. Sessile serrated polyps are rare and no diagnostic accuracy data were available for diminutive sessile serrated polyps from our systematic review of diagnostic studies (Section 4).

For the base case analysis in our economic evaluation, we compare strategies using virtual chromoendoscopy technologies (NBI, i-scan and FICE) with a histopathology assessment strategy. For the comparator **histopathology strategy**, we assume that all polyps are resected and sent for histopathology, and that subsequent screening and surveillence invitations are based on the histopathology results, which are assumed to be 100% accurate.

We refer to the virtual chromoendoscopy strategy used in our base case analysis as the **VC strategy**; it has the following characteristics:

- Diminutive polyps in the whole colon are optically characterised using virtual chromoendoscopy
- Diminutive polyps characterised with high confidence as adenomas are resected and discarded
- Diminutive polyps characterised with high confidence as hyperplastic are left in situ
- Any polyps that cannot be characterised with high confidence are resected and sent to histopathology

The VC strategy is based upon the Detect InSpect ChAracterise Resect and Discard (DISCARD) strategy described in Ignjatovic and colleagues[59] and then subsequently adapted in the two economic models identified by our systematic review of economic evaluations.[111,112] Ignjatovic and colleagues' study[59] was one of the first to evaluate a resect and discard strategy, and they proposed that polyps <10mm in size should be characterised, and if appropriate be discarded or left in situ without histopathology. Subsequent studies and guidance have modified the DISCARD strategy to apply to only diminutive polyps (≤5 mm). The NICE scope, ESGE guidelines,[30] both economic evaluations identified through our systematic review, and our model limit the VC strategy to diminutive polyps.

Our VC strategy does differ from the DISCARD strategy in the way that hyperplastic polyps are dealt with in the proximal colon (see Figure 3 on page 35 above). In the base case analysis, the model does not differentiate between diminutive hyperplastic polyps found in the rectosigmoid colon or other parts of the colon, because the best available diagnostic data from our systematic review was based on polyps in the whole colon. However, we have conducted scenario analyses (section 5.5.2) using what we refer to as the **DISCARD strategy**, which has the following characteristics:

- Any polyp assessed with low confidence is resected and sent to histopathology
- Diminutive polyps in the whole colon characterised with high confidence as adenomas are resected and discarded
- Diminutive polyps in the proximal colon characterised with high confidence as hyperplastic are resected and discarded
- Diminutive polyps in the rectosigmoid colon characterised with high confidence as hyperplastic are left in situ

We assessed each of the virtual chromoendoscopy based strategies (VC and DISCARD) for each of the three technologies (NBI, i-scan and FICE). In addition, we conducted a scenario analysis using the post hoc pooled meta-analysis sensitivity and specificity estimates for the virtual chromoendoscopy technologies (Section 4.1.2).

Following colonoscopy and receipt of histopathology results, patients are assigned a surveillance interval based on their estimated level of risk (see Figure 30). The risk classification of patients used corresponds to British guidelines[28] for determining surveillance intervals following identification of exclusively diminutive adenomas at colonoscopy: low risk (0-2 adenomas); intermediate risk (3-4 adenomas), and high risk (5 or more adenomas).

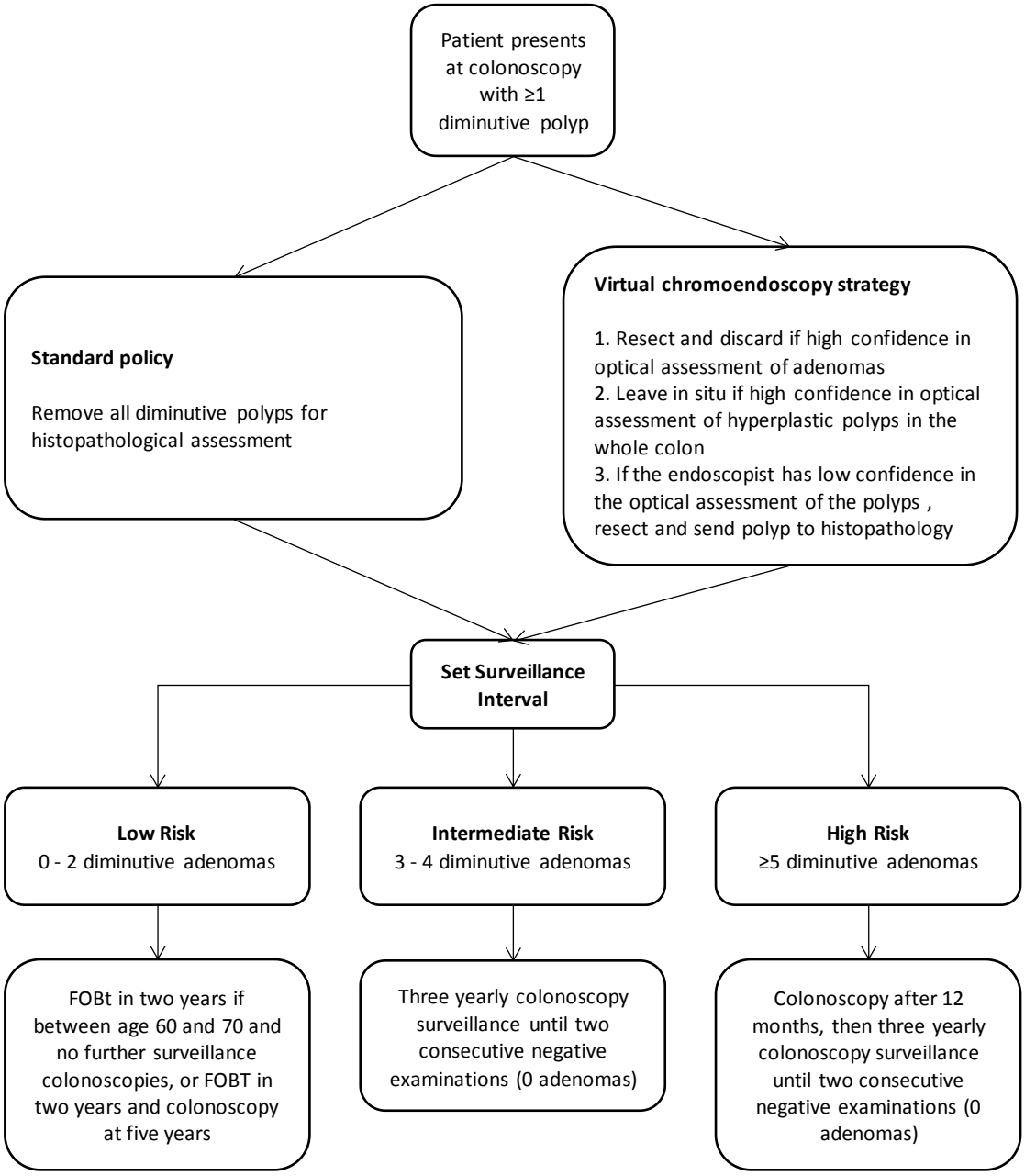


Figure adapted from 'Public health functions to be exercised by NHS England: Service specification No.26, Bowel Cancer Screening Programme'[126]

Figure 30 NHS Bowel Cancer Screening Pathway (with endoscopy policies)

There are four main implications of using a virtual chromoendoscopy strategy (VC or DISCARD) rather than the histopathology strategy:

1. **Initial costs:** Most hospitals already have equipment capable of VCE. There would be additional training costs for endoscopists to use this technology, but conversely the cost of polypectomies and histopathology tests would be reduced. Thus the net effect on the cost of initial diagnosis and management (colonoscopy, polypectomy and histopathology) may be positive or negative.
2. **Hyperplastic polyps resected**: The number of hyperplastic polyps unnecessarily resected and hence the numbers of polypectomy-related adverse events, such as bleeding and bowel perforation, will be reduced. Some hyperplastic polyps will still be resected, because they are not assessed with high confidence or are mischaracterised as adenomas (false positives). Adverse events are associated with a mortality risk, reduced quality of life and costs to the health service.
3. **Missed adenomas**: However, some polyps will be mischaracterised as hyperplastic when they are adenomas (false negatives). Such errors will mean that some adenomas will be left in situ, leading to a small increase in the incidence of colorectal cancer, with associated QALY loss and healthcare costs.
4. **Incorrect follow up**: Some patients may be assigned to the wrong follow up interval (according to the BCSP guidelines, Figure 30): either too long an interval if one or more adenomas are missed (false negatives); or too short an interval if one or more hyperplastic polyps are characterised as adenomas (false positives). In general, a shorter follow up interval will be beneficial for the patient, due to the reduced risk or earlier detection of cancer. However, for patients at very low risk of colorectal cancer, the potential harm from polypectomy-related adverse events could offset these benefits. The incremental cost to the health service of a shorter follow up interval may, in principle, be positive or negative: since increased costs of screening or surveillance may, to some extent, be offset by cost savings from avoided cancer treatment.

The model estimates the proportion of patients likely to experience these various risks, and hence the expected costs and QALYs associated with the alternative colonoscopy strategies.

5.3.2 Model structure

The model consists of a decision tree for patients undergoing colonoscopy. The tree estimates the short term costs and outcomes for the defined population under each decision strategy, from the time when patients are identified as potential candidates for use of virtual chromoendoscopy, up to the time when any polyps identified as adenomas have been removed and patients have been assigned to a follow-up policy. Long-term costs and QALY outcomes at the endpoints of the decision tree were estimated from an existing model: the School of Health and Related Research (ScHARR) Bowel Cancer Screening (SBCS) model, developed by Whyte and colleagues.[121] We chose to use the SBCS model, rather than to develop a new one, because it is a long-standing model, that has been validated, and which was used to inform the introduction of the national bowel cancer screening programme. The SBCS model was adapted for this current assessment, with updated parameters where possible. It was run independently, and the SBCS cost and QALY estimates for various subgroups of patients were entered as parameters at the endpoints of the decision tree model. The structures and assumptions of the decision tree and SCBS models are described below. Input parameters for both models are then discussed in Section 5.4.

5.3.2.1 The decision tree

The decision tree model compares the virtual chromoendoscopy strategies (VC with each of the teachnologies NBI, i-scan, FICE in the base case) with a histopathology strategy for a cohort of patients (the screening population in the base case). The model adopts a life time horizon and an NHS and personal and social services perspective.

Patients enter the model at colonoscopy, having had at least one diminutive polyp, and no non-diminutive polyps, identified. The cohort is divided into four risk categories, based on the number of adenomas that they have:

- No adenomas
- Low risk (LR): 1-2 adenomas
- Intermediate risk (IR): 3-4 adenomas
- High risk (HR): 5 or more adenomas

The model then calculates the proportion of patients in each category expected to have the correct diagnosis and treatment, and the proportions expected to be diagnosed and treated incorrectly. There are essentially three types of error that can occur: patients might have one or more hyperplastic polyp misclassified as an adenoma and unnecessarily resected; they may have one or more adenoma misclassified as a hyperplastic polyp and left in situ; and/or they may be assigned to an incorrect

surveillance interval – either too long or too short. The resulting permutations of diagnostic outcomes for patients are illustrated in Figure 31. It can be seen that there are six main patient outcomes, which are also defined in Table 32.

Table 32 Definitions of diagnostic outcomes for patients

Patient outcomes		**Interpretation**	**Surveillance interval assigned**
CD	Correct diagnosis	All polyps correctly classified (as either adenomas or hyperplastic polyps)	Correct
MAC	Missed adenoma(s) correct surveillance	One or more adenomas identified incorrectly as hyperplastic polyps and left in situ	Correct
MAI	Missed adenoma(s) incorrect surveillance	One or more adenomas identified incorrectly as hyperplastic polyps and left in situ	Incorrect – too long
HPRC	Hyperplastic polyp(s) resected correct surveillance	One or more hyperplastic polyps identified incorrectly as adenomas and resected	Correct
HPRI	Hyperplastic polyp(s) resected incorrect surveillance	One or more hyperplastic polyps identified incorrectly as adenomas and resected	Incorrect – too short
MAHPR	Missed adenoma(s) and hyperplastic polyp(s) resected	One or more hyperplastic polyps identified incorrectly as adenomas and resected and One or more adenomas identified incorrectly as hyperplastic polyps and left in situ	Correct [a]

[a] The probability that a patient who has both false positive and false negative test results is given the wrong surveillance interval is very small, as this would require a total of three or more errors (one false positive and two false negatives, or vice versa).

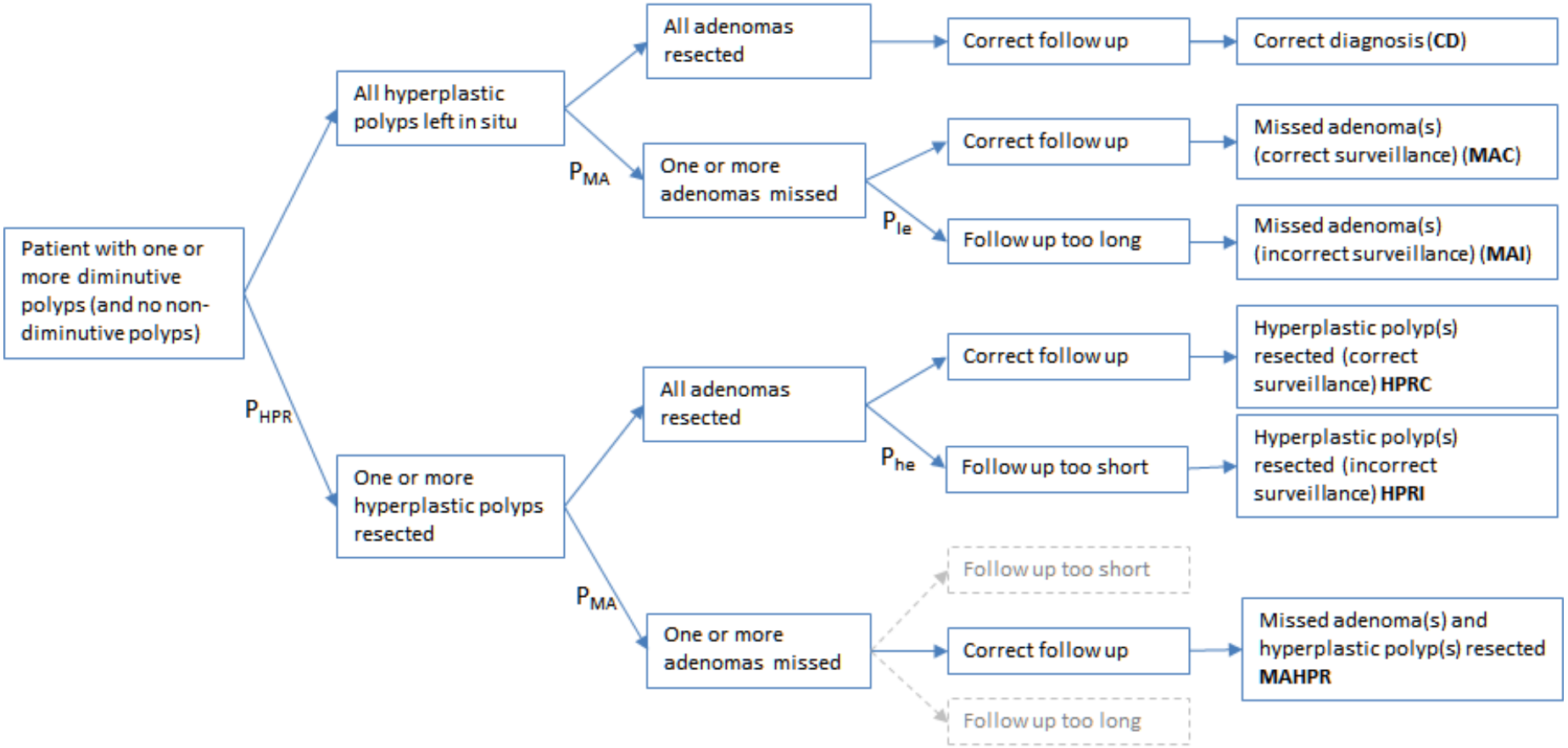


Figure 31 Decision tree showing diagnostic outcomes for patients

The probability of these different outcomes depends on the number of polyps and adenomas that the patient has, the diagnostic accuracy of the colonoscopy technology, and the policies for resecting polyps and assigning surveillance intervals.

In general, if the actual number of adenomas is at the higher end of the risk classification range, then if the patient has one or more hyperplastic polyps identified incorrectly as adenomas, they may be given a shorter surveillance interval than is appropriate. Similarly, if the actual number of adenomas is at the lower end of the risk classification range, then if the patient has one or more adenomas identified incorrectly as hyperplastic polyps and left in situ, they may be given a longer surveillance interval than is appropriate.

Some outcomes are not possible for particular groups of patients: for example, a patient with one hyperplastic polyp and one adenoma (LR) cannot be assigned an incorrect surveillance interval, since even if the hyperplastic polyp is mistaken for an adenoma, they would still be placed in the LR group and be invited (correctly) for routine screening. Other outcomes will be very improbable for some patients: for example, a patient with 9 adenomas (HR) is very unlikely to be diagnosed with less than 5 adenomas, and so is unlikely to be assigned to a surveillance interval that is too long.

It is possible that patients could simultaneously have one or more hyperplastic polyp misdiagnosed as an adenoma (FP) and one or more adenoma misdiagnosed as a hyperplastic polyp (FN). If so, the patient would be at risk of harm from the unnecessary resection(s) and increased risk of cancer due to the adenoma(s) left in situ. However, it is unlikely that they would be assigned to an incorrect surveillance interval, since the errors for individual polyps would be likely to cancel out.

The mathematics behind the estimation of outcome probabilities for patients from polyp-level diagnostic accuracy estimates is explained in section 5.3.2.2 below. But first we continue the overview of the decision tree model, and explain how it links to long-term outcomes from the SBCS model.

Under the histopathology strategy, all patients are assumed to receive the correct diagnosis (33). All polyps including adenomas are resected, so no adenomas are left in situ, and patients are assigned to the correct follow up strategy: routine invitation to screening for those with 0-2 adenomas, three-yearly surveillance for those with 3-4 adenomas, and annual surveillance for those with 5 or more adenomas. The model calculates the resources required for histopathology and polypectomy and the number of adverse events that result from polypectomies, with associated treatment costs and disutilities. Long-term outcomes associated with each diagnostic outcome are taken from the SBCS

model with no adenomas left in situ and all patients assigned to the correct follow up. The SBCS model includes higher adenoma incidence rates for patients who have had adenomas resected than for patients who started without adenomas (normal epithelium), and the rate of recurrence of adenomas is higher for patients who were initially at higher risk. Cancer incidence, and hence costs and outcomes in the SBCS model, also depend on the surveillance interval assigned. A detailed description of the SBCS model is provided below in Section 5.3.2.3.

Table 33 Diagnostic outcomes by initial risk status: Histopathology strategy

Initial risk (adenomas)	Patient outcome	Hyperplastic resected	Adenomas missed	Surveillance interval	Initial SBCS state
LR (0)	CD	All	None	Correct	Normal (screening)
LR (1-2)	CD	All	None	Correct	LR all resected (screening)
IR (3-4)	CD	All	None	Correct	IR all resected (3-yearly)
HR (5+)	CD	All	None	Correct	HR all resected (annual)

With VC, errors in characterisation of polyps are possible, and hence patients may be left with one or more adenomas in situ (due to false negatives), and/or have hyperplastic polyps unnecessarily resected (due to false positives). Errors in polyp characterisation with VC might also cause patients to be allocated to the wrong follow up strategy – with either too long or too short an interval. The diagnostic outcomes for patients under the VC strategy are shown in Table 34. Outcomes that are impossible or very unlikely are omitted from this table.

For patients without any adenomas, there are only two possible outcomes: they may have a correct diagnosis and have no polyps resected (CD); or they may have one or more hyperplastic polyps removed unnecessarily (HPRC). In either case, patients with no adenomas are very unlikely to be assigned the wrong follow up: the probability of the three or more FP results that would be required for them to be incorrectly assessed as IR is very low. Costs and outcomes for this group are therefore taken from the results for patients starting in SBCS model in the 'normal epithelium' health state and following routine screening. There are five possible diagnostic outcomes for patients with 1-2 adenomas. They may be correctly diagnosed (CD); have one or more adenoma missed, but no resections of hyperplastic polyps and be assigned correctly to routine screening (MAC); have no adenoma missed but one or more hyperplastic polyps resected, either with the correct follow up of routine screening (HPRC) or unnecessary 3-yearly surveillance (HPRI); or they may have one or more adenoma missed and also one or more hyperplastic polyp resected with the correct follow up (MAHPR). Patients in this group start in the SBCS model in the 'post polypectomy (low risk adenomas removed)' health state or in the 'low risk adenomas' health state (1-2 diminutive adenomas

in situ). All patients in this group will be invited for routine screening, except those with one or more FP results who are misclassified as IR. Finally, patients with three or more adenomas (IR or HR) have all possible outcomes illustrated in Figure 31. We assume that patients in this group with one or more missed adenomas start in the 'LR adenomas' health state in the SCBS model, with 1-2 adenomas in situ: although it is possible that patients could have 3 or more adenomas missed, this is very unlikely.

Table 34 Diagnostic outcomes by initial risk status: VC strategy

Initial risk (adenomas)	Patient outcome	Hyperplastic resected	Adenomas missed	Follow up interval	Initial SBCS state
LR (0)	CD	None	-	Correct	Normal (screening)
	HPRC	One or more	-	Correct	Normal (screening)
LR (1-2)	CD	None	None	Correct	LR all resected (screening)
	MAC	None	One or more	Correct	LR adenomas (screening)
	HPRC	One or more	None	Correct	LR all resected (screening)
	HPRI	One or more	None	Too short	LR all resected (3-yearly)
	MAHPR	One or more	One or more	Correct	LR adenomas (screening)
IR (3-4)	CD	None	None	Correct	IR all resected (3-yearly)
	MAC	None	One or more	Correct	LR adenomas (3-yearly)
	MAI	None	One or more	Too long	LR adenomas (screening)
	HPRC	One or more	None	Correct	IR all resected (3-yearly)
	HPRI	One or more	None	Too short	IR all resected (annual)
	MAHPR	One or more	One or more	Correct	LR adenomas (3-yearly)
HR (5+)	CD	None	None	Correct	HR all resected (annual)
	MAC	None	One or more	Correct	LR adenomas (annual)
	MAI	None	One or more	Too long	LR adenomas (3-yearly)
	HPRC	One or more	None	Correct	HR all resected (annual)
	HPRI	One or more	None	Too short	HR all resected (annual)
	MAHPR	One or more	One or more	Correct	LR adenomas (annual)

5.3.2.2 Estimating patient outcome probabilities from polyp-level diagnostic accuracy

Probability of test results for an individual polyp

For the individual polyp, there are four possible VCE test results (TP, FP, FN and TN). The probability of these outcomes can be calculated as a function of the proportion of polyps that are adenomas (p), and the sensitivity (Se) and specificity (Sp) of the test, as shown in Table 35.

Table 35 Virtual chromoendoscopy results for an individual polyp

Polyp results		Interpretation	Probability
TP	True positive	Adenoma correctly classified	P(TP) = p . Se
FP	False positive	Hyperplastic polyp identified incorrectly as an adenoma	P(FP) = (1-p) . (1-Sp)
FN	False negative	Adenoma identified incorrectly as a hyperplastic polyp	P(FN) = p . (1-Se)
TN	True negative	Hyperplastic polyp correctly classified	P(TN) = (1-p) . Sp

p = proportion of polyps that are adenomas;

Se = sensitivity of the VCE test (probability that an adenoma is correctly identified); and

Sp = specificity of the VCE test (probability that a hyperplastic polyp is correctly identified).

Probability of test results for multiple polyps

For patients with more than one polyp, the probabilities of different combinations of test results can be estimated using the binomial distribution. For example, the probability that a patient with n polyps has k false positive test results is:

$$P(k\ FP) = \left(\frac{n!}{k!(n-k)!}\right) P(FP)^{k} (1 - P(FP))^{(n-k)}$$

This formula is used in the decision tree model to estimate the probability of the six main diagnostic outcomes shown in Figure 31 and Table 32. This approach does require an assumption that the test results for individual polyps within a patient are independent of one another: thus, for example, the probability that an individual polyp gives a FP test result is assumed to be constant, regardless of whether other polyps in the patient have given an FP result. In practice, the types of polyp within a patient are likely to be clustered, however we have not identified any data to quantify the extent of any such clustering.

Probability that one or more hyperplastic polyps are misidentified as adenomas

The probability that one or more hyperplastic polyps are incorrectly identified as adenomas in a patient with n polyps is:

$$\begin{aligned} P(\text{one or more FP in a patient}) &= 1 - P(\text{no FP in a patient, k=0}) \\ &= 1 - \left(\frac{n!}{0!(n-0)!}\right) P(FP)^{0} (1 - P(FP))^{(n-0)} \\ &= 1 - (1 - P(FP))^{n} \end{aligned}$$

In the cases where one or more polyp is assessed with low confidence (lc is proportion of polyps assessed with low confidence), the above formula can be generalised to:

$$P(\text{one or more FP in a patient}) = 1 - (1 - P(FP))^{n(1-lc)}$$

Probability that one or more adenomas are missed

In a similar way, the probability that one or more adenomas are incorrectly identified as hyperplastic polyps is:

$$P(\text{one or more FN in a patient}) = 1 - (1 - P(FN))^{n(1-lc)}$$

Or, in the cases where the DISCARD strategy is used, and the proportion of polyps in the proximal region is px:

$$P(\text{one or more FN in a patient}) = 1 - (1 - P(FN))^{n(1-lc)(1-px)}$$

Probability of correct / incorrect follow up intervals

Whether patients are given incorrect follow up depends on their actual number of adenomas and the number of FP and FN results. Thus, a patient with five adenomas, who should be invited for annual surveillance, might be mistakenly invited for colonoscopy only once every three years if one or more adenoma was missed. Estimating the probabilities for every possible combination of adenomas, FP and FN results is complicated. However, the probability of being given the wrong surveillance interval is very low for some patients. For example, patients with no adenomas would need to have three more FP results than FN results, before they would move into the range where they might be offered three-yearly surveillance. Similarly, patients with seven adenomas would need three or more FN results than FP results to move from the annual to three-yearly surveillance category. Given the multiplicative nature of the binomial formula, and relative rarity of FP and FN errors, such outcomes are very unlikely. We therefore made a simplifying assumption: that the probability of three or more errors in polyp characterisation (FP and/or FN) within a patient is negligible. .

For each risk category, we estimated the proportion of patients who have the number of adenomas corresponding to the lower and higher ends of the classification range as,

le = % patients at the lower end / % patients in risk classification
he = % patients at the higher end / % patients in risk classification

The probability of patients having one or more missed adenomas and being assigned to an incorrect follow up strategy (too long an interval) is:

P(one or more missed adenoma in a patient and incorrect surveillance) = P_{le} . P_{MA}

Similarly, the probability of patients having one or more hyperplastic polyp misclassified as an adenoma and being assigned to an incorrect strategy (too short an interval) is:

P(one or more MA in a patient and incorrect SI) = P_{he}. P_{HPR}

Table 36 Summary of probability calculations for diagnostic outcomes

Patient outcome	**Interpretation**	**Follow up interval**	**Probability**
CD	Correct diagnosis	Correct	1 – P(MAC) – P(MAI) – P(HPRC) – P(HPRI) – P(MAHPR)
MAC	Missed adenoma (correct surveillance)	Correct	$(1\text{-le}).(1 - (1 - P(FN))^{n(1-lc)(1-px)})$
MAI	Missed adenoma (incorrect surveillance)	Incorrect – too long	$le. (1 - (1 - P(FN))^{n(1-lc)(1-px)})$
HPRC	Hyperplastic polyp resected (correct surveillance)	Correct	$(1\text{-he}). (1-(1- P(FP))^{n(1-lc)})$
HPRI	Hyperplastic polyp resected (incorrect surveillance)	Incorrect – too short	$he. (1-(1- P(FP))^{n(1-lc)})$
MAHPR	Missed adenoma, hyperplastic polyp resected	Correct	$\binom{n!}{2!(n-2)!}P(FP).P(FN). (1-P(FP)-P(FN))^{(n-2)}$

5.3.2.3 SBCS Markov model

The ScHARR Bowel Cancer Screening (SBCS) model[121] describes the development of adenomas and colorectal cancer and subsequent disease progression for the general population of England eligible for bowel cancer screening. It was developed by ScHARR for the NHS Bowel Cancer Screening

Programme. The model is a 'Markov-type' health state transition model, that takes a cohort approach (rather than individual-level simulation). It estimates QALYs and costs for a cohort of 65-year-olds at risk of developing colorectal cancer over a lifetime horizon and using an annual cycle length. Costs were estimated from the perspective of the English NHS, and a discount rate of 3.5% was applied to costs and QALYs. The basic model structure consists of a natural history model; and a screening and surveillance pathway.

The basic natural history model is illustrated in Figure 32. This shows the expected progression of adenomas and CRC in the absence of an active screening and surveillance programme.

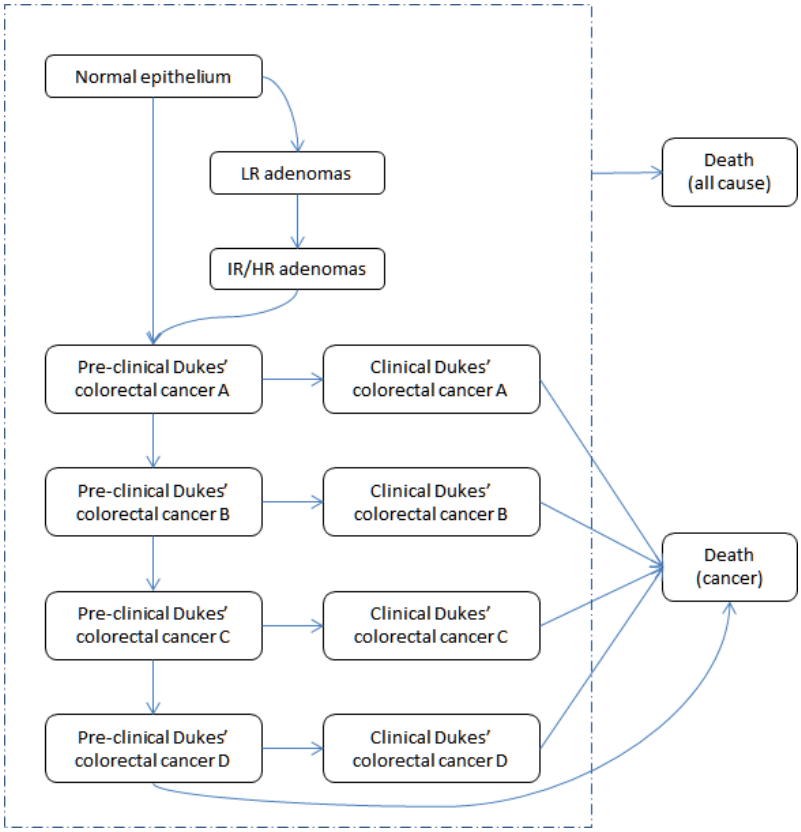


Figure 32 SBCS natural history model

Adapted from Whyte and colleagues[127]

Patients start in one of the pre-cancer health states: normal epithelium (no adenomas); LR adenomas; or IR/HR adenomas. Over time, they may progress through the adenoma-cardinoma route: from normal epithelium to LR adenomas, to IR/HR adenomas, and to pre-clinical Dukes' stage A CRC. It is also possible for patients to transition directly from normal epithelium to pre-clinical stage A CRC (*de novo* cancers). Pre-clinical cancer progresses through the stages, from A to B to C to D, but at some time it is likely to be diagnosed, through chance detection or symptomatic presentation, at which time the patient moves to the related 'clinical' cancer stage. Progression through the clinical cancer

stages is not modelled, instead a stage-specific cancer survival rate is applied. It is also possible for patients with undiagnosed stage D cancer to be fatal. Patients can die from other causes from any of the health states.

The SBCS model was designed to evaluate alternative active screening and surveillance programmes. The post-screening surveillance pathway is illustrated in Figure 33.

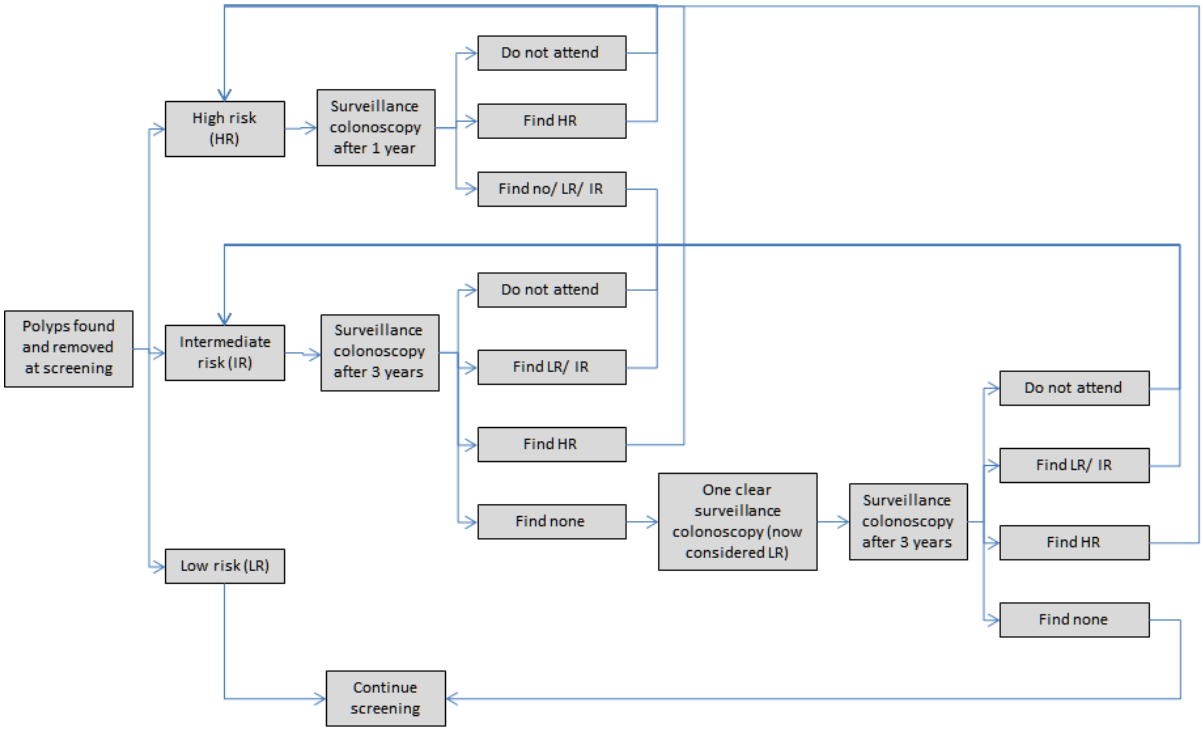


Figure 33 SBCS Surveillance colonoscopy pathway

This shows the assumptions built in to the SBCS model about how patients would be followed up under BSC guidelines, according to findings at an initial colonoscopy after a positive screening result, which reflects the starting point from the end of our decision tree for our base case screening population. Patients assessed to be at low risk following an initial colonoscopy (0-2 diminutive adenomas in our population), or with no adenomas at two successive three-year surveillance colonoscopies, are assumed to be invited for routine screening. The screening pathway in the version of the SBCS model used to generate cost and QALY estimates for the VCE model was chosen to reflect the current NHS Bowel Cancer Screening Programme, with the offer of a home FOBT every 2 years for all men and women aged 60 to 74, and invitation for colonoscopy for patients with an abnormal screening test.

In the SCBS model, colonoscopy is assumed to be standard colonoscopy without virtual chromoendoscopy. However, the model does assume less than perfect sensitivity of colonoscopy for detecting adenomas: 0.77 for LR adenomas and 0.98 for IR/HR adenomas. It also assumes that the cost of histopathology is incurred only for adenomas, a mean of 1.9 per person undergoing colonoscopy. Thus the cost and accuracy of colonoscopy in the SCBS model is possibly more reflective of VCE than with standard colonoscopy.

The simple natural history diagram in Figure 32 does not show all transitions in the SBCS model. In particular, it omits recurrence of adenomas and cancer incidence for patients who have had adenomas removed at colonoscopy. These additional transitions are illustrated in Figure 34.

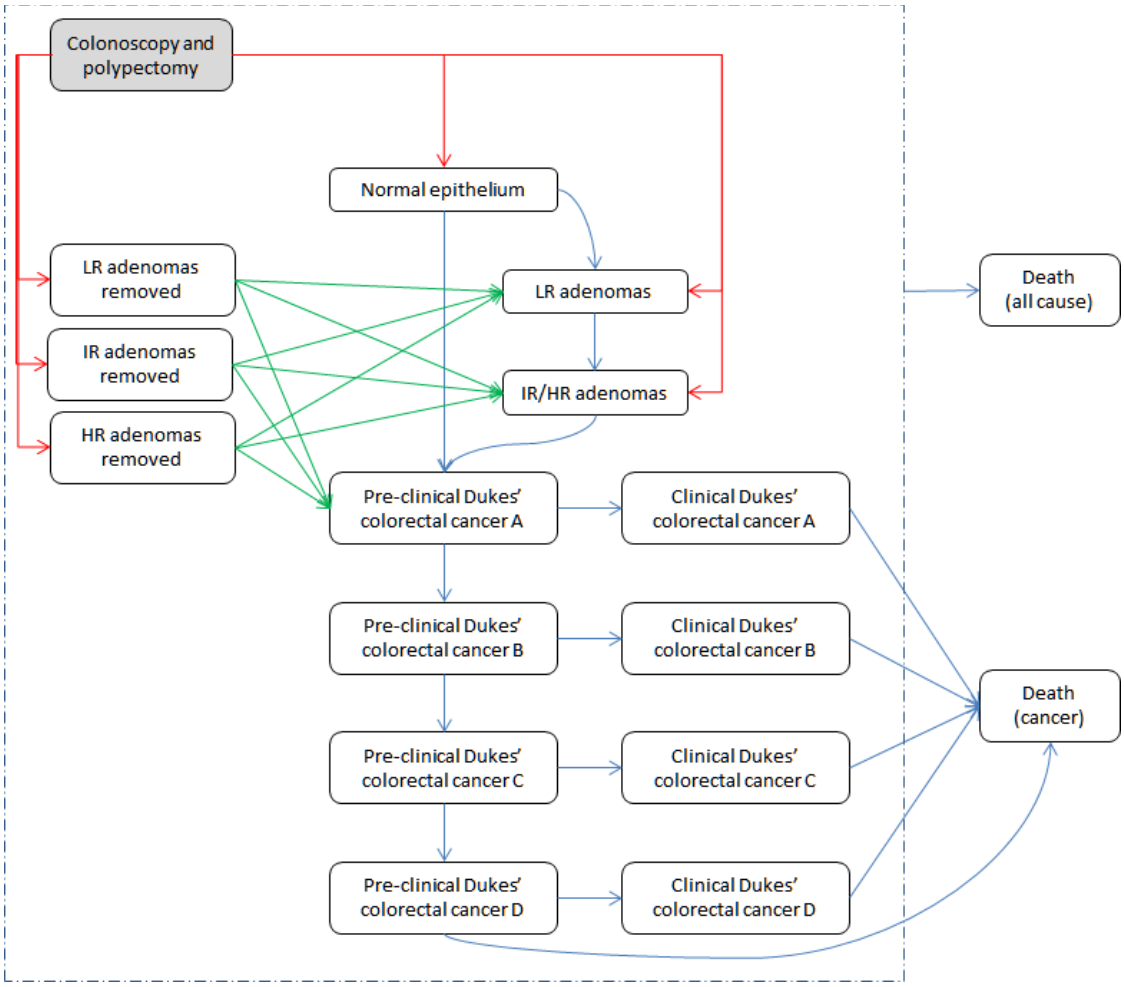


Figure 34 SBCS adenoma recurrence following polypectomy

Following colonoscopy, patients enter the following health states in the SBCS model: patients who started with no adenomas go to the 'normal epithelium' state; patients with 1-2 adenomas left in situ, go to 'LR adenomas'; those with 3 or more adenomas left in situ go to 'IR/HR adenomas'; and patients who have all had adenomas resected go to the LR, IR or HR adenomas removed states,

depending on their initial risk level. Subsequently, patients who have had all adenomas removed may have a recurrence of LR or IR/HR adenomas, and they also have a small chance of 'de novo' cancer, transitioning directly to pre-clinical Dukes' stage A CRC.

Thus, the costs and QALYs for the endpoints of our decision tree were calculated by running the SBCS model with a cohort of 65 year old patients starting in each of the post-colonoscopy health states (normal epithelium, LR adenomas removed, IR adenomas removed, HR adenomas removed, LR adenomas and IR/HR ademomas). The model was run for each possible post-colonoscopy state three times, assuming routine screening, three-yearly surveillance and annual surveillance in turn. Several updates were made to the SBCS model for these analyses. The input parameters are described in section 5.4. Screening and treatment costs were inflated or updated where appropriate (Table 41 and Table 42). Analyses were run assuming the average number of adenomas present in patients with at least one adenoma was 1.9, although the SBCS model does not explicitly simulate the number of polyps. The final cost and QALY estimates from the SBCS model that were used in our decision tree analysis are shown below in Table 46.

5.3.3 Evaluation of uncertainty

The evaluation of the cost-effectiveness of virtual chromoendoscopy technologies is based on uncertain information about variables such as the diagnostic accuracy, polyp demographics, HRQoL and resource use. This uncertainty was evaluated using deterministic and probabilistic sensitivity analyses (PSA). One-way deterministic sensitivity analyses were conducted to evaluate the influence of individual parameters on the model results and to test the robustness of the cost-effectiveness results to variations in the structural assumptions (section 5.5.2.1).

Multi-parameter uncertainty in the model was addressed using PSA (section 5.5.2.3). In the PSA, probability distributions are assigned to the point estimates used in the base case analysis. The model is run for 5000 iterations, with a different set of parameter values for each iteration, by sampling parameter values at random from their probability distributions. The uncertainty surrounding the cost-effectiveness of each treatment is represented using a cost-effectiveness acceptability curve (CEAC) according to the probability that the intervention will be cost effective at a particular willingness to pay threshold. Appendix 9 reports the parameters included in the PSA, the form of distribution used for sampling each parameter, and the upper and lower limits assumed for each variable.

The results of the PSA should be treated with some caution, however, since it does not reflect some important sources of uncertainty or correlations between model parameters. Firstly, we note that the

PSA does not integrate uncertainty over the long-term impact of diagnostic errors on patient outcomes and costs, since we could not obtain correlated samples of cost and QALY outputs from the SBCS model. The PSA also omits correlations between sensitivity and specificity estimates from our bivariate meta-analysis. Statistical advice to the team, indicated that if no threshold effect could be demonstrated between diagnostic sensitivity and specificity of virtual chromoendoscopy, then modelling these parameters as uncorrelated in PSA would have little effect on their uncertainty in comparison to modelling them allowing for correlation. In our meta-analyses (Section 4.1.2), we found that there was no significant evidence of a threshold effect. Therefore, for the PSA we have varied sensitivity and specificity independently. It is most likely that the the consequence of these omissions is that the PSA under-estimates overall uncertainty over the cost-effectiveness of the VCE strategies. In addition, there are uncertainties over some structural assumptions that are not reflected in the PSA.

5.3.4 Model validation

The decision tree model was validated by checking its structure, calculations and data inputs for technical correctness. The model structure was reviewed by clinical experts for appropriateness for the disease and diagnosis. The model was checked for internal consistency by a second health economist. The robustness of the model to changes in input values was tested using sensitivity analyses to ensure that any changes to the input values produced changes to the results of the expected direction and magnitude.

The prediction of correct surveillance intervals was compared between the estimates from the model and those in the published literature. Three studies of NBI[55,56,66] that reported both accuracy of diagnosing individual diminutive polyps and accuracy of assignment of patients to surveillance interval using data from diminutive polyps only were identified by our systematic review of diagnostic studies. In Chandran and colleagues,[55] the diagnostic accuracy was 91.2% whilst the surveillance interval was correctly determined in 98% of patients. In Gupta and colleagues (2012),[56] the diagnostic accuracy was 84.8%, whilst prediction of surveillance interval was accurate in 86.1% to 94.1% of patients if only diminutive polyps were considered. In Paggi and colleagues (2012),[66] diagnostic accuracy for diminutive polyps was 84.0% whilst correct surveillance intervals were applied 85.3% of the time. None of the i-scan or FICE studies identified by our systematic review reported of the accuracy of assignment of patients to a surveillance interval based on diminutive polyps only. The model predicted correct surveillance intervals in 93% to 98% of patients using the virtual chromoendoscopy technologies.

The majority of the estimates of correct surveillance interval prediction identified by our systematic review of diagnostic studies (Section 4.1.3) were based on using virtual chromoendoscopy characterisations for polyps <5mm in size (or in some studies <10mm in size) combined with histopathological assessment of all other polyps (14/17 studies). In these 14 studies[59,62,63,67,68,70,71,74-76,79,82-84] the estimates of correct surveillance interval prediction range between 79.9% and 100% across all virtual chromoendoscopy technologies; only in three of the NBI studies[63,75,76] did some agreements fall below 90.0% . The surveillance interval prediction from our model is broadly consistent with the systematic review findings.

5.4 Model parameters

The following sub-sections report parameters included in the model. The model parameters include polyp and adenoma demographics, diagnostic test accuracy, adverse event rates, health sector costs (such as cost of colonoscopy), HRQoL and long-term epidemiology (such as disease progression). The costs and adverse event parameters have been based upon those previously used in the SBCS model[121] and updated, where necessary.

5.4.1.1 Prevalence of polyps and adenomas

The prevalence of patients presenting with different numbers of polyps and adenomas at colonoscopy were estimated from the literature for three populations: the screening population (base case), and the surveillance and symptomatic populations (used in scenario analyses).

Screening population

We searched for studies that described the distribution of polyps in patients in a bowel screening population. We identified one study by Raju and colleagues[128] who reported data for the distribution of polyps and adenomas per patient. We analysed the distribution of polyps and adenomas to derive the average number of polyps and adenomas for low risk (LR), intermediate risk (IR) and high risk (HR) patients and the frequency of patients in each risk category, assuming all polyps are diminutive.

Raju and colleagues[128] is a retrospective analysis of data from a colon cancer screening programme in the USA. Three hundred and forty three patients underwent colonoscopy between 2009 and 2011. In the study, 46 patients had no polyps, and there were 882 polyps in the remaining 297 patients (2.97 polyps per patient). Of the patients that had polyps, there were 206 patients who had a total of 422 adenomas, i.e. 1.4 adenomas per patient with a polyp, or 2.04 per patient with an adenoma. Thirty percent of patients who had polyps had no adenomas.

We used a graphical data extraction programme (XY Scan)[129] to extract the data from Raju and colleagues. This extraction resulted in a slight overestimation of the number of adenomas (426 instead

of the reported 422) and the number of patients with adenomas (207 instead of 206) in order to keep polyp numbers correct at 882.

In order to calculate the number of polyps per patient in each risk category, we assumed that the overall prevalence of patients with adenomas was evenly distributed across the risk categories, where people had adenomas. The risk stratification was defined according to the current BSG guidelines[28] where people with 1-2 adenomas are low risk, those with 3-4 adenomas are intermediate risk and those with five or more adenomas are high risk. First, we calculated the proportion of patients with the number of adenomas that corresponded with the risk classification and then we calculated a weighted average of the number of polyps and adenomas in these patients. The derivation of the polyp demographics are shown in more detail in Appendix 10. Polyp demographics are shown in Table 37.

Table 37 Prevalence of polyps and adenomas by risk classification for bowel cancer screening patients at colonoscopy

Polyp demographics in patients with at least one polyp	Value	Source
Prevalence of patients with at least one adenoma	0.698	Raju et al.[128]
Prevalence of patients with no adenomas	0.302	Raju et al.[128]
Prevalence of patients with low risk	0.535	Raju et al.[128]
Prevalence of patients with intermediate risk adenoma	0.107	Raju et al.[128]
Prevalence of patients with high risk	0.056	Raju et al.[128]
Average number of polyps	2.97	Raju et al.[128]
Number of polyps, low risk patients	2	Raju et al.[128]
Number of polyps, intermediate risk patients	4.78	Raju et al.[128]
Number of polyps, high risk patients	8.47	Raju et al.[128]
Number of adenomas, low risk patients	1.4	Raju et al.[128]
Number of adenomas, intermediate risk patients	3.34	Raju et al.[128]
Number of adenomas, high risk patients	5.91	Raju et al.[128]

Surveillance population

We were unable to identify any studies that reported the distribution of adenomas in a surveillance population, whereby all patients after colonoscopy had been followed-up for the appropriate surveillance interval as defined by their risk classification. We found several studies that reported the distribution of adenomas at follow-up surveillance for specific subgroups. For example, Lee and colleagues[130] reported the outcome of 12 month surveillance colonoscopy in high risk patients (n=1760) in the NHS Bowel Cancer Screening Programme. Martinez and colleagues[131] reported a

pooled analysis of eight prospective studies comprising 9167 people with previously resected colorectal adenomas during a median follow-up of four years. We found several other studies that reported the distribution of adenomas at various follow-up intervals for patients with more than one adenoma resected.[132,133] In the absence of data that fit our population group, we used these studies, together with an assumption to calculate the distribution of adenomas in this population.

The proportion of patients with no adenomas at follow-up surveillance was similar for Lee and colleagues[130] (49.2%) and Martinez and colleagues[131](53.3%). We chose the estimate from Martinez and colleagues[131] as it was the larger study and not only for high risk patients. We stratified those patients that had low risk, intermediate or high risk adenomas in the same proportion as for the screening population (Table 37). The resulting distribution of adenomas for the surveillance population is shown in Table 38.

Table 38 Proportion of patients by risk category for surveillance and symptomatic populations

Distribution of patients	Surveillance population	Symptomatic population
No adenoma	0.533	0.782
Low risk	0.358	0.125
Intermediate risk	0.072	0.061
High risk	0.037	0.032

Symptomatic population

We identified one relevant study by Mcdonald and colleagues[134] that described the proportion of people who had adenomas in a group of consecutive patients referred from primary care for colonoscopic examination in the NHS. Patients were referred for symptoms including rectal bleeding, change in bowel habits and abdominal pain. No patients were included if they had been referred as a result of the Bowel Cancer Screening Programme. The distribution of adenomas for the symptomatic population is shown in Table 38.

The study also included a small number of patients with irritable bowel syndrome and we have excluded these from our calculation of the distribution of adenomas in the symptomatic population. The study reports the number of people who have no adenomas, low risk adenomas and high risk adenomas. The high risk adenoma group was split between intermediate risk and high risk in the same proportion as for the screening population (Table 38).

5.4.1.2 Diagnostic accuracy

The sensitivity and specificity of histopathology and the virtual chromoendoscopy technologies are taken from the meta-analyses conducted in this report, as described in Section 4. We have assumed that histopathology provides an accurate diagnosis of all polyps (i.e. 100% sensitivity and specificity). The diagnostic accuracy parameters are shown in Table 39 and are for high confidence characterisations of polyps in the whole colon. The proportion of polyps assessed with low confidence is derived from those NBI studies in our systematic review that reported these data and is assumed to be the same for FICE and i-scan.

Scenario analyses were conducted for alternative diagnostic accuracy estimates derived from the systematic review and meta-analysis in section 5.5.2, as follows:

- Sensitivity and specificity for polyps characterised with high confidence in the rectosigmoid colon
- Sensitivity and specificity for polyps characterised with any confidence level in the rectosigmoid colon
- Sensitivity and specificity for polyps characterised with any confidence level in the whole colon
- Sensitivity and specificity for a pooled VCE analysis
- Sensitivity and specificity for endoscopists experienced in the use of NBI

Table 39 Sensitivity and specificity for histopathology, NBI, i-scan and FICE

Parameter	Value	Lower 95% CI	Upper 95% CI	Source
Histopathology sensitivity	1			Assumption
Histopathology specificity	1			Assumption
NBI sensitivity	0.910	0.855	0.945	Meta-analysis
NBI specificity	0.819	0.760	0.866	Meta-analysis
FICE sensitivity	0.814 *	0.732	0.875	Meta-analysis
FICE specificity	0.850 *	0.786	0.898	Meta-analysis
i-scan sensitivity	0.962	0.917	0.983	Meta-analysis
i-scan specificity	0.906	0.842	0.946	Meta-analysis
Proportion low confidence	0.214	0.21	0.22	NBI studies that reported these data in our review

* As there were no data available for sensitivity and specificity for FICE characterisations with high confidence, we have used data from our meta-analysis of FICE with any level of confidence

5.4.1.3 Adverse effects

There are small risks attached to polypectomy such as bowel perforation and bleeding which may lead to hospitalisation and, for those patients who experience perforation, a small risk of death. The probabilities of these adverse effects were taken from the published sources used in the SBCS model and are shown in Table 40.

Table 40 Probabilities of adverse events for perforation and bleeding for patients receiving polypectomy

Parameter	Value	Lower 95% CI	Upper 95% CI	Source
Probability of perforation with polypectomy	0.003	0	0.01	Whyte et al.[121]
Probability of death, for patients with perforation during polypectomy	0.052	0.01	0.11	Gatto et al.[135]
Probability of hospitalisation for bleeding with polypectomy	0.003	0	0.01	Atkin et al.[136]

5.4.1.4 Estimation of costs

Costs were included for colonoscopy, polypectomy, adverse events and histopathology. The unit costs were taken from the NHS Reference costs for 2014/15.[122] A summary of the unit costs is shown in Table 41.

Table 41 Unit costs for colonoscopy and treating adverse events

Parameter	Value	Lower 95% CI	Upper 95%CI	Source
Cost of colonoscopy without polypectomy	£518.36	£340.89	£695.83	HRG 2014-15 FZ51Z, Day case
Cost of colonoscopy with polypectomy	£600.16	£406.24	£794.08	HRG 2014-15 FZ52Z, Day case
Cost of treating bowel perforation (major surgery)	£2,152.77	£902.21	£3,403.33	HRG 2014-15 FZ24E-J Weighted average, non-elective long stay

Cost of admittance for bleeding (overnight stay on medical ward)	£475.54	£327.69	£623.39	HRG 2014-15 FZ38G-P Weighted average, non-elective short stay
Pathology cost per polyp examination	£28.82	£6.78	£50.86	HRG 2014-15 DAPS02

System costs

The equipment and maintenance costs for virtual chromoendoscopy technologies are shown in Appendix 11. These costs are not included in the base case analysis for virtual chromoendoscopy versus histopathology as all equipment and maintenance costs are included within the National Reference Costs for colonoscopy and polypectomy (Table 41). There are differences in the costs between the virtual chromoendoscopy technologies and these are explored in a scenario analysis (section 5.5.2).

Colorectal cancer treatment costs

The SBCS model includes colorectal cancer treatment costs by patient age and Dukes' colorectal cancer staging score. These costs were taken from the study by Pilgrim and colleagues[137] and have been inflated to 2015 prices using The Hospital & Community Health Services (HCHS) index[123] (Table 42).

Table 42 Updates to parameter values in the SBCS model: Bowel cancer screening and colorectal cancer treatment costs (inflated to 2015)

	Dukes' colorectal cancer stage at diagnosis			
Age at diagnosis	A	B	C	D
40-49	£8,871	£8,858	£14,683	£11,862
50-59	£5,789	£7,110	£9,821	£8,557
60-69	£4,686	£5,423	£7,357	£6,596
70-79	£3,220	£3,500	£4,546	£4,423
80-100	£1,398	£1,567	£1,581	£818

5.4.1.5 Training costs

As discussed earlier (Section 1.2.6) in order for endoscopists to accurately use virtual chromoendoscopy, they will need to receive training. This may entail training programmes in the form of video packages and/or supervision from endoscopists experienced in using virtual

chromoendoscopy. Several studies have evaluated training packages that were developed to train endoscopists in the use of NBI.[64,90,138,139]

For example, Ignjatovic and colleagues[138] conducted a prospective education study on a computer-based training module on 21 individuals (novices, trainees, and experienced gastroenterologists) with varying colonoscopy experience in the UK. There was significant improvement in the accuracy in characterisation of polyps after the training. Ignjatovic and colleagues[138] commented that although the NBI learning curve is thought to be relatively short, with an improvement in diagnostic accuracy after as few as 44 polyps, it is not clear how expertise is best transferred to community gastroenterologists and to trainees. McGill and colleagues[64] showed that the performance of endoscopists could be sustained over time by repeating the training module at the mid-point of the study. Meads and colleagues[139] suggest that ongoing training and assessment is necessary to sustain performance.

We assumed the number of days training would be two days per year per endoscopist in common with the NBI study by Solon and colleagues.[116] Using a daily rate for endoscopists of £1104 from PSSRU,[123] and assuming each endoscopist completes 150 endoscopies per year gives a training cost per patient of £14.72.

5.4.1.6 Health-related quality of life

The SBCS model[121] used a study by Ara and Brazier[140] that reported utility values. Ara and Brazier pooled the data from four Health Surveys for England in order to compare self-reported health status and quality of life response for subjects with or without a specified list of health conditions. The mean EQ-5D score for respondents was 0.697, while those without cancer the mean EQ-5D score was 0.798. The mean age for respondents for this health state was 60.9 years.

We conducted a targeted search for other studies reporting the HRQoL for patients with colorectal cancer. The searches sought to identify studies reporting EQ-5D that described the HRQoL in general of patients with colorectal cancer, rather than a specific stage of colorectal cancer, such as metastatic cancer. The searches identified three potentially relevant studies summarized in Table 43. One study was from the UK, one study was from the USA[141] and one from Finland.[142]

Table 43 Summary of HRQoL studies identifed

Study	Year	Country	Study type	Population	EQ-5D values
Djalalov et al.[141]	2014	USA	Systematic review and meta-analysis	26 studies that reported utility weights for CRC health states. 6543 respondents (mean age 62 years)	0.76
Farkkila et al.[142]	2013	Finland	Cross-sectional study	508 Finnish CRC patients (mean age 68 years) Patients were divided into five groups: primary treatment, rehabilitation, remission, metastatic disease and palliative care.	Remission: 0.85; All patients 0.813.
Downing et al.[143]	2015	UK	Population-level study	All individuals diagnosed with CRC in England in 2010 and 2011 who were alive 12 to 36 months after diagnosis were sent a questionnaire. 21,802 of 34,467 patients responded.	Mean EQ-5D values not reported.

CRC - colorectal cancer

Djalalov and colleagues[141] performed a systematic review of utility weights for colorectal cancer. They identified 26 studies providing unique utilities for colorectal cancer health states elicited from 6546 respondents. They included utility assessments including the EQ-5D, HUI3 and time-trade off. The colorectal cancer utility data were analysed using linear mixed-effects models for different variables including colorectal cancer type, stage, and utility measure. They calculated the mean EQ-5D score of the population of people with colorectal cancer to be 0.76. It is unclear if this estimate captures the overall HRQoL for patients with colorectal cancer as there was a greater number of studies included with more severe disease in the meta-analysis, and the overall mean utility score reflects this.

Farkkila and colleagues[142] provide utility values for patients with colorectal cancer in Finland. In this study, patients diagnosed with colorectal cancer received a questionnaire by mail. A total of 508 patients assessed their HRQoL using generic 15D and EQ-5D (with the UK tariff). Patients were divided into five groups: primary treatment, rehabilitation, remission, metastatic disease and palliative care. The patients' HRQoL was compared to population reference values. The study reported an EQ-

5D utility value of 0.813 for all patients with colorectal cancer and 0.85 for patients in cancer remission. The utility values were higher for patients in remission than the standardized general population (non-significant difference). For the purposes of our analysis, we assumed that patients in remission have similar utility to the general population, and therefore the mean decrement for colorectal cancer patients is 0.037.

Downing and colleagues[143] sent a questionnaire to all individuals diagnosed with colorectal cancer in England in 2010 and 2011, who were alive 12 to 36 months after diagnosis and 21,802 patients responded. The questionnaire included questions related to treatment, disease status and HRQoL (EuroQoL). However, Downing and colleagues[143] did not provide mean EQ-5D values.

For our base case analysis, we used HRQoL values from Ara and Brazier,[140] for consistency with the SBCS model. We explored alternative quality of life values from Farkkila and colleagues[142] in a scenario analysis.

5.4.1.7 Disutility

Disutility values were sought for patients who experience adverse events during polypectomy such as bowel perforation or bleeding. However, we were not able to identify values for disutilities for these events from the literature. As an alternative we estimated values for disutility for bleeding by assuming they would be similar to a major gastrointestinal bleed and used the value from Dorian and colleagues[144] of 0.1511 for two weeks, i.e. a total QALY loss of 0.006. Values for perforation were assumed to be the same as for stomach ulcer/abdominal hernia/rupture taken from Ara and Brazier.[140] The disutility value was 0.118 for one month, i.e. total QALY loss of 0.010.

5.4.1.8 Epidemiology of adenoma and cancer progression

Transition probabilities in the SBCS natural history model (progression between the adenoma states, pre-clinical CRC stages and from pre-clinical to clinical CRC stages) and screening test characteristics were estimated using a calibration approach. These parameters are not observable, so they were inferred based on available data on CRC incidence by age and stage in the absence of screening, and from CRC screening datasets. Results are presented in Whyte et al 2012.[121]

The SBCS model uses cancer recurrence rates for people from the NHS bowel cancer screening programme with high risk adenomas and data from a study by Martinez and colleagues[131] for people with low risk adenomas; see Table 44. The proportion of people in the high risk surveillance category who have had a polypectomy requiring annual surveillance is 0.29. Full details of the data and assumptions used are available in Whyte and colleagues.[121]

Table 44 Adenoma recurrence probabilities used in the SBCS model

Description	Probability of transition to	Value
LR adenoma, all adenomas resected	LR adenomas health state	0.100
LR adenoma, all adenomas resected	LR adenomas health state	0.040
LR adenoma, all adenomas resected	CRC health state	[a]
HR adenoma (IR), all adenomas resected	LR adenomas health state	0.163
HR adenoma (IR), all adenomas resected	LR adenomas health state	0.091
HR adenoma (IR), all adenomas resected	CRC health state	[a]
HR adenoma (HR), all adenomas resected	LR adenomas health state	0.188
HR adenoma (HR), all adenomas resected	LR adenomas health state	0.568
HR adenoma (HR), all adenomas resected	CRC health state	[a]
[*]assumed to be the probability of transitioning from normal epithelium to Dukes' A		

CRC - colorectal cancer

To ensure consistency between the model parameters, it is important that the post-polypectomy transition probabilities used align with the other natural history transition probabilities in the model. It was assumed that persons who are undergoing surveillance post-polypectomy are at higher risk of developing adenomas than persons with a normal epithelium, and that polypectomy reduces the risk of developing CRC. Hence restrictions were placed on the post-polypectomy transition probabilities as described in Table 45.

Table 45 SBCS restrictions on transition probabilities post-polypectomy

Restrictions on transition probabilities post polypectomy
Post polypectomy(LR) to LR adenoma **>** Normal epithelium to LR adenoma
Post polypectomy(HR) to LR adenoma **>** Normal epithelium to LR adenoma
Post polypectomy(LR) to HR adenoma **<** LR adenoma to HR adenoma **>** Normal epithelium to HR adenoma
Post polypectomy(HR) to HR adenoma **>** Normal epithelium to HR adenoma
Post polypectomy(LR) to CRC **<** LR adenoma to CRC **>** Normal epithelium to CRC
Post polypectomy(HR) to CRC **<** HR adenoma to CRC **>** Normal epithelium to CRC
Post polypectomy(LR) to LR adenoma< Post polypectomy(HR) to LR adenoma
Post polypectomy(LR) to HR adenoma< Post polypectomy(HR) to HR adenoma
Post polypectomy(LR) to CRC adenoma< Post polypectomy(HR) to CRC adenoma

5.4.1.9 Long-term estimates of costs and QALYs

Table 46 presents the results of the SBCS analyses, showing expected discounted costs and QALYs for patients at each of the diagnostic endpoints from the decision tree model (as listed in Table 34). Estimates are for one person aged 65 years in each diagnostic category, from the end of colonoscopy after a positive FOBT result with removel of polyps if indicated, and then modelled over a lifetime horizon. The costs presented here do not include costs for the initial colonoscopy, polypectomy, histopathology or adverse events, which are modelled in the decision tree. They do include costs for subsequent follow up, including routine screening and surveillance , and for treatment of any incident cancers. Similarly, the QALY estimates do not include effects of any adverse events associated with the initial colonoscopy and polypectomies, but they do include adverse effects associated with subsequent rounds of screening or surveillance, and with incident cancers.

Table 46 Expected lifetime costs and QALYs for 1 person aged 65 undergoing colonoscopy

Initial risk (adenomas)	Patient outcome	Adenomas missed	Hyperplastic polyps resected	Surveillance interval	Costs, £,	QALYs[a]	QALYs[b]
LR (0)	CD	None	None	Invited to screening	109	11.26653	11.27254
	HPRC	None	One or more	Invited to screening	109	11.26653	11.27254
LR (1-2)	CD	None	None	Invited to screening	109	11.26653	11.27254
	HPRC	None	One or more	Invited to screening	109	11.26653	11.27254
	HPRI	None	One or more	3 year surveillance	1075	11.29947	11.30355
	MAI*	One or more	None	Invited to screening	250	11.26399	11.27027
	MAC*	One or more	None	Invited to screening	250	11.26399	11.27027
	HPRMA*	One or more	One or more	Invited to screening	250	11.26399	11.27027
IR (3-4)	CD	None	None	3 year surveillance	1097	11.29934	11.30341
	HPRC	None	One or more	3 year surveillance	1097	11.29934	11.30341
	HPRI	None	One or more	Annual surveillance	1577	11.32057	11.30659

	MAI *	One or more	None	Invited to screening	250	11.26399	11.27027
	MAC	One or more	None	3 year surveillance	1161	11.29891	11.30291
	HPRMA	One or more	One or more	3 year surveillance	1161	11.29891	11.30291
HR (5+)	CD	None	None	Annual surveillance	1584	11.30252	11.30654
	HPRC	None	One or more	Annual surveillance	1584	11.30252	11.30654
	HPRI	None	One or more	Annual surveillance	1584	11.30252	11.30654
	MAI	One or more	None	3 year surveillance	1161	11.29891	11.30291
	MAC	One or more	None	Annual surveillance	1681	11.30152	11.30553
	HPR_MA	One or more	One or more	Annual surveillance	1681	11.30152	11.30553

[a] QALYs using quality of life estimates from Ara and Brazier[140]

[b] QALYs using quality of life estimates from Farkkila et al.[142]

* Results for patients with missed adenomas adjusted to ensure that costs and QALYs are less favourable than if all adenomas had been removed with the same follow up.

Results from the SBCS model were counter-intuitive for patients with one or more adenomas missed and left in situ and routine screening follow up. Estimated QALYs for this group (11.26730) were higher than for patients with all adenomas resected and the same follow up interval (11.26653 for LR). Similarly, long-term cost estimates for patients with routine screening were lower if adenomas were missed (£98) than if all adenomas had been successfully identified and removed (£109). This small inconsistency appears to result from the assumptions about direct (de novo) incidence of cancers from the 'adenomas removed' and 'adenomas in situ' health states (see Figure 34). In the LR group, if all ademonas are removed, the risk of progression to cancer through this direct route compensates for the reduced risk of cancer via the adenoma-carcinoma pathway. To compensate for this effect we adjusted the estimated QALYs and costs for patients with adenomas left in situ and routine screening. We calculated the QALY loss of having adenomas left in situ compared with having all adenomas removed for the HR group with routine screening and similarly with 3-yearly surveillance. Then we calculated the ratio between the 3-year surveillance QALY loss and the routine screening QALY loss. This ratio was then assumed to be the same for the LR group. The same method was used to adjust the cost estimate for LR patients with adenomas lef t in situ and routine screening.

5.5 Results of the independent economic analysis

5.5.1 Base case cost-effectiveness results

The base case analysis patients in the model are those undergoing bowel cancer screening with a starting age of 65 years. The colonoscopy costs are derived from NHS Reference Costs and include the cost of the colonoscopy equipment and its maintenance in the base-case, with all system costs (endoscope, system, and maintenance) identical across interventions. A sensitivity analysis is conducted using costs system, scope and maintenance costs from each manufacturer in section 5.5.2.2.

Table 47 reports the clinical outputs produced by the decision tree model. In the histopathology strategy, all polyps are resected, whilst between 58% and 63% of polyps are resected for FICE and NBI respectively. Virtual chromoendoscopy reduces the number of hyperplastic polyps resected from 1.53 in the histopathology alone strategy to between 0.06 (i-scan) and 0.14 (FICE) but leaves some adenomas in situ (between 0.04 for i-scan and 0.21 for FICE). Virtual chromoendoscopy reduces adverse events due to bleeding and perforations, and deaths from perforations by roughly a third. The correct surveillance interval estimated in the model varies for the virtual chromoendoscopy technologies between 94% (FICE) and 97% (i-scan).

Table 47 Clinical outcomes from the decision tree, for a hypothetical patient receiving colonoscopy

	Histopathology	**NBI**	**FICE**	**i-scan**
Polypectomy	100.00%	63.38%	58.42%	61.84%
Polyps resected	2.97	1.47	1.37	1.45
Hyperplastic polyps resected	1.53	0.13	0.14	0.06
Hyperplastic polyps left in situ	0	1.40	1.39	1.48
Adenomas resected	1.44	1.33	1.22	1.39
Adenomas left in situ	0	0.10	0.21	0.04
Bleeding events	0.003	0.00190	0.00175	0.00186
Perforations	0.003	0.00190	0.00175	0.00186
Perforation deaths	0.000156	0.000099	0.000091	0.000096
Adenomas left in situ (%)	0.00%	7.13%	14.70%	3.04%

Hyperplastic polyps resected (%)	100.00%	8.68%	9.44%	3.68%
Correct Surveillance Interval	100%	94.7%	93.8%	97.4%

The incremental results of the base case deterministic analysis with the long-term model are presented in Table 48. Where an intervention is dominated (more costly and less effective), the calculation of incremental costs for the next least costly intervention is compared to the next non-dominated intervention. Pairwise comparisons to histopathology are also presented for NBI, FICE and i-scan, respectively, for full incremental costs, QALYs, and ICERs.

Table 48 Cost-effectiveness results of the lifetime economic model

	Costs	**Incremental Costs**	**QALYs**	**Incremental QALY**	**ICER (£ per QALY)**
Full incremental results					
Histopathology	£988.95	--	11.2703	--	Dominated
FICE	£901.25	-£87.70	11.2701	-0.0001	
i-scan	£909.74	£8.49	11.2709	0.0008	£10,465.74
NBI	£915.85	£6.11	11.2708	-0.0001	Dominated
Pairwise comparisons					
Histopathology	£988.95		11.2703		
NBI	£915.85	-£73.10	11.2708	0.0005	Dominates
Histopathology	£988.95		11.2703		
FICE	£901.25	-£87.70	11.2701	-0.0001	£671,383 *
Histopathology	£988.95		11.2703		
i-scan	£909.74	-£79.21	11.2709	0.0007	Dominates

* Incremental cost saving per QALY lost.

In pairwise comparisons, NBI and i-scan dominate histopathology, i.e. they are cheaper and more effective. FICE is more cost effective than histopathology, as the ICER for histopathology vs. FICE is greater than £30,000 per QALY. The difference between histopathology and i-scan, the most effective intervention, was 0.25 Quality Adjusted Days per individual. The differences in costs between the virtual chromoendoscopy technologies were less than £15 over a patient lifetime. I-scan is £79 less costly than histopathology and produces 0.0007 more QALYs.

Table 49 shows the costs and QALYs for the intial colonoscopy and for the long-term component for each risk group for NBI vs. histopathology. Most of the cost savings occur for the initial colonoscopy.

For the low risk group, the long-term costs are higher for NBI, due to the small proportion of patients who are assigned to a more frequent surveillance interval. Most of the QALY gains for NBI are from the reduction in deaths from perforation. There are QALY gains for NBI for patients assigned to more frequent surveillance interval, particularly for patients with low risk, and QALY losses for patients with adenomas left in situ and assigned to less frequent surveillance interval.

Table 49 Summary of the costs and QALYs for the intial colonoscopy and the long-term components

	Costs, £			QALYs		
	Histopathology	NBI	Difference	Histopathology	NBI	Difference
Initial colonoscopy	£691.68	£607.46	84.22	-0.00005	-0.00003	-0.00002
0 adenomas	£32.88	£32.88	0.00	3.3986	3.3990	-0.0003
LR adenoma	£58.34	£83.08	-24.74	6.0298	6.0305	-0.0007
IR adenoma	£117.42	£108.36	9.06	1.2095	1.2090	0.0005
HR adenoma	£88.63	£84.07	4.56	0.6324	0.6324	0.0000
Total	£988.95	£915.85	73.10	11.2703	11.2708	-0.0005

5.5.2 Sensitivity analyses

5.5.2.1 One-way deterministic sensitivity analyses

Parameters were varied across a range of lower and upper values. The parameters that were varied in one-way sensitivity analyses are reported in Table 50 and Table 51. Most of the one-way sensitivity analyses use 95% confidence intervals from data identified during our systematic review and targeted parameter searches. However, some data were taken from different ranges, for example to show the variation between studies for these data. The prevalence of adenomas were varied across the possible range for each risk classification.

Table 50 Parameter values used in one-way sensitivity analyses

Parameter	Mean	Lower	Upper	Range definition
NBI Sensitivity	0.910	0.855	0.945	95% CI
NBI Specificity	0.819	0.760	0.866	95% CI
FICE Sensitivity	0.814	0.732	0.875	95% CI
FICE Specificity	0.850	0.786	0.898	95% CI
i-scan Sensitivity	0.962	0.917	0.983	95% CI

Parameter	Mean	Lower	Upper	Range definition
i-scan Specificity	0.906	0.842	0.946	95% CI
Proportion Low Confidence Assessments	0.210	0.105	0.315	Assumed range
Prevalence of adenomas in patients with polyps	0.698	0.600	0.800	Assumed range
Average adenomas in patients that have low risk adenomas	1.395	1	2	Assumed range
Average adenomas in patients that have intermediate risk adenomas	3.341	3	4	Assumed range
Average adenomas in patients that have high risk adenomas	5.913	5	9	Assumed range
Probability of perforation with polypectomy	0.003	0.000	0.010	95% CI
Probability of perforation death	0.052	0.010	0.110	95% CI
Probability of hospitalisation for bleeding	0.003	0.000	0.010	95% CI
Cost of colonoscopy (without polypectomy)	£518.36	£340.89	£695.83	95% CI
Cost of colonoscopy (with polypectomy)	£600.16	£406.24	£794.08	95% CI
Cost of treating bowel perforation (major surgery)	£2,152.77	£902.21	£3,403.33	95% CI
Cost of admittance for bleeding (overnight stay on medical ward)	£475.54	£327.69	£623.39	95% CI
Pathology cost	£28.82	£6.78	£50.86	95% CI
Training cost	£14.72	£10.30	£19.14	95% CI = +/- 30% of mean

Data were not available for the uncertainty around the long-term outcomes. We included one-way sensitivity analyses for these outcomes but used arbitrary ranges. We included the long-term outcomes for patients with incorrect diagnoses, i.e. false negatives and false positives in each risk category, for both costs and QALYs. The ranges used were calculated by adding or subtracting half the difference between a correct diagnosis and the false diagnosis in either costs or QALYs. The ranges used are reported in Table 51.

Table 51 Parameter values used in one-way sensitivity analyses for long-term outcomes for patients with incorrect diagnoses

	Mean	Lower CI	Upper CI	Assumption
Health State Costs				
LR Hyperplastic polyps resected	£1,075	£592	£1,558	CI = 50% of difference between HPR and CD
LR Missed adenoma	£250	£180	£321	CI = 50% of difference between MA and CD
IR Hyperplastic polyps resected	£1,577	£1,337	£1,817	CI = 50% of difference between HPR and CD
IR Missed adenoma	£250	£0	£674	CI = 50% of difference between MA and CD
HR Hyperplastic polyps resected	£1,584	£1,584	£1,584	CI = 50% of difference between HPR and CD
HR Missed adenoma	£1,161	£950	£1,373	CI = 50% of difference between MA and CD
Health State QALY				
LR Hyperplastic polyps resected	11.2830	11.3159	11.2830	CI = 50% of difference between HPR and CD
LR Missed adenoma	11.2627	11.2653	11.2627	CI = 50% of difference between MA and CD
IR Hyperplastic polyps resected	11.3010	11.3042	11.3010	CI = 50% of difference between HPR and CD
IR Missed adenoma	11.2463	11.2817	11.2463	CI = 50% of difference between MA and CD
HR Hyperplastic polyps resected	11.3025	11.3025	11.3025	CI = 50% of difference between HPR and CD
HR Missed adenoma	11.2971	11.3007	11.2971	CI = 50% of difference between MA and CD

LR low risk (1-2 adenomas), IR intermediate risk (3-4 adenomas), HR high risk (≥5 adenomas), HPR hyperplastic polyp resected, MA missed adenoma, CD correct diagnosis

The results of the one-way sensitivity analyses for each virtual chromoendoscopy technology: NBI, FICE, and i-scan (Figure 35 - Figure 37) are presented as pairwise comparisons to histopathology.

For each virtual chromoendoscopy technology, there were 25 parameters evaluated and the 11 most influential parameters on the model results are presented in the corresponding tables. The results show the changes in incremental net monetary benefits, rather than the change in ICERs. As the ICERs are negative, these values are more difficult to interpret.

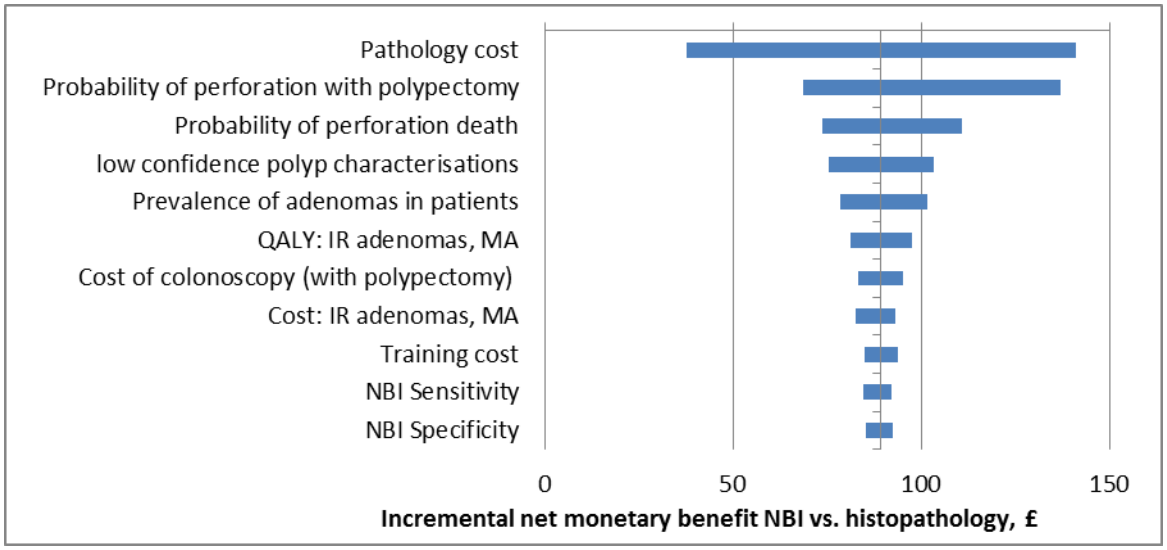


Figure 35 Tornado plot of one-way sensitivity analyses for NBI

For NBI compared to histopatholgy, NBI remained the dominant strategy for all sensitivity analyses. Figure 35 shows that, for NBI compared to histopathology, the most influential parameters on the model results are the pathology cost, the probability of perforation with polypectomy and the proportion of patients who die from perforation, and the long-term QALY estimate for intermediate patients with a missed adenoma.

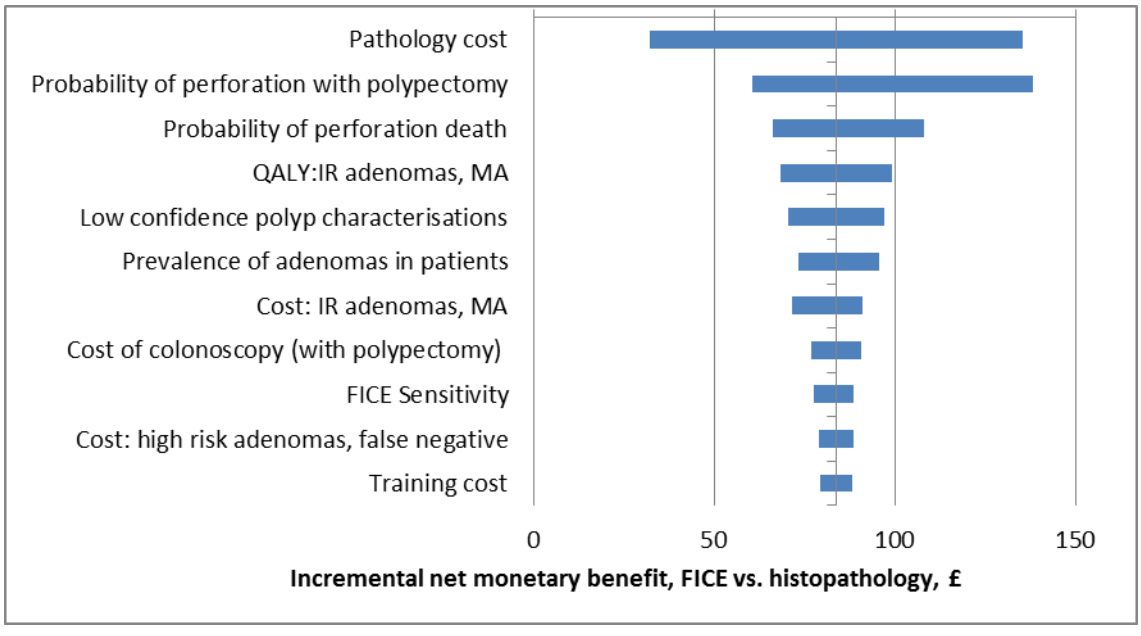


Figure 36 Tornado plot of one-way sensitivity analyses for FICE

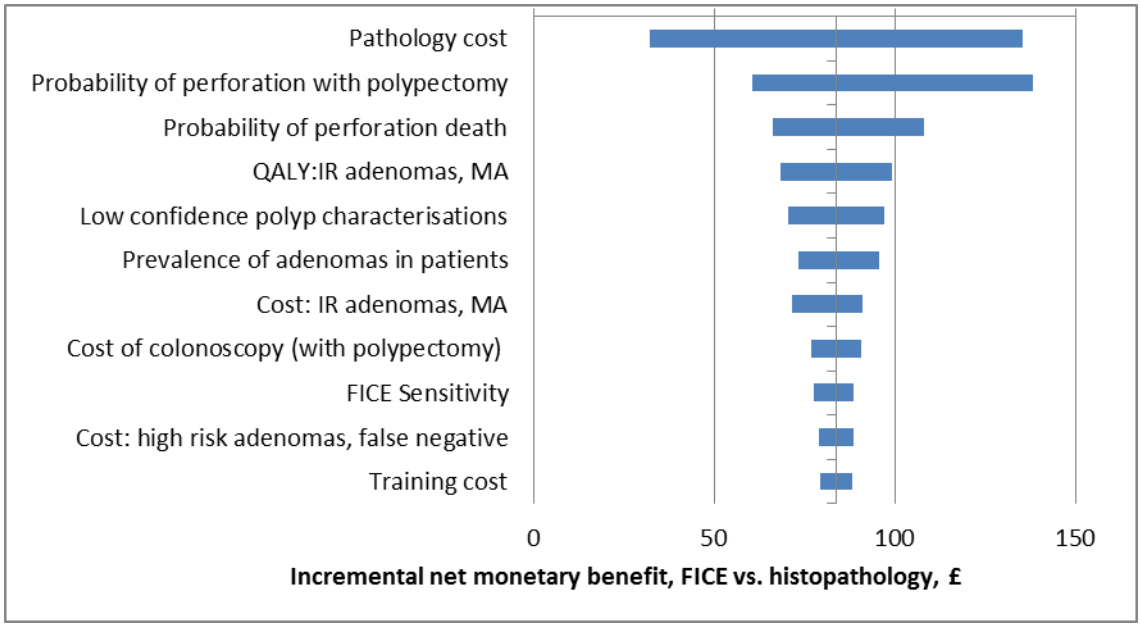


Figure 36 shows that, for histopathology compared to FICE, the most influential parameters on the model results are the pathology cost, the probability of perforation with polypectomy and the proportion of patients who die from perforation, and the proportion of low confidence characterisations made at low confidence. FICE remained more cost effective than histopathology for all sensitivity analyses.

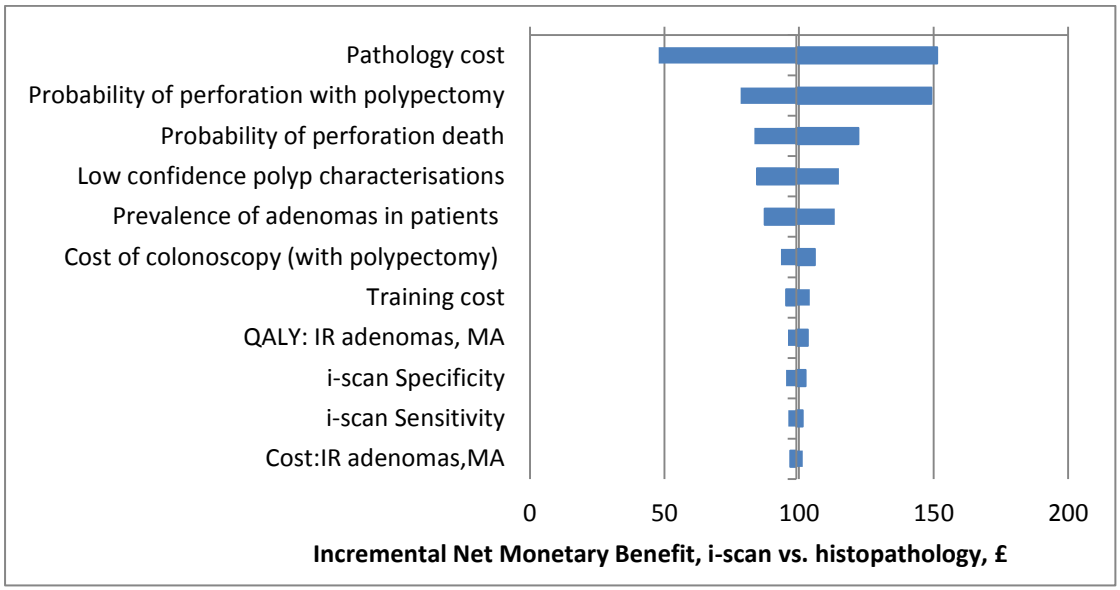


Figure 37 Tornado plot of one-way sensitivity analyses for i-scan

The most influential parameters on the model results for one-way analyses comparing i-scan to histopathology are the pathology cost, the probability of perforation with polypectomy, and the proportion of polyp characterisations made at low confidence.

5.5.2.2 Scenario analyses

In this section, twelve scenario analyses are explored. The descriptions of the scenario analyses are provided in Table 52. Further description of the components of each analysis follow the table.

Table 52 Description of the scenario analyses

#	Analysis	Diagnostic Accuracy (Part of colon – confidence in characterisation)[1]	Other parameters changed
0	Base case	Whole colon – high	
1	Surveillance patients	Whole colon – high	Starting risk distributions changed
2	Symptomatic patients	Whole colon – high	Starting risk distributions changed
3	DISCARD[59]	Rectosigmoid – high	Only polyps in rectosigmoid colon may be left in situ
4	DISCARD[59]	Whole colon – high	Only polyps in rectosigmoid colon may be left in situ
5	DISCARD[59]	Whole colon – any	Only polyps in rectosigmoid colon may be left in situ
6	VC Strategy	Whole colon – any	
7	Costs calculated for each system (endoscope, system, maintenance)	Whole colon - high	Costs for each scope calculated as in Appendix 11.
8	Long-term QALYs derived from SBSC model use alternative utility values	Whole colon - high	Utility values for colorectal cancer derived from Farkkila and colleagues and simulated using SBCS for long-term QALYs (Table 49).
9	Pooled VCE base case	Whole colon - high	
10	NBI, experienced endoscopists	Whole colon - high	
11	NBI, experienced endoscopists	Rectosigmoid – high	Only polyps in rectosigmoid colon may be left in situ
12	Follow-up surveillance	Whole colon - high	Long-term costs and QALYs

[1]FICE diagnostic accuracy is based only on characterisations in the whole colon made at any level of confidence

The population for the base case analysis is for patients referred for colonoscopy following bowel cancer screening. Scenario analyses were used to explore two further populations: patients receiving surveillance colonoscopy following previous adenoma removal (referred to as surveillance patients) (scenario 1), and patients referred for colonoscopy for symptoms suggestive of colorectal cancer (symptomatic patients) (scenario 2). We conduct scenario analyses using alternative starting distributions of patients between risk categories to conduct both of these analyses, the alternative values used in these analyses are reported in Section 5.4.1.1.

For our base case analysis we used the VC strategy. Three scenario analyses using the DISCARD strategy were conducted with different diagnostic accuracy data used for each. The differences between the VC strategy and the DISCARD strategy are described in Section 5.3. Scenario 3 uses diagnostic accuracy data derived from high confidence characterisations in the rectosigmoid colon. Scenario 4 uses diagnostic accuracy data derived from high confidence decisions in the whole colon. Scenario 5 uses diagnostic accuracy data from polyp characterisations made in the whole colon with any level of confidence.

We also conducted a scenario analysis where the VC strategy was applied to the whole colon (Scenario 6), but with diagnostic accuracy data for any level of confidence characterisation instead of diagnostic accuracy from high confidence characterisations in the whole colon (as in the base case); this analysis would represent a worst case scenario on diagnostic accuracy. The diagnostic accuracy data used for Scenarios 3 to 6 are reported in Table 53. All diagnostic accuracy data for NBI and FICE were derived from meta-analyses in Section 4.1.2. For i-scan, diagnostic accuracy for the base case and Scenario 4 was derived from our meta-analysis as reported in Section 4.1.2, whilst diagnostic accuracy for Scenario 3 was derived from Rath and colleagues,[82] and Scenario 5 and 6 were derived from Lee and colleagues.[77]

Table 53 Diagnostic accuracy data used in scenario analyses

	NBI		**FICE**		**i-scan**	
Diagnostic accuracy (colon location – confidence in characterisation)	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Rectosigmoid – high confidence[1]	87.41%	95.26%	81.39%	85.02%	98.10%	94.40%
Whole colon –	90.97%	81.88%	81.39%	85.02%	94.34%	91.53%

	NBI		FICE		i-scan	
Diagnostic accuracy (colon location – confidence in characterisation)	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
high confidence[2]						
Whole colon – any confidence level[3]	88.17%	80.74%	81.39%	85.02%	96.05%	88.15%
[1] Scenario 3 (except FICE); [2] Base case and Scenario 4; [3] Scenario 5 and 6 (and all FICE analyses)						

In the base case analysis, all virtual chromoendoscopy systems have the same cost, as the equipment and maintenance cost for the colonoscopy systems are included in the reference cost of colonoscopy. In this analysis, we investigated the effect on the model results of including the difference in the systems costs compared with the average costs of NBI, FICE and i-scan, using market share data. The net cost differences related to system costs (scope, system and maintenance) from average costs for colonoscopy techniques are reported in Table 54. The calculation of these parameter values is shown in Appendix 11.

Table 54 Net cost difference from the average cost for virtual chromoendoscopy techniques

Intervention	Cost difference	95% CI (Lower)	95% CI (Upper)	Standard Error
NBI	£19.36	£5.08	£33.64	£7.29
FICE	-£61.93	-£81.22	-£42.63	£9.84
i-scan	-£48.27	-£53.22	-£43.32	£2.53

Scenario 8 investigates the effect of alternative utility values, derived through our literature review of quality of life studies, have on the model results. The utility values used to generate these long-term outcomes are reported in Table 55, whilst the long-term QALYs produced through by SBCS model for the alternative utility values are reported in Section 5.4.1.6.

Table 55 Utility values used in the base case analysis and the scenario analysis

Health state	Base Case[140]	Scenario 8[140,142]
No cancer	0.798	0.798
Colorectal cancer	0.697	0.761

Scenario 9 investigates the combined effect ofvirtual chromoendoscopy technologies compared to histopathology. The diagnostic accuracy data for this scenario were taken from our meta-analysis pooling all available studies from high confidence characterisations in the whole colon (described in section 4.1.5) and are shown in Table 58. This scenario is based on a post-hoc meta-analysis used to illustrate a possible class effect of the VCE technologies (NB. It features NBI and i-scan studies, but there was insufficient evidence to include FICE).

Scenarios 10 and 11 use diagnostic accuracy data from studies that reported data for endoscopists experienced in the use of NBI. This scenario is informed by a post-hoc meta-analysis of the sub-set of NBI studies in which endoscopists were experienced in the use of NBI for optical characterisation of polyps. This is in contrast to the base case meta-analysis of NBI studies which included studies of experienced and non-experienced endoscopists. Given the observation of higher diagnostic accuracy according to prior experience of the endoscopist this scenario was conducted to provide a more equal comparison with the meta-analysis of i-scan, given that the majority of studies featured experienced endoscopists. These data are shown in Table 58 and the meta-analysis to derive them is described in section 4.1.2.

Table 56 Diagnostic accuracy data used in scenario analyses for pooled VCE and experienced endoscopists

#	Scenario	Sensitivity	Specificity
9	Pooled VCE base case	91.82%	83.20%
10	NBI, experienced endoscopists (whole colon)	91.83%	82.16%
11	NBI, experienced endoscopists (rectosigmoid)	90.37%	98.14%

In the base case the long-term cost and QALY outcomes, derived from the SBCS model, were estimated assuming the use of standard colonoscopy for any patients requiring follow-up surveillance (i.e. VCE was not used during follow-up colonoscopy). These long-term costs and QALY outcomes do not therefore show the true extent of the future colonoscopies. For example, we would expect there to be future cost savings for virtual chromoendoscopy in any future colonoscopies. We investigated the likely impact on the model results if all patients assigned to the virtual chromoendoscopy group would receive virtual chromoendoscopy technologies for follow-up surveillance (Scenario 12).

The long-term costs and QALYs for the histopathology group were adjusted by an estimate of the differences in costs and QALYs for a follow-up colonoscopy. These were calculated according to the numbers of patients receiving follow-up colonoscopy in each risk group and the additional costs and loss in QALYs at follow-up surveillance, taken from our analysis for the surveillance population (scenario 2, Table 56). From this analysis, the additional cost for each patient receiving histopathology compared to NBI is £84.69 and the loss in QALYs is -0.0007.

We assumed that 20% patients in the low risk group would have a follow-up colonoscopy after 10 years, all intermediate risk patients would have a follow-up colonoscopy after three years and all high risk patients would have a follow-up colonoscopy after one year. Additional costs at colonoscopy were discounted according to how many years until the surveillance colonoscopy. The long-term costs and QALYs for histopathology for the low risk, intermediate risk and high risk groups were then adjusted by the estimates shown in Table 57.

Table 57 Parameters used in follow-up surveillance scenario

Risk group	**Proportion receiving follow-up colonoscopy**	**Time until surveillance colonoscopy**	**Additional cost, discounted @ 3.5% pa**	**Additional discounted QALYs**
Low risk	20%	10 years	£12.01	-0.00015
Intermediate risk	100%	3 years	£76.38	-0.0007
High risk	100%	1 year	£81.82	-0.0007

Results of scenario analyses

Pairwise results of the scenario analyses one to eight are reported for histopathology compared to NBI (Table 58), FICE (Table 59) and i-scan (Table 60).

Table 58 Pairwise results for NBI compared to histopathology

NBI vs. histopathology						
		Histopathology		**NBI**		
#	**Scenario**	**Costs**	**QALY**	**Costs**	**QALY**	**ICER**
0	*Base case*	*£988.95*	*11.2703*	*£915.85*	*11.2708*	*Dominated*
1	Surveillance patients	£925.66	11.2684	£840.97	11.2692	Dominated
2	Symptomatic patients	£910.75	11.2679	£804.35	11.2687	Dominated
3	DISCARD, rectosigmoid – high confidence (diagnostic accuracy)	£988.95	11.2703	£946.84	11.2703	Dominated

4	DISCARD, whole colon – high confidence (diagnostic accuracy)	£988.95	11.2703	£962.08	11.2708	Dominated
5	DISCARD, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£962.38	11.2708	Dominated
6	VC strategy, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£914.29	11.2706	Dominated
7	Costs calculated for each system	£988.95	11.2703	£931.14	11.2708	Dominated
8	Alternate utility values	£988.95	11.2759	£915.85	11.2765	Dominated

The scenarios show that NBI dominates histopathology for all scenarios, i.e. NBI is less expensive and more effective.

Table 59 Pairwise results for FICE compared to histopathology

FICE vs. histopathology						
		Histopathology		**FICE**		
#	**Scenario**	**Costs**	**QALY**	**Costs**	**QALY**	**ICER**
0	*Base case*	*£988.95*	*11.2703*	*£901.25*	*11.2701*	*£671,383*
1	Surveillance patients	£925.66	11.2684	£830.53	11.2687	Dominated
2	Symptomatic patients	£910.75	11.2679	£794.23	11.2684	Dominated
5	DISCARD, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£955.93	11.2705	Dominated
7	VC strategy, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£863.12	11.2701	£963,335
8	Alternate utility values	£988.95	11.2759	£901.25	11.2759	£1,273,941

FICE has fewer scenario analyses because there is only one source of diagnostic accuracy, a meta-analysis of all FICE characterisations in the whole colon at any level of confidence, which eliminates the possibility of conducting Scenarios 3, 4, or 6. For subgroup analysis for surveillance and symptomatic patients and the DISCARD strategy (scenario 5), FICE dominates histopathology. For scenarios 7 and 8 FICE remains cost effective compared to histopathology.

Table 60 Pairwise comparisons of i-scan to histopathology

i-scan vs. Histopathology						
		Histopathology		**i-scan**		
#	**Scenario**	**Costs**	**QALY**	**Costs**	**QALY**	**ICER**
0	*Base case*	*£988.95*	*11.2703*	*£909.74*	*11.2709*	*Dominated*
1	Surveillance patients	£925.66	11.2684	£834.99	11.2693	Dominated
2	Symptomatic patients	£910.75	11.2679	£801.43	11.2689	Dominated
3	DISCARD, rectosigmoid – high confidence (diagnostic accuracy)	£988.95	11.2703	£949.62	11.2706	Dominated
4	DISCARD, whole colon – high confidence (diagnostic accuracy)	£988.95	11.2703	£954.70	11.2707	Dominated
5	DISCARD, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£958.58	11.2708	Dominated
6	VC strategy, whole colon – any confidence level (diagnostic accuracy)	£988.95	11.2703	£913.85	11.2709	Dominated
7	Costs calculated for each system	£988.95	11.2703	£860.82	11.2709	Dominated
8	Alternate utility values	£988.95	11.2759	£909.74	11.2766	Dominated

For all scenario analyses comparing i-scan to hisopatholgy, i-scan was the dominant strategy.

Scenario 9 shows the analysis for pooled VCE compared to histopathology (Table 61). The results for this scenario are similar to the base case analysis for NBI, and VCE dominates histopathology. For the analysis comparing NBI performed by an endoscopist with prior NBI experience to histopathology, the results are also similar to the base case analyses for NBI and VCE.

Table 61 Scenario analyses for all VCE technologies and for endoscopists experienced in NBI

#	**Scenario**		**Costs**	**QALYs**	**ICER (£/QALY)**
9	Pooled VCE, whole	Histopathology	£988.95	11.2703	-

	colon, high confidence	All VCE	£914.96	11.2708	Dominates
10	Experienced endoscopists for NBI, whole colon	Histopathology	£988.95	11.2703	-
		NBI	£916.49	11.2708	Dominates
11	Experienced endoscopists for NBI, rectosigmoid	Histopathology	£988.95	11.2703	-
		NBI	£944.69	11.2703	Dominates

The results for the surveillance scenario where the differences in costs and QALYs between NBI and histopathology in a follow-up colonoscopy were included (Scenario 12), are shown in Table 62. These results are not significantly different to the base case analysis. Compared to the base case analysis, there is an increase in cost savings for NBI of £20 and an increase in incremental QALYs of 0.0003.

Table 62 Results of the follow-up surveillance scenario

	Costs	Incremental Costs	QALYs	Incremental QALY	ICER (£/QALY)
Histopathology	£1,011.75		11.2700		
NBI	£915.85	-£95.91	11.2708	0.0008	Dominates

5.5.2.3 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis was undertaken to provide estimates of cost-effectiveness and the likelihood of cost-effectiveness under joint uncertainty of parameters. In the probabilistic analysis, costs for colonoscopies are assumed to be identical between technologies. The probabilistic sensitivity analysis was undertaken using 5000 simulations. Cost-effectiveness acceptability curves were created using the net-benefit method to represent the probabilities of interventions being the most cost-effective option across a range of cost-effectiveness thresholds. The parameters and the distributions used in the probabilistic sensitivity analysis are shown in Appendix 9. The choice of distributions used in the PSA is based upon common practice.

Results

Table 63 and Figure 38 present the result of the base case analysis using the VC strategy (described in Section 5.3).

Table 63 Full incremental probabilistic cost-effectiveness results for virtual chromoendoscopy (base case)

	Costs	Incremental Costs	QALYs	Incremental QALYs	ICER (£/QALY)
Histopathology	£987.07	--	11.2703	--	Dominated
FICE	£899.74	-£87.33	11.2701	-0.0001	
i-scan	£908.07	£8.34	11.2709	0.0008	£10,298.72
NBI	£914.19	£6.12	11.2708	-0.0001	Dominated

In the base case analysis, i-scan was the most cost-effective technology in 85.2% of analyses at a cost-effectiveness threshold of £20,000 per QALY and in 99.5% of simulations at £30,000 per QALY.

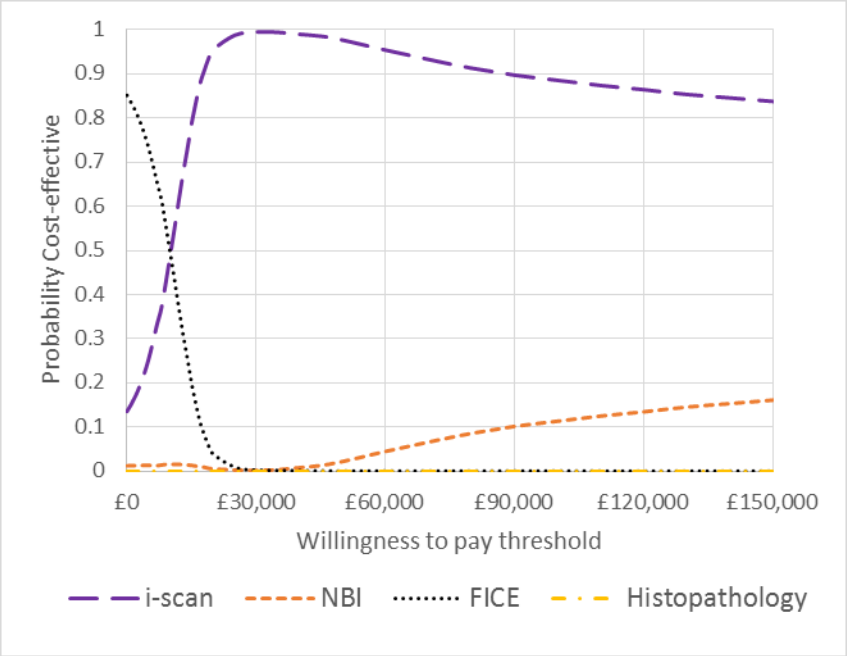


Figure 38 Cost-effectiveness acceptability curves (base case)

5.5.3 Comparison of the economic models

Our systematic review of cost-effectiveness identified two previous economic evaluations by Hassan and colleagues[111] and Kessler and colleagues.[112] Comparing results from these evaluations with our model is difficult, given the differences in design and data used in these studies. Both previous economic evaluations used a similar strategy for virtual chromoendoscopy to that used in our model. They used a resect and discard strategy in the whole colon. Furthermore, Hassan and colleagues included the whole screening population, whereas the population used for Kessler and colleagues and our analysis is for those who had one or more polyps identified. The two previous studies are for a

different health care system (USA) and so there are differences in the health state resource costs used between the models. Also the two previous studies have not presented the results in QALYs.

The proportion of low confidence assessments and the diagnostic accuracy data used in the model are shown in Table 64. The sensitivity of NBI used in the model is similar between the studies but we have used a lower specificity than the other models. Kessler and colleagues assumed that all patients would be assessed with high confidence whereas we assume that only 79% of patients are assessed with high confidence.

Table 64 Diagnostic accuracy parameters used in the economic evaluations

Parameter	**Hassan et al.**[111]	**Kessler et al.**[112]	**Current assessment**
Low confidence assessments	16%	0%	21%
Sensitivity NBI	94%	90%	91%
Specificity NBI	89%	90%	82%

All studies concluded that virtual chromoendoscopy would be cost saving compared to histopathology. The cost saved per person was US$174 versus £74 for our model and the model by Kessler and colleagues[112] respectively over the patient lifetime.

The expected benefit of resect and discard was 0.0005 years of life in Kessler and colleagues[112] compared to 0.0005 QALYs in our model, whilst Hassan and colleagues[111] found there was no difference in life expectancy between groups over the patient's lifetime. The data used for the disease progression to predict life expectancy has not been fully reported in Kessler and colleagues.[112] The cost-effectiveness of the submit all strategy compared to resect and discard all polyps varied and was US$377,460 per life year gained for Kessler and colleagues whilst NBI dominated histopathology in our model. Hassan and colleagues were not able to calculate a value as there was no difference in the life expectancy between the submit all and the resect and discard strategy.

6 ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES

As discussed earlier, it is known that the majority of hospitals that perform endoscopy currently possess endoscopy systems capable of virtual chromoendoscopy. Implementation of the technology will therefore not require large scale replacement of equipment. However, not all systems currently in use comprise fully HD components (i.e. endoscope, light source, video processor, visual display monitor, cabling). Optimum image quality will be attained by fully HD systems, and in some centres this may not be achieved until all equipment is routinely upgraded.

The PIVI statement requires that polyp images taken during virtual chromoendoscopy should be permanently stored and should be of sufficient resolution to support the endoscopists' assessment and clinical decisions.[31] Therefore hospitals would need to implement systems to permit adequate electronic storage of HD images linked to patient's files to allow future re-examination if necessary.

In terms of patient issues and preferences, some patients find colonoscopy to be an uncomfortable experience and therefore may prefer that virtual chromoendoscopy is not used if it may potentially increase the time taken to do the procedure (e.g. the time needed for the endoscopist to inspect the image on the monitor before making a characterisation rather than just resecting it). However, there was very little data from the studies included in our systematic review on differences between procedure times between modes of polyp assessment to provide conclusive evidence.

It is possible that some patients may experience anxiety knowing that a polyp, even one characterised as hyperplastic, has not been resected. Some patients may prefer that all polyps are removed, even when there is negligible risk of them becoming cancerous (notwithstanding the fact that some endoscopists currently leave hyperplastic diminutive polyps in situ, as noted earlier in Section 1 of this report). This would not prohibit virtual chromoendoscopy from being used as part of optical assessment, but would mean that a full DISCARD strategy (i.e. leaving in situ hyperplastic polyps in the rectosigmoid) would not be possible for such patients. If a DISCARD strategy is to be implemented there may be a requirement for patient information about the procedure, and the opportunity for discussion between patient and endoscopist before the colonoscopy.

Although virtual chromoendoscopy is currently used in some centres to characterise colorectal polyps its more widespread use would require greater availability of training and auditing to ensure appropriate use. As discussed earlier, current training practices vary in terms of mode and duration, and studies have illustrated the presence of a learning curve to attain acceptable levels of diagnostic accuracy. The manufacturer of NBI suggests that training of up to two days in duration would be

sufficient for initial training. However, expert clinical advice suggests that for some endoscopists allocating that amount of time for training might not be realistic due to busy work schedules.

Not all endoscopists may want to assume the responsibility for characterising colorectal polyps and leaving those considered to be hyperplastic in situ. If virtual chromoendoscopy is to be recommended in the NHS there may be a need for awareness raising and incentives to encourage greater acceptance and use of this technology in practice.

7 DISCUSSION

7.1 Statement of principal findings

7.1.1 Clinical effectiveness

Thirty studies met the inclusion criteria for the systematic review of test accuracy. These assessed NBI (24 studies), i-scan (5 studies) and FICE (3 studies). Two of these studies assessed two of the technologies of interest in this diagnostic assessment (NBI and i-scan; NBI and FICE). Using the QUADAS criteria, we assessed that the results of the studies are likely to be at a low risk of bias. The evidence we identified meets the decision problem for this diagnostic assessment but there is comparatively little evidence for two of the three technologies being considered (i-scan and FICE). Most of the available evidence evaluated the diagnostic accuracy of NBI for assessing diminutive colorectal polyps. The FICE evidence base was particularly limited. We did not identify any FICE studies that assessed the diagnostic accuracy of endoscopists' real-time high confidence evaluations of diminutive polyps, whereas we found evidence in relation to high confidence assessments made with NBI and i-scan. Some of the included studies explicitly referred to a DISCARD strategy, while others did not.

Most of the included studies reported high sensitivity and specificity (with some exceptions), showing that endoscopists had a high probability of correctly identifying adenomas and hyperplastic polyps when using NBI, i-scan or FICE (sensitivity and specificity results are discussed in more detail below). NPV (that is, the probability that patients who are diagnosed by virtual chromoendoscopy as having a hyperplastic polyp truly do not have an adenoma) was more variable across the NBI studies than the FICE or i-scan studies. There was especially little variation in this outcome across the i-scan studies, where NPV ranged from 93% to 96.30% for all characterisations and 94.74% to 100% for high confidence characterisations. Of the three technologies, i-scan had the most consistently favourable results on this outcome. The greater heterogeneity found among the NBI studies may in part be explained by the larger pool of evidence available for NBI than i-scan and FICE. Additionally, two of the FICE studies were conducted by the same research group, which may have reduced heterogeneity. The heterogeneity in the NBI results may have also been due to variability in the prevalence of adenomas in the populations included in the studies. When prevalence is increased the result is a decrease in the NPV. The more favourable NPV results found for i-scan and variability among the NBI studies may also be explained by the endoscopists' experience in these studies. We note that a range of endoscopists was involved in the NBI studies; some were less experienced in conducting colonoscopy generally and had little or no experience using NBI, ranging to others who were very experienced endoscopists who also had extensive experience of using NBI. By contrast, three of the five i-scan studies included endoscopists with prior experience of i-scan and all the studies were

conducted in single centres often described as academic or specialist centres. The NPV results found in the i-scan studies may therefore not reflect the accuracy that might be achieved by endoscopists working in more generalist or community settings. On the other hand, the large evidence base for NBI may have captured the variability in this outcome that may be observed in practice, where it is likely endoscopists with a range of experience will carry out colonoscopy (although we note that the ESGE guidance recommends that only experienced and adequately trained endoscopists should undertake virtual chromoendoscopy for the real-time assessment of polyps[30]).

Table 65 summarises the key sensitivity and specificity results from the review and the meta-analyses, which we now discuss in more detail. Meta-analysis was conducted where possible but the technologies were not assessed head-to-head in the meta-analyses (as this was not within the decision problem for the assessment, derived from the National Institute for Health and Care Excellence scope), so we cannot comment on how the technologies directly compare to each other statistically.

For all characterisations of polyps (regardless of confidence level) in the whole colon the i-scan (1 study) and FICE (3 studies) results were in the same range of values obtained from the NBI studies (17 and 16 studies for sensitivity and specificity respectively). The summary values from bivariate meta-analysis for sensitivity and specificity of NBI and FICE for all characterisations in the whole colon did not reach 0.90 (i.e. 90%) in either case. Limiting the analysis to high confidence characterisations of polyps in the whole colon, increased the summary sensitivity and specificity values from bivariate meta-analysis which were both over 0.90 for i-scan (2 studies) whereas only the summary value for sensitivity was over 0.90 for NBI (11 studies). As mentioned above, none of the FICE studies analysed outcomes for high confidence assessments of diminutive polyps. As with the NPV results, the higher sensitivity and specificity values seen for i-scan might be explained by the endoscopists in the two i-scan studies being experienced endoscopists working in specialist and academic centres. Therefore, a post-hoc analysis restricting the meta-analysis of high confidence characterisations in the whole colon obtained from studies that reported the endoscopists had prior experience with NBI (4 studies) was conducted. The summary sensitivity and specificity results from this post-hoc analysis of NBI were almost identical to those obtained from all the NBI studies.

Some NBI and i-scan studies provided data on characterisations of polyps in the rectosigmoid colon but no evidence was available for FICE. For all characterisations of polyps (regardless of confidence level) in the rectosigmoid colon the NBI (4 studies) and i-scan (2 studies) results were similar to those obtained from the whole colon. Limiting the analysis to high confidence characterisations of polyps in the rectosigmoid colon increased the summary sensitivity and specificity values from bivariate meta-analysis of NBI and the study estimates from i-scan were also higher (meta-analysis was not

possible for i-scan). A post-hoc analysis restricting the NBI meta-analysis to high confidence characterisations in the rectosigmoid colon obtained from studies that reported the endoscopists had prior experience with NBI (2 studies) increased the summary sensitivity and specificity values further but there was no evidence for i-scan because the single study that reported on high confidence characterisations in the rectosigmoid colon did not report on whether the endoscopist had prior experience using i-scan.

Overall there is evidence showing that, in general, sensitivity and specificity estimates increase when only high confidence characterisations of polyps are considered compared to when all characterisations are considered (i.e. not on the basis of high confidence). It is worth reiterating that the level of confidence with which polyp classifications are made is subjective and is likely to vary between endoscopists. Some endoscopists may refer to the relevant classification system to make a confident polyp characterisation. The studies included in our systematic review did not explicitly state how confidence was achieved. This creates possible uncertainty in the interpretation of diagnostic accuracy based on high confidence characterisations.

We also generated SROC curves to explore the effect of endoscopist experience with NBI on sensitivity and specificity when characterising polyps in the whole colon. This confirmed that endoscopists with prior experience of using NBI to characterise diminutive colorectal polyps achieve higher sensitivity and specificity than endoscopists with no prior experience of using NBI to characterise diminutive colorectal polyps (other than any training that they undertook at the start of the study). It was not possible to discern this effect when comparing the post-hoc meta-analysis of high confidence characterisations in the whole colon made by endoscopists with prior experience of NBI with the meta-analysis of all high confidence characterisations in the whole colon. This maybe because in the pool of 11 NBI studies providing data on high-confidence characterisations in the whole colon three studies included endoscopists with a mix of prior experience and two did not report on prior experience with NBI which would likely have masked any difference between NBI-experienced (4 studies) and NBI-naive endoscopists (2 studies).

Finally, a post-hoc biviariate meta-analysis pooling together all the available evidence for high confidence characterisations of polyps in the whole colon was undertaken which yielded a sensitivity of 0.92 (95% CI 0.87 to 0.95) and a specificity of 0.83 (95% CI 0.78 to 0.87). There were differing opinions among the clinical experts we consulted regarding whether or not it was appropriate to pool evidence from different virtual chromoendoscopy technologies together. The technologies have the same aim (to enhance surface vessel patterns) but achieve this either by filtering the light source (NBI) or by using digital post-processing software to convert white light images such that they appear like

narrow band images (i-scan and FICE). This post-hoc analysis should therefore be treated as illustrative because of the uncertainty regarding whether a class-effect can be assumed and also because the available evidence is predominantly from NBI (11 studies) with only two i-scan studies and none for FICE.

Table 65 Summary of key results

Outcome	**Virtual chromoendoscopy technology**		
	NBI	**i-scan**	**FICE**
All characterisations in the whole colon			
Sensitivity, range across all studies reporting outcome	0.55 to 0.97 (17 studies)	0.95[b] (1 study)	0.74 to 0.88 (3 studies)
Sensitivity, bivariate meta-analysis summary value	0.88 (95% CI 0.83 to 0.92) (16 studies)	Meta-analysis not possible	0.81 (95% CI 0.73 to 0.88) (3 studies)
Specificity, range across all studies reporting outcome	0.62 to 0.95 (16 studies)	0.86[b] (1 study)	0.82 to 0.88 (3 studies)
Specificity, bivariate meta-analysis summary value	0.81 (95% CI 0.75 to 0.85) (16 studies)	Meta-analysis not possible	0.85 (95% CI 0.79 to 0.90) (3 studies)
High confidence characterisations in the whole colon			
Sensitivity, range across all studies reporting outcome	0.59 to 0.98 (13 studies)	0.94 to 0.97[c] (2 studies)	No evidence
Sensitivity, bivariate meta-analysis summary value	0.91 (95% CI 0.85 to 0.95) (11 studies)	0.96 (95% CI 0.92 to 0.98)[d] (2 studies)	No evidence
Specificity, range across all studies reporting outcome	0.44 to 0.92 (12 studies)	0.90 to 0.92[c] (2 studies)	No evidence
Specificity, bivariate meta-analysis summary value	0.82 (95% CI 0.76 to 0.87) (11 studies)	0.91 (95% CI 0.84 to 0.95) (2 studies)	No evidence
High confidence characterisations whole colon by endoscopists with prior experience of the technology (post-hoc analysis)			
Sensitivity, bivariate meta-analysis summary value	0.92 (95% CI 0.89 to 0.94) (4 studies)	0.96 (95% CI 0.92 to 0.98)[d] (2 studies)	No evidence

Specificity, bivariate meta-analysis summary value	0.82 (95% CI 0.72 to 0.89) (4 studies)	0.91 (95% CI 0.84 to 0.95)[d] (2 studies)	No evidence
All characterisations in the rectosigmoid colon			
Sensitivity, range across all studies reporting outcome	0.84 to 0.90 (4 studies)	0.90 to 0.94 (2 studies)	No evidence
Sensitivity, bivariate meta-analysis summary value	0.85 (95% CI 0.75 to 0.91) (3 studies)	Meta-analysis not possible	No evidence
Specificity, range across all studies reporting outcome	0.76 to 0.95 (4 studies)	0.87 to 0.88 (2 studies)	No evidence
Specificity, bivariate meta-analysis summary value	0.87 (95% CI 0.74 to 0.94) (3 studies)	Meta-analysis not possible	No evidence
High confidence characterisations in the rectosigmoid colon			
Sensitivity, range across all studies reporting outcome	0.83 to 0.96 (5 studies)	0.96 (1 study)	No evidence
Sensitivity, bivariate meta-analysis summary value	0.87 (95% CI 0.80, 0.92) (4 studies)	Meta-analysis not possible	No evidence
Specificity, range across all studies reporting outcome	0.88 to 0.99 (5 studies)	0.96 (1 study)	No evidence
Specificity, bivariate meta-analysis summary value	0.95 (95% CI 0.87, 0.98) (4 studies)	Meta-analysis not possible	No evidence
High confidence characterisations in the rectosigmoid colon by endoscopists with prior experience of the technology (Post-hoc analysis)			
Sensitivity, bivariate meta-analysis summary value	0.90 (95% CI 0.71 to 0.97) (2 studies)	No evidence	No evidence
Specificity, bivariate meta-analysis summary value	0.98 (95% CI 0.91 to 1.00) (2 studies)	No evidence	No evidence

Post-hoc pooled analysis of virtual chromoendoscopy technologies: High confidence characterisations in the whole colon	
Sensitivity, bivariate meta-analysis summary value	0.92 (95% CI 0.87 to 0.95) 11 NBI studies, 2 i-scan studies
Specificity, bivariate meta-analysis summary value	0.83 (95% CI 0.78 to 0.87) 11 NBI studies, 2 i-scan studies

[a] All characterisations means not separated by endoscopist confidence level.

[b] One study reported on characterisation of polyps in the distal colon (sensitivity 0.93, specificity 0.83) and one other study reported a per patient analysis of polyps in the last 30 com of colon (sensitivity 0.82, specificity 0.96) but as these outcomes were not for the whole colon they are not directly comparable with the other data in this table row.

[c] One study reported on high confidence characterisations of distal polyps (sensitivity 0.98 and specificity 0.95) but as these data were not for the whole colon they are not directly comparable with the other data in this table row.

[d] The 'High confidence characterisations' result and the 'High confidence characterisations by endoscopists with prior experience of the technology' result are identical because the two studies contributing data to the high confidence meta-analysis were both undertaken by endoscopists with prior experience in using i-scan.

In terms of the other outcomes of interest in this review, none of the studies measured HRQoL, anxiety, number of outpatient appointments or telephone consultations, colorectal cancer or mortality. Only three of the NBI studies and one of the FICE studies reported AEs (e.g. complications of polypectomy such as bleeding). All reported that there were none. Thus, there is only limited data available on AEs in this review. This is an outcome that future studies should consider measuring. A few of the NBI studies reported on the number of polyps that would be resected and discarded if a resect and discard type of management strategy had been in place. Given the limited evidence available it is challenging to determine the number of polyps that would be designated to be left in place, the number of polyps that would be designated to be resected and discarded, and the number of polyps that would be designated for resection and histopathological examination. Likewise, only limited data were available on the length of time to perform the colonoscopy, which means no firm estimates can be made of the additional time it would take during colonoscopy to make real-time assessments of polyp histology.

Table 66 summarises the results of the studies included in this review in relation to the two PIVI requirements that new technologies for the real-time endoscopic assessment of the histology of diminutive colorectal polyps should meet, before a 'resect and discard' strategy could be applied in practice. To reiterate, the criteria specify that for colorectal polyps ≤5 mm in size to be resected and discarded without histopathologic assessment, the endoscopic technology (when used with high

confidence) should have a ≥90% agreement in assignment of post-polypectomy surveillance intervals when compared to decisions based on histopathology assessment of all identified polyps. The criteria also specify that in order for a technology to be used to guide the decision to leave suspected rectosigmoid hyperplastic polyps ≤5 mm in size in place (without resection), the technology should provide ≥90% NPV (when used with high confidence) for adenomatous histology (see section 1.3). Not all the studies that assessed surveillance intervals evaluated these in accordance with the PIVI criteria. We have therefore only included the results here of those studies that clearly calculated concordance of surveillance intervals between virtual chromoendoscopy and histopathology in line with the PIVI requirements. Neither of the two FICE studies that measured surveillance intervals used the PIVI requirements to do this.[83,84] None of the FICE studies examined the NPV for high confidence assessments in the rectosigmoid either. This means that this review did not identify any evidence that enables us to assess how FICE meets the PIVI requirements.

As Table 66 shows, all but one[75] of the NBI and i-scan studies that measured surveillance interval assignment in line with the PIVI criteria[55,62,63,67,68,70,71,75,76,79,82] found a concordance of ≥ 90% between NBI or i-scan and histopathology and thus met this criterion of the PIVI statement (in Ladabaum and colleagues[63] only achieved this for one of the two guidelines used to determine surveillance interval). Most studies did not provide a confidence interval, but where this was reported the lower limit fell below 90% in two of six cases. All the NBI and i-scan studies that measured the NPV of high confidence assessments of diminutive polyps in the rectosigmoid found a ≥ 90% NPV, and thus met the second criterion of the PIVI statement. However, NPV and surveillance interval results for i-scan were only provided by one or two studies respectively, and so the evidence in relation to how i-scan meets the PIVI requirements is limited. Our findings suggest that, on the whole, NBI appears to meet the PIVI criteria, supporting the use NBI to carry out a resect and discard strategy in practice. We note that, in general, where there were discrepancies between the surveillance intervals set following NBI and histopathology, NBI surveillance intervals tended to be shorter than they would have been with histopathology (i.e. patients are seen again sooner).

Table 66 Summary of the review's results in relation to the PIVI criteria

	Assignment of surveillance intervals in accordance with PIVI	**NPV (%), for high confidence assessments of diminutive polyps in the rectosigmoid**
NBI	8 of the 9 studies reporting on this outcome achieved a level of agreement that was ≥ 90%.	92% to 99.4% (Range across 5 studies)

i-scan	2 of the 2 studies reporting this outcome achieved a level of agreement that was ≥ 90%.	97.7% (1 study)
FICE	No evidence	No evidence

It should be noted that our assessment here of the findings of the studies included in this review against the PIVI criteria does not take into account the settings of these studies (i.e. whether they were carried out in specialist, academic settings or routine practice). This could impact on whether virtual chromosendoscopy technologies meet the PIVI criteria. The DISCARD 2 study,[145]which is a large, multicentre prospective UK study, concluded that NBI cannot be recommended for use in routine clinical practice, as when it is used by non-experts in this setting, it did not result in a high enough concordance rate with histopathology for determining surveillance intervals. This study was not included in our systematic review, as it did not meet the inclusion criteria due to only 22% of the colonoscopies being conducted using HD equipment. In this respect it differs from the studies included in this review and the decision problem for this assessment. It is possible that without HD equipment, diagnostic accuracy and appropriate allocation of surveillance intervals may be lower than that achieved when HD equipment is used.

The results of our systematic review have some similarities to those of previous systematic reviews of virtual chromoendoscopy for characterising colorectal polyps, notwithstanding certain differences between reviews in scope and study inclusion criteria.[41-43,146]

For example, the American Society for Gastrointestinal Endoscopy (ASGE) Technology Committee conducted a systematic review to examine whether NBI, i-scan and FICE met the PIVI performance thresholds and therefore whether or not the evidence supported a "diagnosis-and-leave" (ASGE Technology Committee, 2015, p. 1) approach.[146] Literature searches were done on a number of standard health research databases, up to May 2014 (thus the search is around two years older than our literature search). For NBI the review included 19 studies giving estimates of NPV and 10 studies giving estimates of agreement in post polypectomy surveillance intervals. For i-scan there were eight studies of NPV and one study of agreement in post polypectomy surveillance intervals. For FICE there were eight NPV studies and two studies of agreement in post polypectomy surveillance intervals. The majority of the studies used high definition endoscopy systems, and some allowed use of magnification (in contrast with our systematic review).

In the ASGE systematic review[146] the pooled random effects NPV for studies in which an optical characterisation of diminutive polyps with NBI was made with high confidence was 93% (95% CI,

90%-96%). This increased to 95% (95% CI, 92%-98%) when high confidence characterisations were made by endoscopists experienced in optical assessment of colorectal polyps. In our systematic review the majority of NBI studies reported NPV values for high confidence assessments of over 78%, with five studies reporting NPV values of 90% or more.[20,62,67,71,77] (though note that the lower limit of the 95% CI fell below 90% in the majority of studies). The agreement in assignment of post polypectomy surveillance intervals based on optical characterisation of diminutive colorectal polyps with high confidence using NBI was 91% (95% CI, 88%-95%). For i-scan there was no pooled NPV estimate given for high confidence predictions. The overall pooled random effects NPV (any level of confidence prediction) was 84% (95% CI, 76%–91%). A sub-group analysis based on endoscopist experience in performing and interpreting optical biopsies of colorectal polyps reported a pooled random effects NPV of 96% (95% CI, 94-98) for experienced endoscopists compared with a pooled random effects NPV of 72% (95% CI, 69%-76%) for novice endoscopists. As discussed earlier, our systematic review also found that diagnostic accuracy (in terms of sensitivity and specificity) increased in studies (of NBI) involving experienced endoscopists compared to those with less experience. The one i-scan study included in the ASGE review[146] which compared surveillance intervals based on optical assessment compared to histopathology reported an agreement level of 69.5% (95% CI, 63%-75%), thus not meeting the PIVI threshold. The overall pooled random effects NPV for FICE was 80% (95% CI, 76%–85%). This estimate did not improve when restricted to studies of endoscopists experienced in use of optical assessment of colorectal polyps.

Another systematic review, reported by Wanders and colleagues,[41] assessed the diagnostic performance of virtual chromoendoscopy. This review assessed the sensitivity, specificity and NPV of NBI, FICE, and i-scan for optical diagnosis of colonic polyps (in addition to autofluorescence imaging and confocal laser endomicroscopy, which are not within the scope of our systematic review). Key research databases were searched up to January 2013 (thus three years older than our systematic review). The inclusion criteria were broader than our review, permitting studies of diminutive and larger polyps, studies of real time as well as post-procedure image-based virtual chromoendoscopy, studies with or without magnification, and studies with standard or high definition endoscopy systems. However, sub-group analyses were presented based on these criteria, allowing a comparison more aligned to the scope our systematic review to be made. Pooled bivariate meta-analysis sensitivity for the sub-group of five NBI studies with diminutive polyps where the prediction was made with high confidence was 87% (95% CI 78%-93%) and corresponding pooled specificity was 85% (95% CI 74%-92%). These estimates are reported to have been assessed in the context of the PIVI statement, which implies they are based on characterisations of polyps in the rectosigmoid colon. If this is the case then the corresponding NBI pooled sensitivity and specificity estimates for polyps characterised with high confidence in the rectosigmoid in our bivariate meta-analysis are 87% (95% CI 80%-92%)

and 95% (95% CI 87%-98%) respectively (n=four studies). Thus, our estimates are similar in terms of sensitivity but not specificity. A pooled NPV of 83% (95% CI 75%–88%) was reported for NBI, restricted to real time studies (n=35), but not further restricted in terms of diminutive polyps in the rectosigmoid based on high confidence decisions (i.e. in accordance with the PIVI statement), or in terms of the definition status of the endoscopy systems used (standard or high), or magnification status (with or without). The authors suggest that studies of only rectosigmoid NPV are likely to show a good diagnostic performance as the prevalence of non-neoplastic polyps is increased in the rectosigmoid. For FICE bivariate sensitivity and specificity are reported for diminutive polyps, though not stated to be for any particular confidence level (n=four studies). The estimates were 84% (73%-94%) and 87% (79%-94%) respectively, similar to our results (see Table 65). Due to lack of suitable studies no diagnostic accuracy estimates were presented for diminutive polyps characterised with i-scan.

Also of note was that, in the review by Wanders and colleagues,[41] sensitivity and specificity did not differ (statistically) significantly according to whether the EXERA or LUCERA NBI system was used. Even though only the EXERA system is available for use in the UK, the inclusion criteria for our systematic review, based on the National Institute for Health and Care Excellence Scope, allowed studies of both of these systems to be included. (NB. 16 of the NBI studies used EXERA, 5 five used LUCERA and three did not report which system was used – see Table 5). We did not plan to conduct a formal sub-group analyses based on type of NBI system.

7.1.2 Cost-effectiveness

A systematic search of the literature found two economic evaluations of virtual chromoendoscopy compared to histopathology. Both studies compared the resect and discard strategy with current practice of submitting all polyps to histopathology. The evaluations were published in USA. The studies found that there were cost savings for the resect and discard group between US$25 and US$174 per person.

A study by Olympus, the manufacturer of NBI, described a budget impact analysis of NBI in NHS England. The decision tree model has a time horizon of seven years and in each year there is a cohort of patients that undergo endoscopy. The study found that NBI offered cost savings of £141 million over seven years.

We developed an independent cost-effectiveness model comparing NBI, FICE and i-scan with histopathology. The base case analysis uses a virtual chromoendoscopy strategy in a bowel screening population where diminutive polyps in the whole colon are optically characterised. The model uses estimates of diagnostic accuracy from our meta-analysis for diminutive polyps characterised with high

confidence in the whole colon. The results from our economic model suggest that virtual chromoendoscopy is cost saving compared to histopathology with a mean saving of between £73 and £87 per person over their lifetime. There is a small increase in QALYs with NBI and i-scan compared to histopathology of between 0.0005 – 0.0007 QALYs per person, while FICE is associated with 0.0001 QALYs fewer per person than histopathology. NBI and i-scan dominate histopathology, i.e. they are less expensive and more effective. FICE is cost effective compared to histopathology, with a cost saved per QALY lost of £671,383. The model estimates that the correct surveillance interval would be given to 95% of patients with NBI, 94% of patients with FICE and 97% of patients with i-scan. Results are most sensitive to the pathology cost, the probability of perforation with polypectomy and the proportion of patients who die from perforation. Probabilistic sensitivity analyses were conducted for pairwise and incremental comparisons for histopathology with virtual chromoendoscopy technologies. The probabilistic ICERs were similar to the base case deterministic ICERs. At a willingness-to-pay threshold of £20,000 and £30,000, i-scan was most cost effective in 95% and 33% of simulations respectively.

Analyses were also conducted for a surveillance population, who had previously had one or more adenomas detected at an earlier colonoscopy, and a symptomatic patient population which had been referred for colonoscopy with symptoms suggestive of colorectal cancer. These populations had a lower risk of adenomas than the screening population. All virtual chromoendoscopy technologies were less expensive and more effective than histopathology for the surveillance population and symptomatic population analyses.

Analyses were conducted for a DISCARD strategy where diminutive polyps in the rectosigmoid colon are optically characterised. These analyses used the diagnostic accuracy from our meta-analysis for diminutive polyps characterised with high confidence in the rectosigmoid colon (Figure 16). All virtual chromoendoscopy technologies were less expensive and more effective than histopathology. There were smaller differences in costs and QALYs between virtual chromoendoscopy and histopathology for this analysis than for the base case analysis.

The base case results show that the virtual chromoendoscopy technologies are associated with cost savings compared to histopathology and small gains in QALYs. Given the large number of colonoscopies performed every year, the potential cost savings for the NHS are substantial. The cost savings are due to a reduction in the number of polypectomies performed (with a consequent reduction of adverse events from bleeding and perforation) and polyps sent for histopathological examination. Our base case analysis estimated that there would be around 40% fewer polypectomies performed and this would result in between 3% and 15% of adenomas left in situ with virtual

chromoendoscopy and more than 90% fewer hyperplastic polyps resected. The model estimates that virtual chromoendoscopy would lead to incorrect surveillance intervals for between 3% and 6% of patients. The QALY gains are due to the reduction in adverse events, such as perforation, and QALY losses are due to the long-term consequences of not resecting adenomas and patients receiving incorrect surveillance intervals.

The base case analyses indicate that the cost-effectiveness of histopathology compared to virtual chromoendoscopy varies according to the virtual chromoendoscopy technology. The differences in cost-effectiveness between the virtual chromoendoscopy technology are largely attributable to the differences in the diagnostic sensitivity of the technologies, with our meta-analysis calculating sensitivity for i-scan of 0.96 and for FICE of 0.814. We urge caution when comparing between the results of different virtual chromoendoscopy technologies, given the differences in the diagnostic accuracy studies for these technologies in our meta-analyses.

7.2 Strengths and limitations of the assessment

7.2.1 Strengths of the assessment

The strengths of this assessment include that we carried out the systematic review and economic analysis independent of competing interests, and the methods we used were pre-specified in a published protocol. We sought feedback from our expert advisory group on the draft protocol and incorporated their comments into the final version. The protocol was published on the National Institute for Health and Care Excellence website and was discussed by experts in the topic area recruited by National Institute for Health and Care Excellence (specialist members of the appraisal committee). The protocol was also published on the PROSPERO prospective register of systematic reviews website.

We critically appraised all of the diagnostic test accuracy studies included in the review using recognised criteria[37,38] to assess potential risks of bias and to assess the generalisability of the results. Our expert advisory group commented on the protocol and a draft of this report, and we also sought specialist methodological input from the NIHR Complex Reviews Support Unit to conduct this assessment.

Our economic model is in line with current BSG[108] and ESGE[30] guidelines, unlike other models that have examined virtual chromoendoscopy. Hassan and colleagues[111] assume that all patients undergoing screening would have a repeat colonoscopy at 10 years, which is not the recommended surveillance interval in BSG or ESGE guidelines. In Kessler and colleagues,[112] the polyp groups used are inconsistent with both guidelines. Kessler divides patients into four groups by the types of polyps

that patients have, whereas guidelines divide patients into risk groups by the number of adenomas that they have. Solon and colleagues did not examine surveillance intervals, so is not representative of UK practice.[116]

Our model uses the SBCS model to generate long-term outcomes. The SBCS model was developed for the NHS Bowel Cancer Screening Programme.[121] Using long-term outcomes from the SBCS model allows guidance to be consistent across NHS evidence streams.

In line with National Institute for Health and Care Excellence Methodological Guidance,[118] we derived as much of our evidence from systematic searches as feasible. The diagnostic accuracy data were obtained from a robust systematic review and meta-analysis using appropriate bivariate meta-analysis techniques, where possible.[40] Cost data were derived from appropriate NHS sources, and quality of life data were derived from EQ-5D and expressed in QALYs as the primary measure of benefit. Additionally, we conducted a wide variety of sensitivity analyses to explore uncertainty.

7.2.2 Limitations of the assessment

The evidence base for this assessment was particularly limited for FICE and to a lesser extent for i-scan. This limits the conclusions we can draw about the diagnostic accuracy of these technologies for assessing diminutive colorectal polyps in real-time. None of the FICE studies we identified assessed surveillance intervals nor NPV in relation to the PIVI criteria, which meant there was no evidence available to assess how use of FICE meets the PIVI requirements for a resect and discard strategy to be adopted using this technology in practice. Most of the studies included in this review evaluated NBI, but there was heterogeneity in the NBI studies in terms of the original purpose of the studies, country and settings, likely prevalence of adenomas (which can then impact NPV estimates), polyp classification systems used and experience of endoscopists. This makes it difficult to determine the diagnostic accuracy of NBI and to provide clear implications for practice. However, despite this heterogeneity, NBI appears to meet the PIVI requirements (with the caveat that, when reported, the lower limit of 95% confidence intervals was sometimes below the 90% PIVI threshold), supporting its use for a resect and discard strategy in practice.

One limitation of this review is that we did not formally investigate the impact study setting has on diagnostic accuracy estimates. Some research has shown that studies conducted in academic or specialist centres tend find better diagnostic accuracy outcomes than those conducted in generalist settings or community practice.[145] It is not possible to determine from this review how accurate NBI is for the real-time diagnosis of diminutive polyps when used in different settings. We also did not formally investigate the impact of the classification system used for characterising polyps in the

studies. There was much variation in the reporting of the classification schemes used which would have introduced uncertainty in assembling subgroups. Expert clinical advice suggested that diagnostic performance is unlikely to vary according to different schemes as some of the classification schemes are derived from others (e.g. the NICE) classification[20] is based on the Kudo scheme,[22] amongst other schemes). Caution is also advised in the interpretation of our subgroup analysis based on endoscopist's experience with virtual chromoendoscopy, as there was variation between studies in how experience was measured and also there were small numbers of studies in the subgroups.

In order to construct an economic model for histopathology compared to virtual chromoendoscopy, it was necessary to make several assumptions. Firstly, it is not reported in the studies identified how the sensitivity and specificity for individual polyps relates to the surveillance intervals for patients. Whilst some studies in the systematic review of diagnostic accuracy examined correct assignment of surveillance intervals, the data from these studies was insufficient to incorporate in the model. Therefore, we assumed that diagnostic accuracy data for individual polyps was applicable to the entire patient, and assigned patients into risk categories a priori using data from Raju and colleagues.[128] When comparing our modelled outcomes to those found in the systematic review of diagnostic accuracy studies, the model's correct prediction of surveillance intervals was similar to that found in the systematic review (see Section 4 for details). Furthermore, we assumed that the prevalence of adenomas was constant across risk groups with adenomas to predict the number of polyps that patients have. It may be that patients in different risk groups have different ratios of adenomas to polyps. If patients with low risk adenomas have a higher number of polyps per adenoma than patients in the higher risk categories, this would adversely affect the cost-effectiveness of histopathology compared to virtual chromoendoscopy, as more hyperplastic polyps would be resected and sent to histopathology.

The long-term cost and QALY outcomes, derived from the SBCS model, were estimated assuming use of standard colonoscopy for any follow-up surveillance. These long-term costs and QALY outcomes do not therefore show the true extent of the future colonoscopies, for example we would expect there to be future cost savings for virtual chromoendoscopy in any future colonoscopies. It was not feasible to include our decision tree within the SBCS model. However, we included a sensitivity analysis to investigate the likely impact of including virtual chromoendoscopy, which had only a small effect on the model results. This was because the majority of patients were low risk, i.e. few of them would have repeat colonoscopy.

The economic analysis includes only diminutive polyps. Although the decision problem focuses on diminutive polyps, some people with diminutive polyps will also have larger sized polyps (falling into

the ‘small’ and ‘large’ categories). We attempted to incorporate large and small polyps using data from studies identified in the systematic review and meta-analysis as well as targeted searches, but there was insufficient data to allow coherent analysis of larger polyps. In practice, large polyps would be assessed using only histopathology, and the effect would be an increase the number of patients with intermediate and high risk adenoma (i.e. shorter surveillance intervals), and a decrease in the number of polyps characterised as adenomas in intermediate and high risk patients. It is this last feature of the analysis that made assessing large polyps infeasible as no data were available that indicated the number of polyps found in patients with large polyps at intermediate or high risk. Additionally, no information could be identified on what proportion of patients in the intermediate risk category had two or fewer adenomas with one adenoma being large. Including small polyps would only affect the proportion of patients assessed using only histopathology. Surveillance intervals for small polyps are identical to diminutive polyps.

Further, the model does not differentiate between the type of polyp such as depressed polyps or sessile serrated polyps. No diagnostic accuracy data were identified specifically for either type of polyp. Additionally, sessile serrated polyps are rare and no diagnostic accuracy data were available for diminutive sessile serrated polyps from our systematic review of diagnostic studies (Section 4). These polyps may be more likely to be given a low confidence assessment, in which case they would therefore undergo histopathology.

Another uncertainty is the variation in diagnostic accuracy of virtual chromoendoscopy that would occur as a result of polyps that are unable to be successfully retrieved for histopathological analysis (e.g. due to fragmentation). We have noted earlier in this report (Section 1.2) that histopathology, despite being the accepted reference standard, is imperfect. Evidence shows that polyp retrieval failure increases significantly with smaller polyps, particularly those which are diminutive, even when resected by experienced colonoscopists. Lost polyps would be classified as adenomas, even though many would be hyperplastic. A retrospective analysis of 4383 polyps resected from 1495 patients undergoing colonoscopy in the BCSP reported a polyp retrieval failure rate of 6.1%. In our systematic review estimates of polyps not successfully resected for histopathological analysis, where reported, ranged from 0.5% (Basford)[79] to 13% (DISCARD)[3] though in most studies estimates were below 5%. The effect of this is to reduce the diagnostic accuracy of histopathology relative to that of virtual chromoendoscopy.[3] We note that some polyps resected using the virtual chromoendoscopy strategy would also be sent to histopathology. We have not been able to incorporate this uncertainty into our economic analysis due to lack of data to inform how this would affect all of the relevant input parameters. It may lead to a small reduction in the cost of histopathological assessment due to fewer polyps being sent to the laboratory.

The data on recurrence rates post-polypectomy in the SBCS model have several limitations. The transition probabilities reported in Table 44 are not age-dependent; however, the transition probabilities used in the model are age-dependent. The study populations do not reflect the English bowel cancer screening population, are quite small in size, do not use the BSG surveillance guidelines to categorise adenomas, and report highly varying recurrence rates. The SBCS data on recurrence rates for people classified as intermediate or high risk and undergoing one or three yearly surveillance have not been updated with more recent data from the NHS cancer screening programme.

The full uncertainty around the model results have not been explored in the PSA as the long-term outcome parameters have not been varied. These data were not available from the SBCS model.

7.3 Uncertainties

We considered that the participants enrolled in the NBI, i-scan and FICE studies included in the systematic review of diagnostic accuracy are generally likely to be representative of the types of participants who would receive colonoscopy in the UK for screening, surveillance or on account of symptoms experienced. The majority of the studies were conducted in a single centre and so the results of these studies may not be transferrable to other centres. The endoscopists who took part in the NBI studies had a range of experience with endoscopy and NBI across the studies, and it is unclear how this reflects the experience of endoscopists currently working in UK practice. Endoscopists underwent training in NBI in the majority of the NBI studies but it is unclear how representative this training may be of any received in current UK practice. Relatedly, three of the five i-scan studies were conducted by endoscopists with prior experience of using i-scan, in single centres often described as academic or specialist centres. The results of these studies may therefore not be applicable to less experienced endoscopists working in more generalist or community settings. As we did not explore the effect of the study setting on the results from the NBI studies, it is unclear how generalisable the NBI findings are to specialist and generalist centres in the UK. The European (ESGE) guidance[30] recommends that only experienced and adequately trained endoscopists should undertake virtual chromoendoscopy for the real-time assessment of diminutive colorectal polyps. Our review suggests that better diagnostic accuracy (i.e. sensitivity and specificity) outcomes are obtained by more experienced endoscopists, supporting the need for endoscopists to have adequate experience and training in these technologies to use them for real-time diagnosis.

Most of our studies reported diagnostic accuracy derived from expert endoscopists, so the results may not be generalizable to endoscopists with less experience with virtual chromoendosopy technologies. It may be that the level of expertise in endoscopists is lower than in the studies, which would result in lower diagnostic accuracies seen in clinical practice.

The long-term outcomes from the SBCS model include disease progression for patients with small (6-9mm) and large (>10mm) adenomas. It is likely that this overestimates the cancer rates in patients with diminutive polyps who would receive different management due to the use of virtual chromoendoscopy technology. It may be that cancer rates are lower in these patients than predicted by the SBCS model, which would result in lower QALY losses for people treated with virtual chromoendoscopy and therefore increase the cost-effectiveness of histopathology compared to virtual chromoendoscopy.

The FICE diagnostic accuracy data does not include data on polyp characterisations made with high confidence or polyp characterisations made in the rectosigmoid colon, so these cost-effectiveness results are not directly comparable to those of the other virtual chromoendoscopy technologies. More data on the diagnostic accuracy of FICE is necessary to adequately represent its cost-effectiveness.

We have not included within the analysis any benefits to patients in the case where they are informed of the results more quickly or do not have to attend follow-up consultation. There may also be a reduction in anxiety that some patients may experience whilst waiting for results. There was insufficient evidence on these factors to include within the economic analysis.

8 CONCLUSIONS

8.1 Implications for service provision

This assessment found that virtual chromoendoscopy technologies (i.e. NBI, i-scan and FICE), using HD systems without magnification, have potential for use in practice for the real-time assessment of diminutive colorectal polyps. The studies identified in this review suggest that on the whole NBI and i-scan meet the PIVI requirements for these technologies to be used in practice to carry out a 'resect and discard' strategy. Data for i-scan supporting this, though, were limited, and most data were from studies involving endoscopists with prior i-scan experience working in specialist or academic centres. It was unclear how generalisable the NBI results in relation to the PIVI were to UK routine practice settings, as we did not investigate the impact of study setting. Due to limited evidence, it is unclear which of the three virtual chromoendoscopy technologies perform the best. NBI and i-scan had generally better diagnostic accuracy outcomes than FICE, but, again, a greater proportion of i-scan studies were known to involve endoscopists with prior experience of i-scan. Diagnostic accuracy results for NBI were more heterogeneous, but we found that endoscopists with prior experience of NBI achieved higher diagnostic accuracy results than endoscopists with no prior NBI experience. Our findings suggest, as per the ESGE guidance,[30] that virtual chromoendoscopy should be undertaken by experienced and adequately trained endoscopists. This has implications for practice in terms of the need to provide training. Virtual chromoendoscopy technologies were cost saving compared to histopathology. NBI and i-scan were more effective than histopathology. FICE was cost effective compared to histopathology.

Uptake of virtual colonoscopy for the assessment of diminutive polyps in practice will likely depend on the willingness of colonoscopists to take on the responsibility for characterising polyps and the provision of equipment for NBI, i-scan and FICE. We understand that most endoscopes used in the UK have this technology available, although not all centres may have HD equipment. We did not find any studies measuring patient HRQoL, anxiety or the acceptability of virtual chromoendoscopy to patients, so it is unclear how comfortable patients would be with virtual chromoendoscopy being used to assess their polyps. Some patients may experience anxiety knowing that a hyperplastic polyp has not been resected. Some patients may prefer that all polyps are removed, even when there is negligible risk of them becoming cancerous.

8.2 Suggested research priorities

More research is needed to assess the diagnostic accuracy performance of i-scan and FICE when used without magnification to assess diminutive colorectal polyps in real-time, as there is currently only limited evidence available regarding these two technologies. We also suggest that evaluations of the performance of these technologies in generalist, routine practice settings would be informative,

particularly as the i-scan literature is currently limited, and most studies involved endoscopists with prior experience of i-scan working in specialist or academic centres. Multi-centre studies, across a range of settings, would also be informative.

Randomised head-to-head comparisons of NBI, FICE and i-scan would be useful to directly compare outcomes when these technologies are used without magnification to assess diminutive colorectal polyps in real-time. We only identified two head-to-head studies in this review, and so we could only narratively comment on which technologies may perform better. (NB. head to head comparisons of the technologies were not within the decision problem for this assessment, but they may nonetheless be informative to endoscopists interested in using them).

Further studies evaluating the effect of endoscopist experience and training on diagnostic accuracy outcomes when using these technologies would be useful. Endoscopist experience and training is an important consideration and we found few studies that compared the performance of endoscopists with different levels of training and experience, limiting the extent to which we could investigate the influence of this on outcomes in this review.

Future studies should measure adverse effects of polypectomy to provide clearer information about the potential harms of these technologies when used to carry out a 'resect and discard' strategy compared to histopathological assessment of all polyps. We suggest that it would be ideal if future studies also included measures of HRQoL and patient anxiety, as it is currently unclear how patients will respond to the use of these technologies in practice.

Longitudinal data from studies following-up patients over time since their colonoscopy procedure was carried out are needed to quantify the impact of these technologies on colorectal cancer incidence, longer-term HRQoL and mortality.

9 REFERENCES

1. NHS Choices. *Bowel polyps*. http://www.nhs.uk/Conditions/polyps-bowel/Pages/Introduction.aspx (accessed 13/01/2016).
2. Kang YK. Diminutive and Small Colorectal Polyps: The Pathologist's Perspective. *Clinical Endoscopy* 2014;47(5):404-8.
3. Wang LM, East JE. Diminutive polyp cancers and the DISCARD strategy: Much ado about nothing or the end of the affair? *Gastrointestinal Endoscopy* 2015;82(2):385-88.
4. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. *Gastrointestinal Endoscopy* 2003;58(6 Suppl):S3-43.
5. Williams JG, Pullan RD, Hill J, Horgan PG, Salmo E, Buchanan GN, et al. Management of the malignant colorectal polyp: ACPGBI position statement. *Colorectal Disease* 2013;15 Suppl 2:1-38.
6. Cancer Research UK. *Bowel Cancer (colorectal cancer)*. http://www.cancerresearchuk.org/about-cancer/type/bowel-cancer/ (accessed 07/07/2016).
7. Ballinger AB, Anggiansah C. Colorectal cancer. *BMJ* 2007;335(7622):715-8.
8. National Institute for Health and Care Excellence. Colorectal cancer: the diagnosis and management of colorectal cancer (CG131): National Institute for Health and Care Excellence (NICE), 2011.
9. van Putten PG, Hol L, van Dekken H, Han van Krieken J, van Ballegooijen M, Kuipers EJ, et al. Inter-observer variation in the histological diagnosis of polyps in colorectal cancer screening. *Histopathology* 2011;58(6):974-81.
10. NHS Choices. *Endoscopy*. http://www.nhs.uk/conditions/endoscopy/pages/introduction.aspx (accessed 13/01/2016).
11. Varadarajulu S, Banerjee S, Barth BA, Desilets DJ, Kaul V, Kethu SR, et al. GI endoscopes. *Gastrointestinal Endoscopy* 2011;74(1):1-6.e6.
12. Manfredi MA, Abu Dayyeh BK, Bhat YM, Chauhan SS, Gottlieb KT, Hwang JH, et al. Electronic chromoendoscopy. *Gastrointestinal Endoscopy* 2015;81(2):249-61.
13. Olympus. EVIS LUCERA ELITE Concept Brochure. Japan: Olympus Medical Systems Corp., 2013:8.
14. Olympus. EVIS EXERA III Brochure. USA: Olympus America Inc., 2012:9.
15. Olympus. EVIS LUCERA SPECTRUM family brochure. 2009:7.
16. Fujinon Fujifilm. Fujinon Video endoscopes 500 series scope. Electronic video endoscopy system 4400. Japan: Fujinon Corporation, 2009:13.
17. Pentax Medical. *What is i-scan?* https://i-scanimaging.com/index.php?pg=What is i-scan?&l_id=1 (accessed 13/01/2016).

18. Pentax Medical. *i-scan imaging*. https://www.pentaxmedical.com/pentax/en/95/1/i-scan-imaging/ (accessed 13/01/2016).
19. American Society for Gastrointestinal Endoscopy (ASGE) Technology Committee. High-definition and high-magnification endoscopes. *Gastrointestinal Endoscopy* 2014;80(6):919-27.
20. Hewett DG, Kaltenbach T, Sano Y, Tanaka S, Saunders BP, Ponchon T, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. *Gastroenterology* 2012;143(3):599-607.e1.
21. Longcroft-Wheaton GR, Bhandari P. A Novel Classification System (NAC) for the Characterisation Colonic Polyps Using FICE Without Optical Magnification: A New Classification System (NAC). *Gastrointestinal Endoscopy* 2012;75(4):322-22.
22. Kudo S, Tamura S, Nakajima T, Yamano H, Kusaka H, Watanabe H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointestinal Endoscopy* 1996;44(1):8-14.
23. Wada Y, Kudo SE, Kashida H, Ikehara N, Inoue H, Yamamura F, et al. Diagnosis of colorectal lesions with the magnifying narrow-band imaging system. *Gastrointestinal Endoscopy* 2009;70(3):522-31.
24. IJspeert JE, Bastiaansen BA, van Leerdam ME, Meijer GA, van Eeden S, Sanduleanu S, et al. Development and validation of the WASP classification system for optical diagnosis of adenomas, hyperplastic polyps and sessile serrated adenomas/polyps. *Gut* 2016;65(6):963-70.
25. Basford PJ, Longcroft-Wheaton GR, Bhandari P. The learning curve for in-vivo characterisation of small colonic polyps: Number needed to train (NNT) is 200 polyps. *Gastrointestinal Endoscopy* 2013;1):AB528.
26. Rees C, Gibson ST, Rutter MD, Baragwanath P, Pullan R, Feeney M, et al. UK Key Performance Indicators & Quality Assurance Standards for Colonoscopy. London, 2015.
27. National Institute for Health and Care Excellence (NICE) Diagnostics Assessment Programme. *Virtual chromoendoscopy for real-time assessment of colorectal polyps during colonoscopy - Final scope*. https://www.nice.org.uk/guidance/indevelopment/gid-dg10004/documents (accessed 4/7/2016).
28. Cairns SR, Scholefield JH, Steele RJ, Dunlop MG, Thomas HJ, Evans GD, et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). *Gut* 2010;59(5):666-89.
29. National Institute for Health and Care Excellence (NICE). Colorectal cancer prevention: colonoscopic surveillance in adults with ulcerative colitis, Crohn's disease or adenomas. London, 2011.

30. Kaminski MF, Hassan C, Bisschops R, Pohl J, Pellise M, Dekker E, et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2014;46(5):435-49.
31. Rex DK, Kahi C, O'Brien M, Levin TR, Pohl H, Rastogi A, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointestinal Endoscopy* 2011;73(3):419-22.
32. Picot P, Cooper K, Pickett K, Rose M, Shepherd J. *Virtual chromoendoscopy for the real-time assessment of colorectal polyps in vivo*. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016037767.
33. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York Publishing Services Ltd.: CRD, 2009.
34. Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Chapter 11. Interpreting results and drawing conclusions. In: Deeks J, Bossuyt P, Gatsonis C (eds.) *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version*; 2010.
35. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y, . Chapter 10. Analysing and presenting results. In: Deeks J, Bossuyt P, Gatsonis C (eds.) *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* 2010.
36. National Institute for Health and Care Excellence. Diagnostics Assessment Programme manual. Manchester, 2011.
37. Reitsma J, Rutjes A, Whiting P, Vlassov V, Leeflang M, Deeks J. Assessing methodological quality. In: Deeks J, Bossuyt P, Gatsonis C (eds.) *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* 2009.
38. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3(25).
39. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health* 2015;18(4):103-9.
40. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58(10):982-90.
41. Wanders LK, East JE, Uitentuis SE, Leeflang MM, Dekker E. Diagnostic performance of narrowed spectrum endoscopy, autofluorescence imaging, and confocal laser endomicroscopy for optical diagnosis of colonic polyps: a meta-analysis. *Lancet Oncology* 2013;14(13):1337-47.

42. McGill SK, Evangelou E, Ioannidis JPA, Soetikno RM, Kaltenbach T. Narrow band imaging to differentiate neoplastic and non-neoplastic colorectal polyps in real time: a meta-analysis of diagnostic operating characteristics (Provisional abstract). *Gut* 2013;62(12):1704-13
43. Guo CG, Ji R, Li YQ. Accuracy of i-scan for optical diagnosis of colonic polyps: A meta-analysis. *PLoS ONE [Electronic Resource]* 2015;10 (5) (no pagination)(0126237).
44. Harbord R, Whiting P. metandi: Meta–analysis of diagnostic accuracy using hierarchical logistic regression. *The Stata Journal* 2009(2):18.
45. Takwoingi Y. Meta-analysis of test accuracy studies in Stata -a bivariate model approach. Version 1.1 Aprl 2016. 2016. http://methods.cochrane.org/sdt/. (accessed 5/7/2016).
46. Review Manager (RevMan) [Computer program] [program]. 5.3.5 version. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
47. Kang HY, Kim YS, Song JH, Yang SY, Lim SH. Comparison of narrow band imaging (NBI) and fujinon intelligent color enhancement (FICE) in predicting small colorectal polyp histology. *Journal of Gastroenterology and Hepatology (Australia)* 2014;29:81.
48. Paggi S, Amato A, Rondonotti E, Spinzi G, Radaelli F. Do new generation narrow band imaging colonoscopes improve the prediction of histology for diminutive polyps? Preliminary results of an observational prospective study. *Gastrointestinal Endoscopy* 2014;1):AB368.
49. Paggi S, Rondonotti E, Amato A, Spinzi G, Radaelli F. Narrow band imaging in the prediction of surveillance intervals after polypectomy in community practice: Ready for (a European) prime time. *Gastrointestinal Endoscopy* 2014;1):AB164.
50. Patel S, Schoenfeld P, Kim H, Kim Y, Ward E, Bansal A, et al. Learning Curves for the Real-Time Histologic Characterization of Diminutive Colorectal Polyps With Narrow Band Imaging (NBI) Using Cumulative Summation Analysis (CUSUM): Implications for Resect and Discard Strategy. *American Journal of Gastroenterology* 2015;110:S598-S99.
51. Patel SG, Rastogi A, Schoenfeld P, Bansal A, Hosford L, Myers A, et al. Cost-savings associated with the resect and discard strategy for diminutive polyps: Results from a prospective multicenter study evaluating real-time characterization of diminutive colorectal polyp histology using narrow band imaging (NBI). *Gastrointestinal Endoscopy* 2016;1):AB421.
52. Patel SG, Schoenfeld PS, Ward EK, Kim H, Bansal A, Hosford L, et al. A prospective multicenter study evaluating real-time characterization of diminutive colorectal polyp histology using narrow band imaging (NBI): Implications for the resect and discard strategy. *Gastrointestinal Endoscopy* 2015;1):AB149.
53. Szura M, Bucki K, Pasternak A, Matyja A, Kulig J. Two-stage optical system for colorectal polyp assessment. *Surgical Endoscopy and Other Interventional Techniques* 2014;28:336.

54. Aihara H, Kumar N, Ryou M, Burakoff R, Gergi MA, Ryan MB, et al. Prospective evaluation of a simplified narrowband imaging scoring system for a differential diagnosis of colorectal lesions. *Surgical Endoscopy* 2016;30(8):3598-603.
55. Chandran S, Parker F, Lontos S, Vaughan R, Efthymiou M. Can we ease the financial burden of colonoscopy? Using real-time endoscopic assessment of polyp histology to predict surveillance intervals. *Internal Medicine Journal* 2015;45(12):1293-9.
56. Gupta N, Bansal A, Rao D, Early DS, Jonnalagadda S, Edmundowicz SA, et al. Accuracy of in vivo optical diagnosis of colon polyp histology by narrow-band imaging in predicting colonoscopy surveillance intervals. *Gastrointestinal Endoscopy* 2012;75(3):494-502.
57. Henry ZH, Yeaton P, Shami VM, Kahaleh M, Patrie JT, Cox DG, et al. Meshed capillary vessels found on narrow-band imaging without optical magnification effectively identifies colorectal neoplasia: a North American validation of the Japanese experience. *Gastrointestinal Endoscopy* 2010;72(1):118-26.
58. Hewett DG, Huffman ME, Rex DK. Leaving distal colorectal hyperplastic polyps in place can be achieved with high accuracy by using narrow-band imaging: an observational study. *Gastrointestinal Endoscopy* 2012;76(2):374-80.
59. Ignjatovic A, East JE, Suzuki N, Vance M, Guenther T, Saunders BP. Optical diagnosis of small colorectal polyps at routine colonoscopy (Detect InSpect ChAracterise Resect and Discard; DISCARD trial): a prospective cohort study. *Lancet Oncology* 2009;10(12):1171-8.
60. Ikematsu H, Matsuda T, Osera S, Imajoh M, Kadota T, Morimoto H, et al. Usefulness of narrow-band imaging with dual-focus magnification for differential diagnosis of small colorectal polyps. *Surgical Endoscopy* 2015;29(4):844-50.
61. Iwatate M, Sano Y, Hattori S, Sano W, Hasuike N, Ikumoto T, et al. The addition of high magnifying endoscopy improves rates of high confidence optical diagnosis of colorectal polyps. *Endoscopy International Open* 2015;3(2):E140-5.
62. Kaltenbach T, Rastogi A, Rouse RV, McQuaid KR, Sato T, Bansal A, et al. Real-time optical diagnosis for diminutive colorectal polyps using narrow-band imaging: the VALID randomised clinical trial. *Gut* 2015;64(10):1569-77.
63. Ladabaum U, Fioritto A, Mitani A, Desai M, Kim JP, Rex DK, et al. Real-time optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions. *Gastroenterology* 2013;144(1):81-91.
64. McGill SK, Soetikno R, Rastogi A, Rouse RV, Sato T, Bansal A, et al. Endoscopists can sustain high performance for the optical diagnosis of colorectal polyps following standardized and continued training. *Endoscopy* 2015;47(3):200-6.

65. Paggi S, Rondonotti E, Amato A, Fuccio L, Andrealli A, Spinzi G, et al. Narrow-band imaging in the prediction of surveillance intervals after polypectomy in community practice. *Endoscopy* 2015;47(9):808-14.
66. Paggi S, Rondonotti E, Amato A, Terruzzi V, Imperiali G, Mandelli G, et al. Resect and discard strategy in clinical practice: a prospective cohort study. *Endoscopy* 2012;44(10):899-904.
67. Patel SG, Schoenfeld P, Kim HM, Ward EK, Bansal A, Kim Y, et al. Real-Time Characterization of Diminutive Colorectal Polyp Histology Using Narrow-Band Imaging: Implications for the Resect and Discard Strategy. *Gastroenterology* 2016;150(2):406-18.
68. Pohl H, Bensen SP, Toor A, Gordon SR, Levy LC, Anderson PB, et al. Quality of optical diagnosis of diminutive polyps and associated factors. *Endoscopy* 2016.
69. Rastogi A, Keighley J, Singh V, Callahan P, Bansal A, Wani S, et al. High accuracy of narrow band imaging without magnification for the real-time characterization of polyp histology and its comparison with high-definition white light colonoscopy: a prospective study. *American Journal of Gastroenterology* 2009;104(10):2422-30.
70. Repici A, Hassan C, Radaelli F, Occhipinti P, De Angelis C, Romeo F, et al. Accuracy of narrow-band imaging in predicting colonoscopy surveillance intervals and histology of distal diminutive polyps: results from a multicenter, prospective trial. *Gastrointestinal Endoscopy* 2013;78(1):106-14.
71. Rex DK. Narrow-band imaging without optical magnification for histologic analysis of colorectal polyps. *Gastroenterology* 2009;136(4):1174-81.
72. Rogart JN, Jain D, Siddiqui UD, Oren T, Lim J, Jamidar P, et al. Narrow-band imaging without high magnification to differentiate polyps during real-time colonoscopy: improvement with experience. *Gastrointestinal Endoscopy* 2008;68(6):1136-45.
73. Shahid MW, Buchner AM, Heckman MG, Krishna M, Raimondo M, Woodward T, et al. Diagnostic accuracy of probe-based confocal laser endomicroscopy and narrow band imaging for small colorectal polyps: a feasibility study. *American Journal of Gastroenterology* 2012;107(2):231-9.
74. Sola-Vera J, Cuesta R, Uceda F, Morillo E, Perez E, Pico MD, et al. Accuracy for optical diagnosis of colorectal polyps in clinical practice. *Revista Espanola de Enfermedades Digestivas* 2015;107(5):255-61.
75. Vu HT, Sayuk GS, Hollander TG, Clebanoff J, Edmundowicz SA, Gyawali CP, et al. Resect and Discard Approach to Colon Polyps: Real-World Applicability Among Academic and Community Gastroenterologists. *Digestive Diseases and Sciences* 2014;60(2):502-08.
76. Wallace MB, Crook JE, Coe S, Ussui V, Staggs E, Almansa C, et al. Accuracy of in vivo colorectal polyp discrimination by using dual-focus high-definition narrow-band imaging colonoscopy. *Gastrointestinal Endoscopy* 2014;80(6):1072-87.

77. Lee CK, Lee SH, Hwangbo Y. Narrow-band imaging versus I-Scan for the real-time histological prediction of diminutive colonic polyps: a prospective comparative study by using the simple unified endoscopic classification. *Gastrointestinal Endoscopy* 2011;74(3):603-9.
78. Kang HY, Kim YS, Kang SJ, Chung GE, Song JH, Yang SY, et al. Comparison of Narrow Band Imaging and Fujinon Intelligent Color Enhancement in Predicting Small Colorectal Polyp Histology. *Digestive Diseases & Sciences* 2015;60(9):2777-84.
79. Basford PJ, Longcroft-Wheaton G, Higgins B, Bhandari P. High-definition endoscopy with i-Scan for evaluation of small colon polyps: the HiSCOPE study. *Gastrointestinal Endoscopy* 2014;79(1):111-8.
80. Hoffman A, Kagel C, Goetz M, Tresch A, Mudter J, Biesterfeld S, et al. Recognition and characterization of small colonic neoplasia with high-definition colonoscopy using i-Scan is as precise as chromoendoscopy. *Digestive & Liver Disease* 2010;42(1):45-50.
81. Pigo F, Bertani H, Manno M, Mirante V, Caruso A, Barbera C, et al. I-Scan high-definition white light endoscopy and colorectal polyps: Prediction of histology, interobserver and intraobserver agreement. *International Journal of Colorectal Disease* 2013;28(3):399-406.
82. Rath T, Tontini GE, Naegel A, Vieth M, Zopf S, Guenther C, et al. High-definition endoscopy with digital chromoendoscopy for histologic prediction of distal colorectal polyps. *BMC Gastroenterology* 2015;15.
83. Longcroft-Wheaton G, Brown J, Cowlishaw D, Higgins B, Bhandari P. High-definition vs. standard-definition colonoscopy in the characterization of small colonic polyps: results from a randomized trial. *Endoscopy* 2012;44(10):905-10.
84. Longcroft-Wheaton GR, Higgins B, Bhandari P. Flexible spectral imaging color enhancement and indigo carmine in neoplasia diagnosis during colonoscopy: a large prospective UK series (Structured abstract). *European Journal of Gastroenterology and Hepatology* 2011;23(10):903-11.
85. Rastogi A, Pondugula K, Bansal A, Wani S, Keighley J, Sugar J, et al. Recognition of surface mucosal and vascular patterns of colon polyps by using narrow-band imaging: interobserver and intraobserver agreement and prediction of polyp histology. *Gastrointestinal Endoscopy* 2009;69(3 Pt 2):716-22.
86. Rastogi A, Bansal A, Wani S, Callahan P, McGregor DH, Cherian R, et al. Narrow-band imaging colonoscopy-a pilot feasibility study for the detection of polyps and correlation of surface patterns with polyp histologic diagnosis. *Gastrointestinal Endoscopy* 2008;67(2):280-86.
87. Rastogi A, Rao DS, Gupta N, Grisolano SW, Buckles DC, Sidorenko E, et al. Impact of a computer-based teaching module on characterization of diminutive colon polyps by using narrow-band imaging by non-experts in academic and community practice: a video-based study. *Gastrointestinal Endoscopy* 2014;79(3):390-8.

88. Rastogi A, Early DS, Gupta N, Bansal A, Singh V, Ansstas M, et al. Randomized, controlled trial of standard-definition white-light, high-definition white-light, and narrow-band imaging colonoscopy for the detection of colon polyps and prediction of polyp histology. *Gastrointestinal Endoscopy* 2011;74(3):593-602.
89. East JE, Suzuki N, Saunders BP. Comparison of magnified pit pattern interpretation with narrow band imaging versus chromoendoscopy for diminutive colonic polyps: a pilot study. *Gastrointestinal Endoscopy* 2007;66(2):310-6.
90. Raghavendra M, Hewett DG, Rex DK. Differentiating adenomas from hyperplastic colorectal polyps: narrow-band imaging can be learned in 20 minutes. *Gastrointestinal Endoscopy* 2010;72(3):572-6.
91. Sano Y, Ikematsu H, Fu KI, Emura F, Katagiri A, Horimatsu T, et al. Meshed capillary vessels by use of narrow-band imaging for differential diagnosis of small colorectal polyps. *Gastrointestinal Endoscopy* 2009;69(2):278-83.
92. Sano Y, Emura F, Ikematsu H. Narrow-band Imaging. In: Waye JD, Rex DK, Williams CB (eds.) *Colonoscopy principles and practice*. 2nd ed. Hoboken: Wiley-Blackwell; 2009 p514-26.
93. Hayashi N, Tanaka S, Hewett DG, Kaltenbach TR, Sano Y, Ponchon T, et al. Endoscopic prediction of deep submucosal invasive carcinoma: validation of the narrow-band imaging international colorectal endoscopic (NICE) classification. *Gastrointestinal Endoscopy* 2013;78(4):625-32.
94. East JE, Suzuki N, Bassett P, Stavrinidis M, Thomas HJ, Guenther T, et al. Narrow band imaging with magnification for the characterization of small and diminutive colonic polyps: pit pattern and vascular pattern intensity. *Endoscopy* 2008;40(10):811-7.
95. Kudo S, Hirota S, Nakajima T, Hosobe S, Kusaka H, Kobayashi T, et al. Colorectal tumours and pit pattern. *Journal of Clinical Pathology* 1994;47(10):880-85.
96. Oba S, Tanaka S, Sano Y, Oka S, Chayama K. Current status of narrow-band imaging magnifying colonoscopy for colorectal neoplasia in Japan. *Digestion* 2011;83(3):167-72.
97. Sano W, Sano Y, Iwatate M, Hasuike N, Hattori S, Kosaka H, et al. Prospective evaluation of the proportion of sessile serrated adenoma/polyps in endoscopically diagnosed colorectal polyps with hyperplastic features. *Endoscopy International Open* 2015;3(4):E354-8.
98. Kim YS, Kim D, Chung SJ, Park MJ, Shin CS, Cho SH, et al. Differentiating small polyp histologies using real-time screening colonoscopy with Fuji Intelligent Color Enhancement. *Clinical Gastroenterology & Hepatology* 2011;9(9):744-49.e1.
99. Teixeira CR, Torresini RS, Canali C, Figueiredo LF, Mucenic M, Pereira Lima JC, et al. Endoscopic classification of the capillary-vessel pattern of colorectal lesions by spectral estimation technology and magnifying zoom imaging. *Gastrointestinal Endoscopy* 2009;69(3 Pt 2):750-6.

100. Cancer Council Australia Colonoscopy Surveillance Working Party. Clinical Practice Guidelines for Surveillance Colonoscopy – in adenoma follow-up; following curative resection of colorectal cancer; and for cancer surveillance in inflammatory bowel disease., 2011.
101. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* 2008;134(5):1570-95.
102. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 2012;143(3):844-57.
103. Winawer SJ, Zauber AG, Fletcher RH, Stillman JS, O'Brien MJ, Levin B, et al. Guidelines for Colonoscopy Surveillance After Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *Gastroenterology* 2006;130(6):1872-85.
104. Rex DK, Ahnen DJ, Baron JA, Batts KP, Burke CA, Burt RW, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. *The American journal of gastroenterology* 2012;107(9):1315-29.
105. European Commission. European Guidelines for Quality assurance in Colorectal Cancer Screening and Diagnosis - First Edition. In: Segnan NP, J;von Karsa,L., editor. Luxembourg: Publications Office of the European Union, 2010.
106. Atkin WS, Valori R, Kuipers EJ, Hoff G, Senore C, Segnan N, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. *Endoscopy* 2012;44 Suppl 3:Se151-63.
107. Hassan C, Quintero E, Dumonceau JM, Regula J, Brandao C, Chaussade S, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2013;45(10):842-51.
108. Coe SG, Thomas C, Crook J, Ussui V, Diehl N, Wallace MB. Colorectal surveillance interval assignment based on in vivo prediction of polyp histology: impact of endoscopic quality improvement program. *Gastrointestinal Endoscopy* 2012;76(1):118-25.e1.
109. Davila RE, Rajan E, Baron TH, Adler DG, Egan JV, Faigel DO, et al. ASGE guideline: colorectal cancer screening and surveillance. *Gastrointestinal Endoscopy* 2006;63(4):546-57.
110. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313(7052):275-83.

111. Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. *Clinical Gastroenterology & Hepatology* 2010;8(10):865-9, 69.e1-3.
112. Kessler WR, Imperiale TF, Klein RW, Wielage RC, Rex DK. A quantitative assessment of the risks and cost savings of forgoing histologic examination of diminutive polyps. *Endoscopy* 2011;43(8):683-91.
113. Longcroft-Wheaton G, Bhandari P. The cost impact of in vivo diagnosis of diminutive polyps: Experience from a screening endoscopy programme. *Gut* 2011;60:A30.
114. Longcroft-Wheaton G, Bhandari P. THE COST IMPACT OF IN VIVO DIAGNOSIS OF DIMINUTIVE POLYPS: EXPERIENCE FROM A SCREENING ENDOSCOPY PROGRAMME. *Gut* 2011;60:A30-A30.
115. McGill SK, Soetikno RM, Yokomizo L, Goldhaber-Fiebert JD, Owens D, Kaltenbach T. Optical diagnosis of small colorectal polyps with resect and discard strategy is cost saving. *Gastrointestinal Endoscopy* 2013;1):AB168.
116. Solon C, Klausnitzer R, Blissett D, Ihara Z. Economic value of narrow band imaging versus white light endoscopy for the characterization of diminutive polyps in the colon: systematic literature review and cost-consequence model. *J Med Econ* 2016:1-27.
117. Ness RM, Holmes AM, Klein R, Dittus R. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. *American Journal of Gastroenterology* 2000;95(7):1800-11.
118. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. *NICE, London* 2013.
119. Rastogi A, Bansal A, Rao DS, Gupta N, Wani SB, Shipe T, et al. Higher adenoma detection rates with cap-assisted colonoscopy: a randomised controlled trial. *Gut* 2012;61(3):402-8.
120. Butterly LF, Chase MP, Pohl H, Fiarman GS. Prevalence of clinically important histology in small adenomas. *Clinical Gastroenterology & Hepatology* 2006;4(3):343-8.
121. Whyte S, Chilcott J, Halloran S. Reappraisal of the options for colorectal cancer screening in England. *Colorectal Disease* 2012;14(9):e547-61.
122. Department of Health. NHS reference costs 2013-2014. *http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_122803* 2015.
123. Curtis L. Unit Costs of Health and Social Care 2015: Personal Social Services Research Unit, 2015.
124. Sharara N, Adam V, Crott R, Barkun AN. The costs of colonoscopy in a Canadian hospital using a microcosting approach. *Canadian Journal of Gastroenterology* 2008;22(6):565-70.

125. Bisschops R, Tejpar S, Willekens H, Hertogh G, Cutsem E. I-SCAN detects more polyps in lynch syndrome (HNPCC) patients: A prospective controlled randomized back-to-back study. *Gastrointestinal Endoscopy* 2012;75(4 suppl. 1):Ab330.
126. Department of Health. Public health functions to be exercised by NHS England: Service specification No.26, Bowel Cancer Screening Programme: Department of Health, 2013.
127. Whyte S, Chilcott J, Cooper K, Essat M, Stevens J, Wong R, et al. Re-appraisal of the options for colorectal cancer screening FULL REPORT: School of Health and Related Research (ScHARR), 2011:58.
128. Raju GS, Vadyala V, Slack R, Krishna SG, Ross WA, Lynch PM, et al. Adenoma detection in patients undergoing a comprehensive colonoscopy screening. *Cancer Med* 2013;2(3):391-402.
129. XY Scan [program]. 4.1 version, 2016.
130. Lee TJ, Nickerson C, Goddard AF, Rees CJ, McNally RJ, Rutter MD. Outcome of 12-month surveillance colonoscopy in high-risk patients in the National Health Service Bowel Cancer Screening Programme. *Colorectal Disease* 2013;15(8):e435-42.
131. Martinez ME, Baron JA, Lieberman DA, Schatzkin A, Lanza E, Winawer SJ, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology* 2009;136(3):832-41.
132. Laiyemo AO, Pinsky PF, Marcus PM, Lanza E, Cross AJ, Schatzkin A, et al. Utilization and yield of surveillance colonoscopy in the continued follow-up study of the polyp prevention trial. *Clinical Gastroenterology & Hepatology* 2009;7(5):562-7; quiz 497.
133. Saini SD, Kim HM, Schoenfeld P. Incidence of advanced adenomas at surveillance colonoscopy in patients with a personal history of colon adenomas: a meta-analysis and systematic review. *Gastrointestinal Endoscopy* 2006;64(4):614-26.
134. McDonald PJ, Digby J, Innes C, Strachan JA, Carey FA, Steele RJ, et al. Low faecal haemoglobin concentration potentially rules out significant colorectal disease. *Colorectal Disease* 2013;15(3):e151-9.
135. Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR, Neugut AI. Risk of perforation after colonoscopy and sigmoidoscopy: a population-based study. *J Natl Cancer Inst* 2003;95(3):230-6.
136. Atkin WS. Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial. *The Lancet* 2002;359(9314):1291-300.
137. Pilgrim HTP, Chilcott J, Bending M, Trueman P, Shorthouse AJ, Tappenden J. The costs and benefits of bowel cancer service developments using discrete event simulation. *Journal of the Operational Research Society* 2009:1305-14.
138. Ignjatovic A, Thomas-Gibson S, East JE, Haycock A, Bassett P, Bhandari P, et al. Development and validation of a training module on the use of narrow-band imaging in differentiation of

small adenomas from hyperplastic colorectal polyps. *Gastrointestinal Endoscopy* 2011;73(1):128-33.

139. Mead RJ, Duku M, Longcroft-Wheaton G, Pearl D, Basford P, Bhandari P. Endoscopic training: Colon polyp assessment. *Gut* 2011;60:A121-A22.
140. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;14(4):539-45.
141. Djalalov S, Rabeneck L, Tomlinson G, Bremner KE, Hilsden R, Hoch JS. A Review and Meta-analysis of Colorectal Cancer Utilities. *Med Decis Making* 2014;34(6):809-18.
142. Farkkila N, Sintonen H, Saarto T, Jarvinen H, Hanninen J, Taari K, et al. Health-related quality of life in colorectal cancer. *Colorectal Disease* 2013;15(5):e215-22.
143. Downing A, Morris EJ, Richards M, Corner J, Wright P, Sebag-Montefiore D, et al. Health-related quality of life after colorectal cancer in England: a patient-reported outcomes study of individuals 12 to 36 months after diagnosis. *Journal of Clinical Oncology* 2015;33(6):616-24.
144. Dorian P, Kongnakorn T, Phatak H, Rublee DA, Kuznik A, Lanitis T, et al. Cost-effectiveness of apixaban vs. current standard of care for stroke prevention in patients with atrial fibrillation. *Eur Heart J* 2014;35(28):1897-906.
145. Rees CJ, Rajasekhar PT, Wilson A, Close H, Rutter MD, Saunders BP, et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. *Gut* 2016.
146. Abu Dayyeh BK, Thosani N, Konda V, Wallace MB, Rex DK, Chauhan SS, et al. ASGE technology committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointestinal Endoscopy* 2015;81(3):502-16.

10 APPENDICES – SEE SEPARATE DOCUMENT

Appendix 1 Search strategy

Appendix 2 Study selection worksheet

Appendix 3 Data extraction tables

Appendix 4 Table of excluded studies with rationale

Appendix 5 Ongoing studies

Appendix 6 Studies excluded from the systematic review of cost-effectiveness studies

Appendix 7 Data extraction forms of included economic evaluations

Appendix 8 Data extraction of company's economic evaluation

Appendix 9 Parameters and distributions used in the probabilistic sensitivity analysis

Appendix 10 Derivation of the distribution of adenomas in patients undergoing colonoscopy

Appendix 11 System costs (scope, system, maintenance)