Faecal calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation

N Waugh, E Cummins, P Royle, N-B Kandala, D Shyangdan, R Arasaradnam, C Clar and R Johnston

Faecal calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation

N Waugh,[1]* E Cummins,[2] P Royle,[1] N-B Kandala,[1] D Shyangdan,[1] R Arasaradnam,[3] C Clar[1] and R Johnston[2]

[1]Warwick Evidence, Division of Health Sciences, Warwick Medical School, Coventry, UK
[2]McMDC Ltd, Glasgow, UK
[3]University Hospital Coventry and Warwickshire, Coventry, UK

*Corresponding author

Health Technology Assessment

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

Abstract

Faecal calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation

N Waugh,[1]* E Cummins,[2] P Royle,[1] N-B Kandala,[1] D Shyangdan,[1] R Arasaradnam,[3] C Clar[1] and R Johnston[2]

[1]Warwick Evidence, Division of Health Sciences, Warwick Medical School, Coventry, UK
[2]McMDC Ltd, Glasgow, UK
[3]University Hospital Coventry and Warwickshire, Coventry, UK

*Corresponding author

Background: Irritable bowel syndrome (IBS) is common, and causes pain, bloating and diarrhoea and/or constipation. It is a troublesome condition that reduces the quality of life but causes no permanent damage. Inflammatory bowel disease (IBD) comprises mainly ulcerative colitis (UC) and Crohn's disease (CD). Both cause serious complications and may lead to sections of the bowel having to be removed, although this is more common with CD. The presenting symptoms of IBS and IBD can be similar. Distinguishing them on clinical signs and symptoms can be difficult. Until recently, colonoscopy was often required to rule out IBD. In younger people, > 60% of colonoscopies showed no abnormality. Faecal calprotectin (FC) is a protein released by the white blood cells, neutrophils, found in inflamed areas of the bowel in IBD. Determining the level of FC in stool samples may help distinguish IBS from IBD.

Objective: To review the value of FC for distinguishing between IBD and non-IBD.

Data sources: Sources included MEDLINE, EMBASE, The Cochrane Library, Web of Science, websites of journals and the European Crohn's and Colitis Organisation (conference abstracts 2012 and 2013), and contact with experts.

Review methods: Systematic review and economic modelling. Review Manager (RevMan) version 5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) was used for most analysis, with statistical analyses done in Stata version 12 (StataCorp LP, College Station, TX, USA). Forest plots and receiver operating characteristic curves were produced. Quality Assessment of Diagnostic Accuracy Studies was used for quality assessment. Economic modelling was done in Microsoft Excel 2010 (Microsoft Corporation, Redmond, WA, USA).

Limitations: Studies were often small, most used only one calprotectin cut-off level, and nearly all came from secondary care populations.

Results: Twenty-eight studies provided data for 2 × 2 tables and were included in meta-analyses, with seven in the most important comparison in adults (IBS vs. IBD) and eight in the key comparison in paediatrics (IBD vs. non-IBD). Most studies used laboratory enzyme-linked immunosorbent assay (ELISA) tests. For distinguishing between IBD and IBS in adults, these gave pooled sensitivity of 93% and specificity of 94% at FC cut-off level of 50 μg/g. Sensitivities at that cut-off ranged from 83% to 100%, and specificities from 60% to 100%. For distinguishing between IBD and non-IBD in paediatric populations with ELISA tests, sensitivities ranged from 95% to 100% at cut-off of 50 μg/g and specificities of 44–93%. Few studies used point-of-care testing but that seemed as reliable as ELISA, though perhaps less specific.

The evidence did not provide any grounds for preferring one test over others on clinical effectiveness grounds. FC testing in primary care could reduce the need for referral and colonoscopies. Any quality-adjusted life-year gains are likely to be small because of the low prevalence of IBD and the high sensitivities of all of the tests, resulting in few false negatives with IBD. However, considerable savings could accrue. Areas of uncertainty include the optimum management of people with borderline results (50–150 µg/g), most of whom do not have IBD. Repeat testing may be appropriate before referral.

Conclusions: Faecal calprotectin can be a highly sensitive way of detecting IBD, although there are inevitably trade-offs between sensitivity and specificity, with some false positives (IBS with positive calprotectin) if a low calprotectin cut-off is used. In most cases, a negative calprotectin rules out IBD, thereby sparing most people with IBS from having to have invasive investigations, such as colonoscopy.

Study registration: PROSPERO CRD 42012003287.

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

Contents

Glossary

Accuracy Accuracy is the probability that the test yields a correct result: (true positive + true negative)/ (positive + negative).

Bloating Fullness or swelling in the abdomen that often occurs after meals.

Constipation A condition in which bowel movements are infrequent, hard and dry, and elimination of faeces is difficult and infrequent.

Cost impact The total cost to the person, the National Health Service or to society.

Cost minimisation analysis A type of economic evaluation used to compare the difference in costs between programmes that have the same health outcome.

Cost-effectiveness analysis An economic study design in which alternative interventions are compared in terms of cost per unit of effectiveness.

Cost-effectiveness model An explicit mathematical framework which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.

Cost-effectiveness The cost per unit of benefit of an intervention. Benefits of different interventions are measured using a single outcome (e.g. life-years gained, quality-adjusted life-years gained, deaths avoided, heart attacks avoided, cases detected).

Cost-of-illness/economic burden studies An analysis of the total costs incurred by a society due to a specific disease.

Cost–consequences analysis A type of economic evaluation, whereby both outcomes and costs of alternative interventions are described, without any attempt to combine the results.

Cost–utility analysis A form of cost-effectiveness analysis in which the units of effectiveness are quality-adjusted life-years.

Crohn's disease A chronic inflammatory disease of the digestive tract that can involve any part of it – from the mouth to the anus. It typically affects the terminal ileum as well as demarcated areas of large bowel, with other areas of the bowel being relatively unaffected. It is often associated with autoimmune disorders outside the bowel, such as rheumatoid arthritis.

Diagnostic odds ratio It is defined as the ratio of the odds of the test being positive if the subject has a disease relative to the odds of the test being positive if the subject does not have the disease.

Discounting Discounting is a method for adjusting the value of costs and outcomes that occur in different time periods into a common time period, usually the present.

Dominance An intervention is said to be dominated if there is an alternative intervention that is both less costly and more effective.

Fagan's nomogram A graph that uses pre-test probability of inflammatory bowel disease and likelihood ratios to estimate the probability of a patient with a positive test having the condition. Examples are shown on pages 39–55.

False negative Incorrect negative test result – number of diseased persons with a negative test result.

False positive Incorrect positive test result – number of non-diseased persons with a positive test result.

Functional bowel disorder In medicine, the term 'functional bowel disorder' refers to a group of disorders that are characterised by chronic abdominal complaints without a structural or biochemical cause that could explain symptoms.

Incremental cost-effectiveness ratio The difference in the mean costs of two interventions in the population of interest divided by the difference in the mean outcomes in the population of interest.

Index test The test of which performance is being evaluated.

Inflammatory bowel disease General term for any disease characterised by inflammation of the bowel. Two of the most common inflammatory bowel diseases are ulcerative colitis and Crohn's disease.

Likelihood ratios Likelihood ratios combine the information from sensitivity and specificity. The likelihood ratio for a positive test is the probability of an individual with inflammatory bowel disease having a positive test (sensitivity) divided by the probability of an individual without inflammatory bowel disease having a positive test (1 minus specificity). The likelihood ratio for a negative test is the probability of an individual with inflammatory bowel disease having a negative test divided by the probability of an individual without inflammatory bowel disease having a negative test. So, the likelihood ratio for a negative test = 1 − sensitivity/specificity. In those with a positive test, likelihood ratio positive values of ≥ 10 are usually regarded as strong evidence of a disease being present.

Meta-analysis Statistical techniques used to combine the results of two or more studies and obtain a combined estimate of effect.

Negative predictive value The negative predictive value of a calprotectin test is the probability that a patient with a negative calprotectin test does not have inflammatory bowel disease.

Positive predictive value The positive predictive value is defined as the probability that a patient with a positive calprotectin test has inflammatory bowel disease.

Probabilistic sensitivity analysis Probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (e.g. Monte Carlo simulation).

Quality of life An individual's emotional, social and physical well-being, and his or her ability to perform the ordinary tasks of living.

Quality-adjusted life-years An index of survival that is adjusted to account for the patient's quality of life during this time. Quality-adjusted life-years have the advantage of incorporating changes in both quantity (longevity/mortality) and quality (morbidity, psychological, functional, social and other factors) of life. Used to measure benefits in cost–utility analysis. The quality-adjusted life-years gained are the mean quality-adjusted life-years associated with one treatment minus the mean quality-adjusted life-years associated with an alternative treatment.

Receiver operating characteristic curve A graph that illustrates the trade-offs between sensitivity and specificity, which result from varying the diagnostic threshold.

Reference standard The best currently available diagnostic test(s), against which the index test is compared.

Sensitivity (of a test) The proportion of individuals classified as positive by the gold (or reference) standard, who are correctly identified by the study test.

Sensitivity analysis A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multiway simple sensitivity analysis (scenario analysis): two or more parameters are varied at the same time, and the overall effect on the results is evaluated.

Specificity (of a test) The proportion of individuals classified as negative by the gold (or reference) standard, who are correctly identified by the study test.

Threshold sensitivity analysis The critical value of parameters above or below which the conclusions of the study will change are identified.

True negative Correct negative test result – number of non-diseased persons with a negative test result.

True positive Correct positive test result – number of diseased persons with a positive test result.

Utility A measure of the strength of an individual's preference for a specific health state in relation to alternative health states. The utility scale assigns numerical values on a scale from 0 (death) to 1 (optimal or 'perfect' health). Health states can be considered worse than death and thus have a negative value.

Visceral hypersensitivity Enhanced perception or enhanced responsiveness within the gut.

List of abbreviations

6-MP	6-mercaptopurine	FOBT	faecal occult blood testing
AQoL	Assessment of Quality of Life	GI	gastrointestinal
ASA	aminosalicylic acid	GP	general practitioner
ASCA	anti-*Saccharomyces cerevisiae* antibodies	HASA	high-dose aminosalicylic acid
AUC	area under the curve	HASB	high-dose aminosalicylic acid with beclometasone
BaFT	barium follow-through	HAST	high-dose aminosalicylic acid with a topical aminosalicylic acid
BSG	British Society of Gastroenterology		
BWUS	bowel wall ultrasonography measurement	Hb	haemoglobin
		HCSC	Hospital and Community Services Costs
CBT	cognitive behavioural therapy		
CD	Crohn's disease	HLA	human leucocyte antigen
CDAI	Crohn's Disease Activity Index	HODaR	Health Outcomes Data Repository
CE	conformité Européenne (European Conformity)	HRQoL	health-related quality of life
		HUI	health–utility index
CEAF	cost-effectiveness acceptability frontier	IBD	inflammatory bowel disease
		IBDQ	Inflammatory Bowel Disease Questionnaire
CEP	Centre for Evidence-based Purchasing		
		IBS	irritable bowel syndrome
CI	confidence interval	IBS-A	irritable bowel syndrome-alternating type
CRC	colorectal cancer		
CRP	C-reactive protein	IBS-C	irritable bowel syndrome-constipation
CT	computed tomography	IBS-D	irritable bowel syndrome-diarrhoea
DOR	diagnostic odds ratio	IBS-M	irritable bowel syndrome-mixed
ELISA	enzyme-linked immunosorbent assay	ICC	intraclass correlation coefficient
		ICER	incremental cost-effectiveness ratio
EQ-5D	European Quality of Life-5 Dimensions	IDDM	insulin-dependent diabetes mellitus
		IP	intestinal permeability
ERG	Evidence Review Group	IQR	interquartile range
ESR	erythrocyte sedimentation rate	LASA	low-dose aminosalicylic acid
FBC	full blood count	LR	likelihood ratio
FC	faecal calprotectin	MRI	magnetic resonance imaging
FDA	Food and Drug Administration	NA	not available
FL	faecal lactoferrin		

NICE	National Institute for Health and Care Excellence		SBFT	small bowel follow-through
NLR	negative likelihood ratio		SD	standard deviation
NPV	negative predictive value		SE	standard error
NSAID	non-steroidal anti-inflammatory drug		SF-36	Short Form questionnaire-36 items
NTAC	NHS Technology Adoption Centre		SF-6D	Short Form questionnaire-6 Dimensions
OBD	organic bowel disease		SG	standard gamble
OR	odds ratio		SROC	summary receiver operating characteristic
pANCA	perinuclear anti-neutrophil cytoplasmic antibodies		SSRI	selective serotonin reuptake inhibitor
PCDAI	Paediatric Crohn's Disease Activity Index		TA	technology appraisal
PIBD	paediatric inflammatory bowel disease		TCA	tricyclic antidepressant
			TNF	tumour necrosis factor
PLR	positive likelihood ratio		TRFIA	time-resolved fluorimetric immunoassay
POCT	point-of-care testing		TSH	thyroid-stimulating hormone
PPI	proton pump inhibitor		TTG	tissue transglutaminase (a test for coeliac disease)
PPV	positive predictive value			
PSA	probabilistic sensitivity analysis		TTO	time trade-off
QALY	quality-adjusted life-year		UC	ulcerative colitis
QoL	quality of life		UCAI	Ulcerative Colitis Activity Index
QUADAS	Quality Assessment of Diagnostic Accuracy Studies		UCDAI	Ulcerative Colitis Disease Activity Index
RCT	randomised controlled trial		VAS	visual analogue scale
ROC	receiver operating characteristic		YHEC	York Health Economics Consortium
RR	relative risk			

Scientific summary

Background

Lower abdominal symptoms – such as pain, diarrhoea and bloating – are common and are usually due to irritable bowel syndrome (IBS), a troublesome condition that interferes with activities of daily life but which does not have serious consequences. Around 10% of the population have symptoms suggestive of IBS, although only about half consult their general practitioners (GPs).

The symptoms of IBS can resemble those of inflammatory bowel disease (IBD), mainly Crohn's disease (CD) and ulcerative colitis (UC). These diseases have serious complications, including a high risk of complications requiring surgery and an increased risk of colorectal cancer.

However, the symptoms of IBD can be different in children, many of whom present with non-specific symptoms, such as mild abdominal discomfort, lethargy, weight loss or growth impairment. In a large UK and Ireland study, only 25% of children with CD presented with the usual triad of diarrhoea, abdominal pain and weight loss. Delays in diagnosis were common, with over one-quarter of patients with CD taking over 1 year to be diagnosed. About 25% of people with IBD develop it under the age of 17 years.

Irritable bowel syndrome is often diagnosed on the basis of signs and symptoms, without a need for further investigations, but distinction from IBD on clinical grounds is often not possible. Distinguishing between IBD and IBS has often required referral to specialist care for colonoscopy, an invasive and unpleasant investigation requiring sedation, usually carried out on a day case basis, at a cost of around £650 (including specialist referral and day case endoscopy). Some centres have reported that > 60% of colonoscopies in younger patients have been normal, and in retrospect, not necessary.

Calprotectin is a protein released by the white blood cells involved in inflammation of the bowel. It is stable in faeces and can be measured by laboratory tests, and more recently by 'point-of-care testing' (POCT). It indicates inflammation in the bowel.

This review examines the clinical effectiveness and cost-effectiveness of FC testing in helping to distinguish between 'functional' disorders, such as IBS and 'organic' disorders, such as IBD. In adults, the differentiation is most often between IBS and IBD. In children, there is a different range of conditions.

Perspectives on the use of calprotectin testing will vary with setting. GPs will see far more cases of IBS than IBD, and for them calprotectin testing offers evidence to rule out IBD. A negative calprotectin will imply IBS. So GPs will be looking for parameters such as sensitivity (for IBD) and negative predictive value (NPV), to provide a basis for a decision not to refer. Gastroenterologists in adult clinics will be seeing a selected group of patients, referred by GPs, with a suspicion of IBD. Gastroenterologists will be looking for positive evidence of IBD in order to decide whether to proceed to further investigations, including colonoscopy and biopsy, and possibly also gastroscopy and other tests. They may find a positive predictive value or a positive diagnostic odds ratio more useful.

It should be noted that diagnosis will be made on the whole clinical picture, not on the basis of calprotectin results alone.

The same general principles will apply to the different case mix seen in paediatric gastroenterology. The proportion with IBD is higher, but a normal or near-normal calprotectin level may contribute to a decision not to proceed to invasive procedures such as endoscopy.

Methods

Systematic review and economic modelling. A broad search strategy was run in several databases. Studies that provided sufficient data for calculation of sensitivity, specificity and other diagnostic outcomes were identified. Review Manager (RevMan) version 5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) was used to generate paired forest plots and receiver operating characteristic (ROC) curves. Stata 12 (StataCorp LP, College Station, TX, USA) was used to produce likelihood ratios, areas under the curve (AUC) and nomograms. The quality of studies was assessed using Quality Assessment of Diagnostic Accuracy Studies. We sought studies in which the reference test was endoscopy with histology.

Results

Clinical effectiveness of calprotectin testing

The primary studies presented data for different groups of conditions, some providing a direct comparison of IBS and IBD, but others comparing a wider range of organic conditions.

Nearly all of the evidence comes from studies in specialist care.

Seven studies gave results that compared IBS and IBD, at eight cut-off levels, ranging from 8 to 150 µg/g, all in adults. Sensitivity was consistently high (usually 100% at levels of < 50 µg/g, ranging from 83% to 100% at a cut-off of 50 µg/g) but specificity was more varied (51% to 100%), especially at lower levels of FC.

Eleven studies, mostly from paediatrics, reported IBD versus non-IBD, with eight cut-off levels. They showed consistently high sensitivity at lower cut-offs: nearly all over 90%, with most at the 50 µg/g cut-off having sensitivities of 100%. Specificity was much more varied, ranging from 44% to 93% at a 50 µg/g cut-off.

Two reports by the York Health Economics Consortium (YHEC) were very useful. The first, from 2010, concluded that FC was a reliable marker of inflammation of the bowel, that high sensitivity was very important and that false positives were preferable to false negatives, that the cut-off should be 50 µg/g, and, in economic terms, that calprotectin dominated (more correct diagnoses at less cost) blood tests such as erythrocyte sedimentation rate and C-reactive protein.

The second YHEC report provided data on the use of calprotectin testing in routine care and on how it contributed to final diagnosis. One finding was that when GPs were sure a patient had IBS, they were usually right – 95% of such patients had normal calprotectin levels.

Choice of cut-off levels

The commonest level for defining normality was 50 µg/g. If sensitivity was deemed of paramount importance (in order not to miss any cases of IBD), that level could be used. Some adults with IBS have raised calprotectin levels and would be 'false positives', who might be referred for endoscopy as '?IBD'. However, there is some evidence that organic pathology is rare with levels of < 150 µg/g, and clinical consensus is that if there are adults with IBD but calprotectin levels of < 150 µg/g, they are likely to have low-grade IBD and would come to no harm if simply monitored by repeated calprotectins, with referral if the level rose.

In theory, a very sensitive approach might lead to people with IBS being false positives and undergoing endoscopy, and a less sensitive approach might mean missing a few people with IBD, with more serious consequences. In practice, clinicians will apply clinical nous and observation, and that will reduce colonoscopies in false positives. Decisions will not be made purely on FC results.

In paediatric age groups, with a different spectrum of conditions, a cut-off level of 50 μg/g gives almost 100% sensitivity but specificity varying from 44% to 94%. One study reported that a cut-off level of 100 μg/g gave sensitivity of only 86%, specificity 91%. Another study recommended a cut-off level of 200 μg/g as being most useful in routine practice.

Cost-effectiveness

NHS Technology Adoption Centre pilot study
Results from the pilot project show that calprotectin testing could reduce costs of referral and investigation of patients under 60 years with chronic diarrhoea by over 60%, if all patients with negative tests are managed in primary care as IBS, with those with borderline and positive tests being referred to gastroenterology.

This reduction is similar to the proportions of colonoscopies reported as normal from some other UK centres.

Review of previous studies
Previous economic analyses have typically concluded that calprotectin testing is cost saving compared with the situation without it. Given test specificities and the assumed prevalences of IBD in the presenting population, the additional cost of the calprotectin testing is more than offset by the reduction in the cost of unnecessary colonoscopies.

External Assessment Group: primary care
The External Assessment Group (EAG) developed a de novo cost-effectiveness model for the use of calprotectin testing for distinguishing between IBS and IBD in the primary care setting. This had an initial sequence of tests, with associated sensitivities and specificities, with positive results being referred to outpatient assessment and colonoscopy. Those testing positive were assumed to go on to an outpatient appointment and colonoscopy. Colonoscopy was associated with a slight risk of bleeds and perforation, with the latter having a very small mortality risk.

Subsequent to testing, patients could receive induction and maintenance treatment for IBS, CD and UC. False negatives could spend a period of time being unsuccessfully treated for IBS before re-presenting for testing.

A key uncertainty in the modelling was whether calprotectin testing would result in a wider group of patients being considered for testing than in the absence of calprotectin testing. This was explored through an alternative presenting population scenario analysis that doubled the number who would be tested compared with the number who would have been previously considered for referral in the absence of calprotectin testing.

The base case of the modelling assumed that without calprotectin testing all of those referred from primary care would go through an outpatient assessment and on to receive a colonoscopy.

Without calprotectin testing, GP clinical assessment can be highly sensitive in referring IBD. However, this may be at the cost of low specificity, with many 'false positives' (people with IBS) referred to gastroenterology. GPs without calprotectin testing might refer about 20%, most without IBD. The rates of false positives referred after calprotectin testing would be much lower: 5.1% and 5.6%, respectively.

Faecal calprotectin testing is estimated to result in cost savings. In theory, small quality-adjusted life-year (QALY) gains could accrue but these are too small to be significant, because of the low prevalence of IBD and the high sensitivities of all the tests, resulting in few false negatives with IBD. Sensitivity analyses suggest that calprotectin testing results in patient gains and remains cost saving compared with GP assessment without calprotectin testing, up to an IBD prevalence of 25%. At this point,

the less-than-perfect sensitivity ELISA testing results in very slight QALY losses compared with GP assessment without calprotectin testing, although cost savings of around £63 per patient on average remain.

The cost savings from calprotectin testing would be much reduced if the numbers tested were double those referred in the absence of calprotectin testing, resulting in slight cost savings or broad cost neutrality. Increases beyond a doubling would be likely to result in additional costs from calprotectin testing.

The savings from increased specificity of GP referral when given access to calprotectin testing depend largely on reduced colonoscopies. Scenario analysis shows that increasing the specificity of specialist assessment reduces the number of colonoscopies, with the cost savings from calprotectin testing falling. With a 95% specificity for outpatient assessment, the cost savings fell to around £10. Given lack of data, no modelling of repeat testing after indeterminate results was done. The impact of this on costs would mainly be determined by the calprotectin levels among patients with IBS who had an indeterminate result from their first test. If levels fell, the second test would result in fewer referrals and so could result in cost savings.

Secondary care

The model developed was also applied to differentiating IBD from non-IBD in the mainly paediatric secondary care setting.

Despite the higher IBD prevalence in the paediatric population, the main test differences still lie in the number of colonoscopies. Without calprotectin testing, all 52.1% of non-IBD patients receive a colonoscopy compared with 13.5% for the ELISA with the 50 μg/g cut-off, and only 9.4% for ELISA with the 100 μg/g cut-off.

The additional ELISA test costs are more than offset by the savings from reduced colonoscopies. Compared with all having a colonoscopy, ELISA with the 50 μg/g cut-off is estimated to save £205 on average, whereas ELISA with the 100 μg/g cut-off is estimated to save £240. Trivial QALY gains of around 0.001 QALYs may occur with ELISA compared with universal colonoscopy, these being slightly larger the 50 μg/g cut-off owing to its better sensitivity. But given the additional average £35 cost, the cost-effectiveness estimate using the 50 μg/g cut-off compared with the 100 μg/g cut-off is £35,000 per QALY. It should be stressed that the QALY differences between the strategies are very small and they may be better considered as equivalent.

Research needs

There is a lack of studies in primary care populations, and on the proportion of patients with lower bowel symptoms that would be tested if FC testing was available to GPs.

Many people have intermediate calprotectin levels (50–150 μg/g) and follow-up studies are required to determine the most useful cut-off level.

Some people with IBS have raised calprotectin levels. The reasons for that are not clear.

Conclusions

The National Institute for Health and Care Excellence (NICE) scope raised questions, abbreviated in italic text below:

Is calprotectin testing a reliable way of differentiating inflammatory diseases of the bowel from non-inflammatory ones?

Yes. The majority of younger adult patients seen with lower abdominal symptoms in general practice have IBS, and the absence of inflammation as indicated by a negative calprotectin test means that IBD is very unlikely. They could be managed in primary care and spared further investigations.

What are the optimal cut-offs for use in primary and secondary care?

The same cut-off should be used in primary and secondary care – currently 50 µg/g for ELISA tests but needing to be reviewed as evidence accumulates. This is based on ensuring high sensitivity, and not missing people with IBD. People with borderline levels of 50–150 µg/g could be monitored initially, with repeat calprotectin testing but some of this group will progress to definite IBD.

How do the rapid point-of-care tests compare with the laboratory tests?

There is currently insufficient evidence on either diagnostic reliability or cost-effectiveness considerations for preferring one test over another.

How will calprotectin testing perform in primary care?

Sensitivity and specificity will be as good in primary care but the lower prevalence will increase the NPV. The main benefit would be to confirm the clinical diagnosis of IBS by GPs. Making calprotectin testing available to GPs could reduce the number of younger adults referred to specialist care, and the need for unpleasant invasive investigations, such as colonoscopy.

Impact in secondary care?

In secondary care, calprotectin testing could considerably reduce the number of colonoscopies required. In various studies, over 60% of colonoscopies in this group of adult patients have been normal.

Calprotectin testing can also reduce the need for colonoscopy in children who do not have IBD, and could reduce diagnostic delays in those who do. It could also reduce loss of work time for parents and loss of school time for children.

Study registration

This study is registered as PROSPERO CRD 42012003287.

Funding

Chapter 1 Introduction

The conditions

Chronic abdominal pain or discomfort, accompanied by diarrhoea or constipation, is common. The symptoms can be due to a number of different conditions, some more serious than others. The conditions include irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD). The commonest forms of the latter are ulcerative colitis (UC) and Crohn's disease (CD, sometimes called regional ileitis, but that term is misleading because CD can have a much wider distribution).

Lower bowel symptoms are very common in general practice. Most patients have IBS, a troublesome and painful condition that reduces the quality of life (QoL) but which does not have serious effects in terms of structural damage to the bowel. However, some patients have IBD, which can lead to serious complications. Most patients with CD will require surgery within 5 years. It is important to distinguish IBD from IBS so that patients with the former can be appropriately managed and monitored. IBD is characterised by inflammation of the bowel, which is not seen in most patients with IBS.

Unfortunately, the symptoms of IBD and IBS are often similar, and, until recently, definitive diagnosis was often made only after invasive colonoscopy and perhaps other investigations. Faecal calprotectin (FC) testing identifies patients with inflammation of the bowel, who need referral to specialist care. The majority of younger patients with lower gastrointestinal (GI) symptoms have IBS, and if the absence of inflammation can be ruled out by a negative calprotectin test, they can then be managed in primary care and spared further investigations.

The most common symptoms of IBS include recurrent colicky abdominal pain or cramping felt in the lower abdomen and relieved by defecation. There may be abdominal distension (bloating) and altered bowel habit – episodes of diarrhoea and constipation. Features supporting a diagnosis of IBS include:

- symptoms > 6 months
- bloating
- associated with other, non-GI problems
- symptoms worsened by stress
- no weight loss.

The Rome criteria (Rome II,[1] Rome III[2]) subdivide IBS into diarrhoea predominant (IBS-D), constipation predominant (IBS-C) or mixed (IBS-M), with roughly one-third of patients in each group.

Irritable bowel syndrome is very common – affecting perhaps 15% of the UK population – although many people who have it never consult their general practitioners (GPs) about it. IBS-D is the commonest form.

It is commonest in young women, with an odds ratio (OR) in women to men of 1.7.[3] The IBS-C form is commoner in women than in men. The underlying mechanism is not known. People who have it are constitutionally well and do not lose weight. It is a troublesome but not a serious condition, in the sense that it does not lead to serious adverse events. But it can be painful and disruptive of normal activities, and people with IBS have a reduced QoL, reported to be reduced by 26%,[4] and 30% if severe.[5] QoL is reduced because of symptoms that disturb work and sleep, and anxiety. It leads to 9–22 lost days of work per year.[6] Akehurst et al.[7] report that in the Trent Region, people with IBS had reduced QoL compared with age-matched, sex-matched and socially matched controls, reflected in every dimension of both Short form-36 (SF-36) and European Quality of Life-5 Dimensions (EQ-5D). They had more time off work than people without IBS, and imposed £123 more costs per year on the NHS.[7] The effect on QoL depends on

severity of symptoms, with those meeting the Rome II criteria faring worse than those meeting Rome I criteria.[8]

The British Society of Gastroenterology (BSG) commissioning report noted:[9]

> *While IBS is not a life-threatening condition, it is a major cause of ill-health and disability, disrupting social activity and work. The large number of patients affected, the need to screen out other diseases, and absenteeism and impairment in the workplace all constitute a major cost to the health service and society at large.*

The cause of IBS is not known in most people but it sometimes follows an episode of infectious gastroenteritis ('food poisoning'). It is often associated with anxiety and depression, and bouts may be triggered by a period of stress.

An important point is that the symptoms of IBS, such as pain, can be quite severe, and may make sufferers think they have something more serious. As the BSG stated:[10]

> *People fear that they may have cancer or that the doctor is missing something more serious. Surely something as simple as IBS would not make me feel so dreadful.*

As we note later, this may affect referrals, if people seek reassurance by asking general practitioners (GPs) to refer them to specialist care.

Conversely, many people with IBD do not consult their doctors until they have had symptoms for some time. A study from Germany reported that patients with CD and UC waited for almost 8 months, on average, before consulting a physician.[11]

Coeliac disease

Coeliac disease is a disease of the small bowel, resulting from an immune reaction to the wheat gluten and similar proteins found in rye, barley and, to a lesser extent, oats. Coeliac disease can be ruled out by testing for autoantibodies at an early stage, so is not relevant to calprotectin testing. It could be classed as an inflammatory disease of the bowel but the inflammatory cells are mainly lymphocytes, so calprotectin is not high [but can be modestly raised in children (D Wilson, Royal Hospital for Sick Children, Edinburgh, 2013, personal communication)].

Ulcerative colitis

Ulcerative colitis is characterised by inflammation of the colon, sometimes intense, with bloody diarrhoea, but is often much milder. The cause is not known, but it appears that some people are more genetically susceptible than others.[12] Around 10% of people with UC have a first-degree relative with the condition. The concordance in monozygotic twins is also around 10%.

Curiously, cigarette smoking may confer some protection, or reduce severity.[12] The risk is also moderately reduced in people who have had appendicitis and appendix removal, under the age of 20 years.

Ulcerative colitis may involve an abnormal immune response to the microbacteria that normally live in the gut, known as commensals. UC is sometimes triggered by episodes of gastroenteritis caused by organisms such as *Salmonella*, *Shigella* and *Campylobacter*, but more by changes in the natural gut flora than direct effects of these organisms.

Ulcerative colitis typically starts in the rectum and spreads upwards through the colon. The natural history is of relapse and remission. At first presentation, most patients have mild disease, and only 10% have severe disease. About half will continue to have mild disease or remission, but in about one-fifth of patients, UC will be chronic and continuous, and more likely to become extensive, throughout the colon.

The aim of treatment in active disease is to secure a remission and then maintain that. Different drugs are used to induce, and then maintain, remission. There is an increased risk of colorectal cancer, so surveillance for that is part of care.

Crohn's disease

Crohn's disease can present in different ways, depending on which part of the intestinal tract is affected. Like UC, it is a relapsing and remitting inflammatory disease. However, it can affect any part of the GI tract – it is a much more extensive disease. Also like UC, there is a genetic susceptibility, with concordance in 35% of monozygotic twins.[13] The cause is unknown but it appears to be commoner in those with a 'westernised' lifestyle. Like UC, it may occur after infectious gastroenteritis and is associated with disturbances in the usual gut flora. The histological features include some similar to tuberculosis but no mycobacteria have been shown to be responsible. There are around 60,000 people with CD in the UK, of whom 20–30% are aged under 20 years.[14] The incidence is highest in the age range 15–30 years. About 25% of cases have onsets under age 17 years.

The pattern of symptoms in children is different. A prospective survey was carried out in the UK and Ireland by the British Paediatric Surveillance Unit, the BSG Research Unit and the Paediatric Register of Inflammatory Bowel Disease. A total of 739 cases were reported in children under the age of 16 years, making it the largest such study. The commonest presenting symptoms of CD were abdominal pain, weight loss and diarrhoea, but 44% did not report diarrhoea and only 25% reported the classical triad of abdominal pain, diarrhoea and weight loss. Other symptoms at presentation included lethargy and anorexia. Paediatric IBD (PIBD) is often more extensive at diagnosis than in adults.

The UK and Ireland survey found that delays in diagnosis of CD in children were common; 18% had a pre-diagnosis duration of symptoms of 1–3 years, and 9% of more than 3 years. Only 9% had isolated small bowel disease.

The delay in diagnosing PIBD has changed little over the last 20 years. What has changed is the incidence. Henderson et al.[15] reported a rise in Scotland of 76% from 1990–5 to 2003–8, and a fivefold increase over the last 40 years, especially in CD. This rise may not apply to the same extent in the rest of the UK, as there is a north–south gradient within Scotland,[16] but, internationally, rates have been increasing.[17]

Symptoms of CD include diarrhoea, pain, and blood or mucus in stools. Other presentations include anaemia due to disturbance of iron metabolism, and extraintestinal disease, such as arthritis, which may appear before any intestinal symptoms. Diagnosis is usually based on histology after biopsies taken during endoscopy. Differential diagnosis includes other causes of abdominal pain, such as IBS. Symptoms may be different in children, in whom growth retardation may be a feature that can precede bowel symptoms.[18]

The outlook for CD is worse than for UC. Only 10% have prolonged remission. Based on past experience, about 20% require hospital admission each year, and half will have required surgery within 10 years of diagnosis. This compares with the 10–30% of adults with UC who will require a colectomy in the first 10 years.[19,20]

The outlook in paediatric UC has been worse, with 20% of 113 children in one study[21] having to have their colon removed by 5 years of duration.

Newer drugs such as the 'biological' agents (infliximab and adalimumab) may reduce admission rates and the need for surgery.[22,23]

There are three main serious intestinal complications of CD. The first is stricture (narrowing) of the bowel. This can lead to intestinal obstruction, and CD can present as an 'acute abdomen' requiring surgery, sometimes mimicking appendicitis. The second is fistulas, which are abnormal connections between sections of bowel, or between bowel and bladder. The third is colorectal cancer, and surveillance is required.

In both UC and CD, some people have active disease but no symptoms. This has been noted following the introduction of colorectal cancer screening using faecal occult blood testing (FOBT). Positive screenees are referred for colonoscopy. Butcher *et al.*[24] reported that amongst 5350 such people who had colonoscopy, 66 were found to have unsuspected IBD (UC–CD 2 : 1), of whom about half had no symptoms. Some had quite extensive UC.[24]

Esch *et al.*[25] reported that some people with CD have no symptoms but are found by chance during investigations for other reasons. However, most developed symptoms over time (mean 3–4 years; range 2 months to 9 years) and one-quarter required surgery. They concluded that initially silent CD requires similar monitoring to initially symptomatic CD.

The treatments and the aims of treatments have changed in recent times. Schoepfer *et al.*[26] comment that the aims have evolved from relieving symptoms towards mucosal healing. They consider that this has been driven by the arrival of new medications, such as the anti-tumour necrosis factor (TNF) drugs that can induce and maintain mucosal healing. A New Zealand consensus conference concluded that early use of infliximab at induction led to higher rates of mucosal healing.[27] Economic modelling by Ananthakrisnan *et al.*[28] suggests that treatment aimed at mucosal healing is cost-effective compared with aiming only at relief of symptoms, because over a 2-year follow-up period the mucosal healing group would have fewer hospital admissions and less surgery than the symptom suppression group. This results in a cost per quality-adjusted life-year (QALY) of around £33,000, based on straight conversion of US dollars (US$) to British pound sterling (£).

The arrival of more effective new drugs increases the importance of prompt diagnosis of CD, and it could be argued that they should be used earlier in the treatment pathway. However, National Institute for Health and Care Excellence (NICE) technology appraisal (TA) 187[14] recommends use of the anti-TNF drugs, infliximab and adalimumab, only in people whose disease has not responded to conventional therapy with steroids or with immunosuppressive agents, such as azathioprine and 6-mercaptopurine (6-MP). Response means relief of symptoms.

The ratio of CD to UC varies between adults and children. In adults, the ratio of CD to UC is 2 : 3,[29] whereas in children the ratio is much higher, at 2.3 : 1.[30]

There are other forms of colitis, such as collagenous colitis and lymphocytic colitis, sometimes combined as 'microscopic colitis', which can cause persistent non-bloody diarrhoea, but these are usually seen in older people. The mean age at diagnosis in a Swedish study[31] was 64 years for collagenous and 59 years for lymphocytic.

Differential diagnosis
Some features of CD, UC, IBS and coeliac disease are compared in *Appendix 1*. The key point is that distinguishing among them by purely clinical means – signs and symptoms – can be difficult. Ford and colleague carried out a systematic review of the usefulness of symptoms and symptom scores for diagnosing IBS.[32] They concluded that individual symptoms (lower abdominal pain, passage of mucus per rectum, feeling of incomplete evacuation, passage of looser stools at onset of abdominal pain, abdominal pain relieved by defecation and patient report abdominal bloating) have limited usefulness for diagnosing IBS. They also concluded that composite scores such as the Manning and Kruis criteria had only modest accuracy, and noted that these scores were developed based on secondary care populations and might be less applicable to the patient mix seen in primary care. They also noted that around 40% of patients in the studies underlying the scores had some form of organic disease, suggesting an element of spectrum bias.

Jellema *et al.*[33] carried out a systematic review of the accuracy of symptom-based criteria for IBS (Manning, Kruis, Rome I and II and others). They included 25 studies, but only three were carried out only on primary

care patients. Jellema et al.[33] concluded that none of the criteria could reliably exclude organic disease. However, there is a school of thought that asserts that:[1]

> A positive diagnosis of IBS should be reached using symptom-based clinical criteria, not after excluding organic disease by exhaustive investigation.

This is echoed in the NICE scope:[34]

> In the majority of cases the diagnosis of IBS can be made on the basis of clinical history alone.

The systematic review done by the National Collaborating Centre for Nursing and Supportive Care for the NICE guideline group on IBS (p. 101) quotes Jeong et al.:[35]

> It is amazing to see the expensive, dangerous and extensive workups to which healthy patients are subjected by physicians searching for an organic cause in patients who obviously suffer from IBS.

The review lists many possibly investigations (pp. 100–1) but these did not include calprotectin.

General practitioners in the NTAC Durham Dales pilot study[36] were good at diagnosing IBS – if a GP thought a patient had IBS, the GP was right in 95%, using a negative FC as confirmation of diagnosis. Note that this does mean that 1 in 20 patients had a diagnosis other than IBS, with raised calprotectin suggesting IBD.

It may be useful to consider new presentations separately. Many GPs will feel confident about the diagnosis in recurrent IBS, when they know the patient well and they have presented with similar symptoms on previous occasions, perhaps after anxiety or stress. They may not feel a need to refer such patients. However, with new presentations there will be more diagnostic uncertainty, and the proportion referred to secondary care to exclude IBD may be higher. Calprotectin testing may be most useful in new presentations.

A survey of GPs from around Bristol found that most GPs were fairly confident ('8 out of 10', where '10' was most confident) that they could diagnose IBS at the first visit and most did not investigate the under 45 years age range further. As only a small proportion was referred for specialist investigation, there may have been some false negatives with IBD.[37]

However, many patients are referred to gastroenterology, for definite diagnosis, which is usually/often based on endoscopy and histology of biopsies. Some studies report that some patients with IBS are very anxious, and require the reassurance of a hospital 'check-up'. In one small study (54 patients) from Cardiff, the main reason for referral was diagnostic uncertainty (37/54) but the second reason was for 'confirmation of IBS' (17/54).[38]

In various studies, the proportion of patients referred for further investigation, in whom abnormal findings are reported on colonoscopy, is low. Kok et al.[39] noted reports that only 22–37% had organic bowel disease (OBD).

The ability of GPs to correctly identify IBS, in a considerable proportion of people with lower abdominal symptoms, has implications for the spectrum of patients in whom calprotectin testing might be used. IBS is very common, and one estimate is that 90% of patients seen, in general practice, with chronic lower abdominal symptoms, have IBS. This high prevalence of IBS in general practice groups has led to concern that results from studies carried out in secondary care may not be applicable to patients seen in primary care. A much higher proportion of patients in secondary care studies may have IBD. However, if GPs are referring only selected patients to specialist clinics, the prevalence of IBD among referrals will be higher, with the spectrum of referred patients more similar to that in the studies from secondary care.

Endoscopy may involve (1) colonoscopy, involving inspection of the whole colon; (2) sigmoidoscopy, inspection of only the distal part of the bowel (the sigmoid colon); or (3) gastroscopy, visualising the oesophagus, stomach and upper part of the small bowel. There are some sections of the small bowel that cannot currently be reached by widely available forms of endoscopy. In those situations, options include capsule camera endoscopy (the 'camera pill'), and imaging methods including ultrasound and magnetic resonance imaging (MRI).

Long delays in diagnosing IBD have been reported. Burgmann *et al.*[40] from Manitoba reported that 42% of a group of people with known IBD, had had GI symptoms for more than 3 years before the diagnosis of IBD, with some having symptoms for as long as 11 years before IBD diagnosis. Delays were much commoner in older age groups, with an incorrect diagnosis in around half of the over-64-year-olds compared with only around 10% in younger adults.

Diagnosis may be complicated by some patients having IBS before developing IBD. Because IBS is so common, this is not unexpected. It may also be that some such patients had IBD from the outset. However, it may be that the risk of IBD is raised in people that have had IBS. Porter *et al.* (2012)[41] reported that in patients who had had what they called 'well-defined IBS', as confirmed by negative colonoscopies, the relative risk (RR) of later IBD was 15. They suggest that some patients had microscopic colitis with normal appearance on colonoscopy, whereas others might have had CD restricted to small bowel.

National Institute for Health and Care Excellence clinical guideline 61 (irritable bowel syndrome in adults)

The NICE clinical guideline (CG) 61[35] makes recommendations for adults with IBS. The guideline recommends that patients with IBS should to be encouraged to manage their symptoms by themselves initially, and be given information on general lifestyle, physical activity, diet and symptom-targeted medication.

The advice on diet should be tailored according to the patient's symptom (diarrhoea, constipation). If diarrhoea is the predominant symptom, then patients should be advised to limit intake of high-fibre food, limit the consumption of fresh fruit and avoid eating insoluble fibre. Patients should also avoid consumption of sorbitol (an artificial sweetener) found in sugar-free sweets and drinks.

If the predominant symptom is constipation, then patients should be advised not to consume starch that resists digestion in the small intestine and reaches the colon intact. If patients need high dietary fibre then they should take soluble fibre such as ispaghula powder or foods, such as oats, which are rich in soluble fibre.

Some of the advice that relates to all types of IBS includes having frequent meals and eating slowly; not skipping meals or having long gaps between meals; drinking at least eight cups of fluid per day, especially water; restricting tea and coffee to three cups per day; avoiding insoluble fibre.

If patients continue to have symptoms and severity increases, then pharmacological intervention is recommended, but no length of time before this is specified in the NICE guideline.[35]

Irritable bowel syndrome-diarrhoea

Pharmacological intervention

First-line treatment
Antispasmodic agents should be taken as and when required, alongside dietary advice.

Second-line treatment

Tricyclic antidepressant drugs (TCAs) started at a low dose taken at night. If TCAs are ineffective then selective serotonin reuptake inhibitors (SSRIs) can be tried. After prescribing TCAs or SSRIs, patients should be followed up after 4 weeks and then at 6- to 12-monthly intervals thereafter.

Psychological interventions

If patients do not respond after 12 months of pharmacological therapy, they may be referred for psychological interventions, such as cognitive behavioural therapy (CBT).

Irritable bowel syndrome-constipation

Pharmacological intervention

First-line treatment

Ispaghula powder Further management in patients with IBS-constipation is similar to those with IBS-diarrhoea.

The reason for including the above summary is because it shows that IBS may be treated in a stepwise way. Each step may take time to be tried, and many patients will not respond to the first or later therapies. The importance of this is because a patient with IBD, which is misdiagnosed as IBS, may go through a time-consuming series of treatments for IBS, before clinical suspicion leads to referral to gastroenterology or paediatrics. IBS can cause considerable pain and discomfort, sometimes more than IBD.

Calprotectin

Calprotectin is a protein found in some cells, most notably the group of white blood cells called neutrophils. It binds to calcium, and is then a stable compound not broken down in the intestines.

In people with bowel conditions that cause inflammation, the increased number of neutrophils in the bowel leads to an increase in FC. It can therefore be used as an indication of inflammation. There are now tests to detect or measure the level of calprotectin in faeces. It appears stable in faeces for at least 7 days (though not all agree). It is also reproducible from day to day in individuals. Naismith et al. (2013)[42] obtained stool samples on three consecutive days from 143 patients with CD, and found low day-to-day variation. They concluded that clinical decisions could be made on a single calprotectin result.

Moum et al.[43] reported considerable variability in FC levels in patients with CD, in samples taken on two consecutive days. However, the variability was seen mainly at high levels, with little in the borderline region of 50–200 mg/l (normal is < 50 mg/l).

There can be false positives from the taking of non-steroidal anti-inflammatory drugs (NSAIDs) but these can be avoided by asking patients to stop taking the drugs before calprotectin testing.

In a Finnish study[44] that compared medication use amongst people with IBD, and the general population, people with IBD had almost a fourfold increase in use of proton pump inhibitors (PPIs) (OR 3.9) and a slight increase in the use of NSAIDs (OR 1.17). However, not all studies have reported increases with NSAIDs. In a study amongst those with borderline calprotectin levels (> 50 but < 150 g/g), Demir et al. (2012)[45] found no significant difference with NSAID use. Conversely, Turvill (2012)[46] reported that 14% of people referred from primary care with intestinal symptoms, and who had raised calprotectin, had a final diagnosis of NSAID enteropathy.

There can also be false positives after chest infections (because of the white blood cells in swallowed sputum) and after bleeding into the bowel.

The proposed role of FC testing in this appraisal is to aid differential diagnosis in people with lower GI symptoms (pain, bloating, diarrhoea, change in bowel habit). The aim is to distinguish between those with inflammatory conditions and those with no inflammation. Many of those with inflammation will have IBD but others may have cancer or other conditions. Most of those with no inflammation will have IBS.

Knowledge of the presence or absence of inflammation will affect the decision on referral for further investigation. The absence of inflammation may lead to a presumption of IBS, to be managed in primary care. The presence of inflammation would be likely to trigger referral to gastroenterology for further investigation, likely to include endoscopy.

Hence there could be two benefits. Those with IBS would not be referred and could therefore escape further investigations especially colonoscopy. Those with inflammation might be referred more promptly and receive appropriate treatment earlier.

Faecal calprotectin could be part of a pre-referral work-up in general practice, such as outlined in *Figure 1*. In the second box, 'TTG' (tissue transglutaminase) refers to testing for coeliac disease. The term 'red flag' is used to refer to symptoms or signs that might be due to cancer, including anaemia, rectal bleeding, unexplained weight loss, abdominal masses, and change in bowel habit in patients of over 60 years of age. A family history of bowel cancer might also be a red flag item.

The age cut-off of 45 years is somewhat arbitrary but was used in the BSG guidelines for diarrhoea in 2002.

The first stage involves excluding patients with 'red flag' signs or symptoms (*Table 1*). These could be indicative of cancer and are indications for rapid referral. However, many people with proven IBS also have red flags. Whitehead *et al.*[47] report data from the Puget Sound Health Cooperative.

Rectal bleeding may be due to haemorrhoids ('piles'), which are common (around 20%) in patients with IBS, especially those with IBS-C.

The next stage involves blood tests, one of which is TTG, a test for coeliac disease. This means that coeliac disease can be confirmed or ruled out at this stage. At present this stage also involves measurement of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), which are markers for inflammation. However, these do not localise the inflammation to the bowel, whereas FC does.

One issue to be considered is whether ESR and CRP should be done at the same time as TTG and full blood count (FBC), on the grounds that they are cheaper, and can done at the first visit. If negative ESR and CRP could rule out inflammatory conditions of the bowel, a presumptive diagnosis of IBS could be given, and referral for further investigations would not be made at this stage. However, a number of studies have reported that CRP and ESR have poor sensitivity and/or specificity[48,49] (R Arasaradnam, University Hospital Coventry and Warwickshire, 2012, unpublished data), meaning that they are negative in many people with active CD. The report by the York Health Economics Consortium (YHEC)[50] for the Centre for Evidence-based Purchasing (CEP) concluded that FC testing dominated (i.e. was both more effective and less costly than) ESR and CRP. More recently, Mascialino *et al.*,[51] from one of the manufacturers of calprotectin tests, Thermo Fisher Scientific (Phadia AB, Uppsala, Sweden) also concluded that FC testing dominated ESR and CRP, after taking into account all costs, in primary and secondary care, including reductions in endoscopies. Their estimate for the UK was that FC saved at least £100 per patient investigated compared with ESR and CRP.
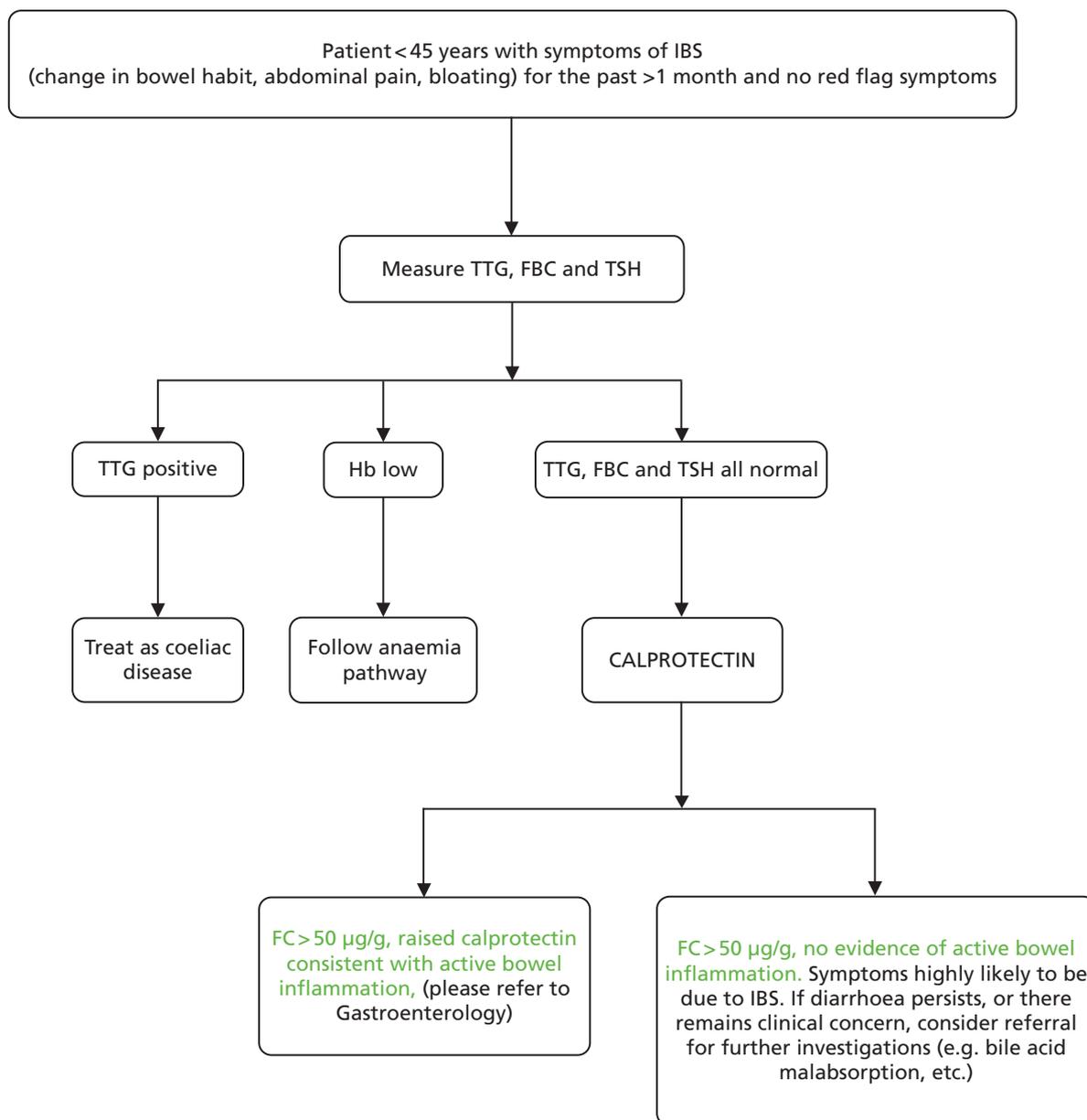
FIGURE 1 Possible pathways in patients with symptoms of IBS.
FBC, full blood count; Hb, haemoglobin; TSH, thyroid-stimulating hormone.

TABLE 1 Red flag indicators in IBS and cancer

Indicators	IBS (%)	GI cancer (%)
Blood in stools	15	14
Unintended weight loss	21	56
Onset of symptoms at > 50 years	32	67
Family history of cancer	20 (unclear but presumed colon cancer)	39 (colon cancer)

Decision problem

The aim of this review is to examine the clinical effectiveness and cost-effectiveness of FC testing in distinguishing between 'functional' disorders, such as IBS, with which sufferers will not come to serious harm, and 'organic' disorders, such as IBD, which require referral to specialist care. In adults, the differentiation is most often between IBS and IBD. In children, there is a different range of conditions.

If calprotectin is a reliable way of detecting inflammation of the bowel, or its absence, then those patients in whom the test shows normal levels could be spared referral to specialist care and the often invasive and unpleasant investigations, such as colonoscopy, which may follow.

Population

The population is patients with lower GI symptoms that are chronic, defined as persisting for at least 6–8 weeks. The upper age limit is 60 years, as per the NICE scope.[34] Symptoms in adults include abdominal pain or discomfort, bloating or change in bowel habit. Some will be newly presenting in primary care; others may already have been referred to specialist care.

Children (under 17 years) are a separate group with a different mix of conditions.

The main focus would ideally be in primary care, because that is where people with lower bowel symptoms first present. FC testing has not been widely available in, or to, primary care, and hence much of the differential diagnosis has been done in hospital clinics.

This could potentially give rise to problems reflecting selection for referral. For example, there may be three groups of people with IBS:

1. Those who do not seek help or advice from GPs but self-treat, as required, with over-the-counter medications, such as laxatives and analgesics: 'self-managed'.
2. Those who do present to their GPs but whose symptoms are not considered such as to require referral: 'GP-managed'.
3. Those referred by GPs to specialist clinics – the 'referred' group. At present it is estimated that only about 25% of patients are referred to secondary care. However, as in the past referral was often followed by colonoscopy, the threshold for testing with calprotectin may be rather lower than the threshold for referral, and many more than 25% may be tested with calprotectin (later, we assume 50%).

Evaluation of tests on only the referred group could, at least in theory, cause spectrum bias problems if the prevalence of IBS was less, and that of IBD higher, as parameters influenced by prevalence might differ from the GP-managed group. This could be important if FC testing was recommended and made more widely available. However, as noted above, GPs are highly selective in whom they refer.

Testing will be used mainly for IBS-D and not IBS-C.

Intervention

Faecal calprotectin tests (*Table 2*). These are of two types:

1. laboratory testing, using mainly enzyme-linked immunosorbent assay (ELISA) methods
2. point-of-care testing (POCT), which can be used in primary care or secondary care.

Laboratory methods are quantitative. Point-of-care tests may be quantitative or semi-quantitative.

The point-of-care tests can give faster results, within about 30 minutes. Extraction of the faecal sample is always manual, so some time costs are irreducible.

TABLE 2 List of FC tests

Manufacturer	Test	Platform
Bühlmann, Laboratories, Schönenbuch, Switzerland	EK-CAL calprotectin ELISA test	ELISA – quantitative Monoclonal antibody. Two versions with low range used for FC levels up to 600 µg/g (range 10–600 µg/g). The manufacturer's recommended cut-off level is 50 µg/g for adults and children aged between 4 and 17 years
Bühlmann, Laboratories, Schönenbuch, Switzerland	Quantum Blue calprotectin test	Rapid test – Immunoassay designed for the quantitative determination of FC in combination with the BÜHLMANN Quantum Blue® Reader. There are two versions, LF-CAL25 with range: 30–300 µg/g, and LF-CHR 25 with range: 100–1800 µg/g The manufacturer's cut-off value of the LF-CAL 25 is 50 µg/g. The manufacturer recommends re-testing samples if results are between 30 and 70 µg/g. This zone is regarded as 'grey zone' and the values corresponds to the 2.5th–97.5th percentile of imprecision around the cut-off of 50 µg/g
Calpro, Lysaker, Norway	Calpro calprotectin ELISA test (ALP) – formerly known as the PhiCal test	ELISA – quantitative The two versions have different ranges: CALP 0100 up to 1250 mg/kg, CALP 0170 up to 2500 mg/kg
	CALP 0100 and CALP 0170	Quantitative ELISA using polyclonal rabbit antibody. Recommended cut-off of 50 µg/g
Eurospital, Trieste, Italy	Calprest	ELISA – quantitative, polyclonal The cut-off level is 50 mg/kg. The manufacturer suggests retesting after a short period of time in patients with FC levels of between 50 and 100 mg/kg
Eurospital, Trieste, Italy	CalFast	Rapid test – Quantitative determination of FC in combination with a dedicated reader
Immundiagnostik AG, Bensheim, Germany	ELISA (K6927)	ELISA – quantitative Quantitative ELISA, using two monoclonal antibodies. Recommended cut-off of 50 mg/kg, and can also be used in children aged 4–17 years. The manufacturer recommends that laboratories establish their own reference range
Thermo Fisher Scientific, Uppsala, Sweden	EliA Calprotectin	EliA – quantitative In contrast with ELISA, EliA measures the presence of target antibodies by fluorescence signal detection EliA platform is a fully automated test, said by the manufacturer to reduce technician workload, time and cost
Preventis (sister company to Immundiagnostik), Bensheim, Germany	KST11005 CalDetect Calprotectin Rapid test (version 1 – CalDetect) *Also referred to as the 'PreventID CalDetect' test*	POCT – immunochromatographic rapid test A semiquantitative test with three lines corresponding to calprotectin 'negative', calprotectin ≤ 15 µg/g, calprotectin 15–60 µg/g and calprotectin > 60 µg/g stool
Preventis (sister company to Immundiagnostik), Bensheim, Germany	CalDetect Calprotectin Rapid test (version 3 – CalScreen)	POCT – immunochromatographic rapid test A yes/no test with only one test line corresponding to the cut-off value of 50 µg/g stool (no inflammation = < 50 µg/g and inflammation present = ≥ 50 µg/g)

Comparators

The main comparator is clinical assessment, which can be supplemented by ESR and CRP, which can indicate inflammation, but not localise it. There are two options for ESR and CRP testing:

1. If GPs have access to FC testing, they could use that in people with suspected IBS. So FC would replace ESR and CRP testing.
2. If normal ESR and CRP could exclude inflammation of the bowel, they might be used as part of the initial work-up. However, the evidence suggests that normal CRP results can occur in the presence of active inflammation.

The limitations of ESR and CRP are:

- Negative tests do not exclude IBD, so if symptoms persist, patients would still require further investigation.
- Positive tests might be due to other, non-GI inflammations, so further investigations would be needed to localise the inflammation.

In one survey carried out in 2010,[26] 89% of gastroenterologists considered calprotectin to be more accurate than CRP and ESR for distinguishing between IBS and IBD. A review by Burri and Beglinger (2012)[52] noted that ESR and CRP had low sensitivity.

As noted earlier, CRP and ESR have poor sensitivity for IBD.

Therefore, there seems little point in doing these tests even if calprotectin was not available. As noted previously, the YHEC report[50] noted that CRP and ESR were economically dominated by calprotectin. These tests are therefore not examined further.

Outcomes

Depending on data availability, these may include:

- referral rates
- numbers of colonoscopies with/without FC testing
- proportion of colonoscopies with no abnormal findings
- duration from onset of symptoms to definite diagnosis of IBD – late diagnosis of CD
- cost
- adverse events such as complications of colonoscopy
- QoL and hence QALYs.

Modelling approach

A set of possible pathways is shown in *Figure 2*:

- No FC testing available. Clinical assessment and simple tests in primary care followed by decision on referral or symptomatic treatment/therapeutic trial in those thought to have IBS.
- Laboratory testing available to GPs. The laboratory just reports the results.
- 'Laboratory plus' where the GP provides clinical details along with the test request and gastroenterologist or clinical biochemist provides commentary and advice.
- POCT available in primary care. If it is negative, the GP manages the patient on a presumptive diagnosis of IBS. If the result is positive, the GP can refer to gastroenterology for further investigation. If indeterminate, the GP can either repeat test or refer.
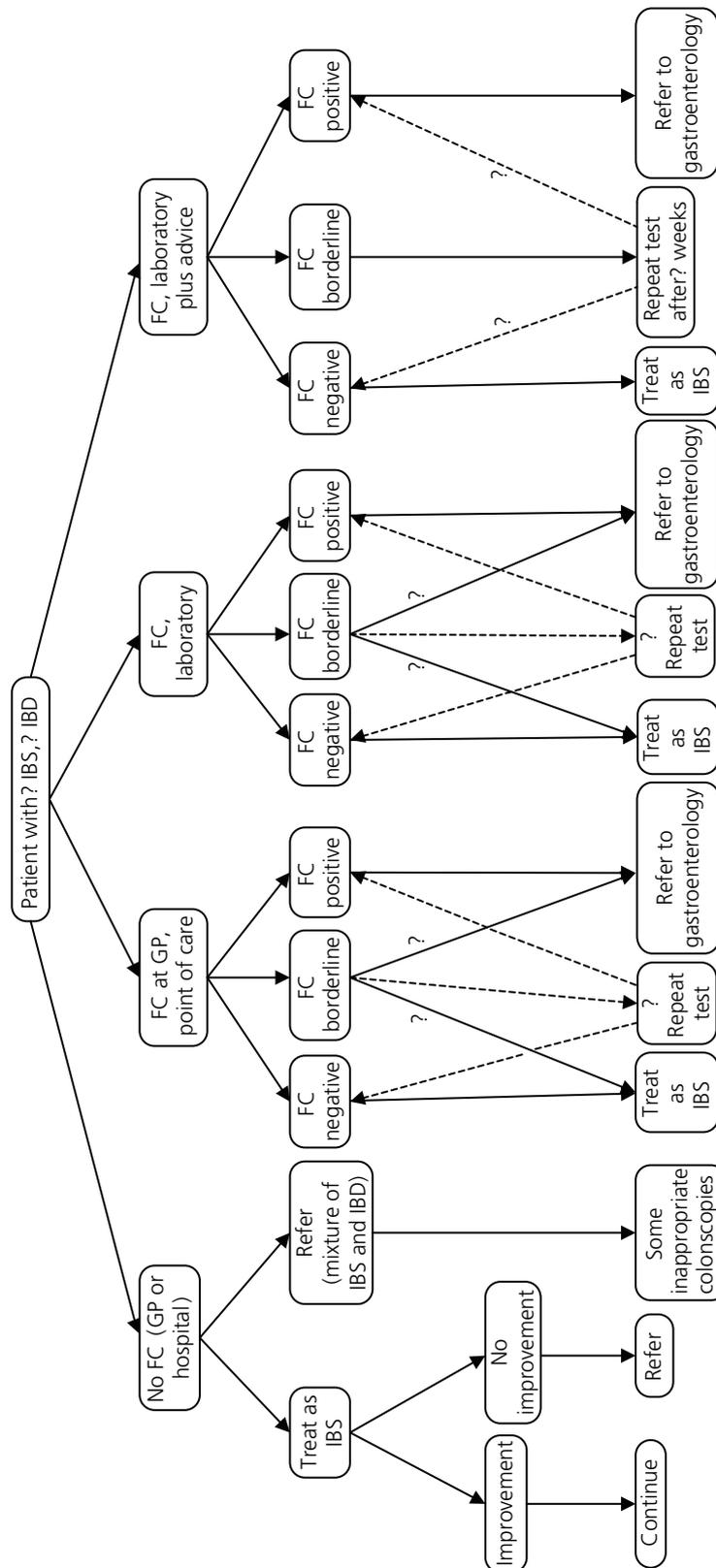
FIGURE 2 Service options.

Methods

The inclusion criteria were studies comparing FC as a guide to inflammation of the lower intestine, ideally with histology as the reference test, in *newly presenting* patients. Exclusion criteria included studies of FC for monitoring activity of IBD or response to treatment in people with known IBD.

We also identified, appraised and summarised recent systematic reviews.

The databases searched for diagnostic studies included the databases MEDLINE, EMBASE, The Cochrane Library and Web of Science from their inception up to March 2013. Also, additional sources of grey literature were searched, the reference lists of relevant articles checked, and experts were contacted for unpublished data. Full details of the search strategy are shown in *Appendix 2*.

The selection was done in three stages, based on fulfilling each of following criteria:

1. Were the patients newly diagnosed?
2. Was an acceptable reference standard used?
3. Were the appropriate outcomes reported, i.e. were sensitivity and specificity data reported or was it possible to derive a 2 × 2 table to determine them?

The hierarchy of evidence based on reference tests was:

- gold standard – endoscopy (usually colonoscopy) and histology
- endoscopy and results by disease but no mention of histology – biopsies presumed to have been done
- endoscopy with report that no biopsies done. Camera endoscopy included here
- no endoscopy but diagnosis by imaging methods, for example thickened gut wall on computed tomography (CT)
- clinical follow-up for 6 months.

Studies were grouped according to the conditions being compared, with most weight being given to:

- studies comparing IBS with IBD
- studies comparing IBD with all non-IBD conditions.

Data were extracted from the included studies for 2 × 2 tables, with FC as the screening test and bowel histology as the reference test. If studies fulfilled the other inclusion criteria but data for 2 × 2 tables were not available, we reported what data were available, such as calprotectin ranges, medians and interquartile ranges (IQRs) to compare groups with different conditions.

In papers where the numbers of true and false positives and negatives were not reported, but data on sensitivity and specificity and the total numbers of people with and without disease were reported, the data for the 2 × 2 table were calculated using the Calculator function in Review Manager (RevMan) version 5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark).

Data on five covariates, FC cut-off level, make of test, age (adult or paediatric), setting (primary or secondary care) and type of test (ELISA or POCT), were extracted for each study and entered into RevMan.

All calprotectin levels were reported in micrograms per millilitre (or equivalent), except by Tibble *et al.* (2002),[49] who used a non-commercial in-house ELISA, with levels reported in milligrams per litre. On the basis of data in previous systematic reviews, results were converted to micrograms per millilitre by multiplying by a factor of five.[53,54]

Statistical methods

RevMan was used for data entry and analysis to generate forest plots, and MedCalc version 12.3.0 (MedCalc Software, Mariakerke, Belgium) for producing statistical data based on the 2 × 2 tables, including positive likelihood ratio (PLR), negative likelihood ratio (NLR), positive predictive value (PPV), negative predictive value (NPV) and disease prevalence.

Studies that provided sufficient data for calculation of sensitivity, specificity and other diagnostic outcomes were identified, and data were entered into RevMan for the generation of paired forest plots and receiver operating characteristic (ROC) curves. Further statistical analysis was performed in Stata 12 (StataCorp, College Station, Texas, USA) to produce likelihood ratios (LRs), area under the curve (AUC) and nomograms. Our intention was to examine the performance of calprotectin testing over a range of values, starting with the level recommended by the manufacturers, which is most often 50 µg/g. Where sufficient studies reported results at the same values, we aimed to pool data for each value.

Meta-analysis was performed in accordance with previously reported guidelines for meta-analyses of diagnostic tests using the Stata command 'metandi'.[55,56] Pooled estimates for values among different diagnoses were obtained with 95% confidence intervals (CIs), assuming a bivariate model.

If there were sufficient studies, we planned to pool data at the same cut-off levels from ELISA and point-of-care tests separately, and compare them. However, only ELISA tests were pooled.

Quality assessment

Quality assessment of studies was done using items adapted from the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) I tool.[57]

Quality assessment items used

1. Was the spectrum of patients representative of the patients who will receive FC testing in practice?
2. Is the reference standard likely to classify the target condition correctly? The reference standard for confirmation of bowel inflammation was histology of biopsies obtained at endoscopy.
3. Is the time period between FC measurement and obtaining tissue for histology short enough to be reasonably sure that the target condition did not change between the two tests? We regarded an acceptable delay between tests as being ≤ 3 months or less.
4. Did the whole group receive verification by histology? If not, were results for those who did receive histology reported separately?
5. Did patients receive endoscopy and histology irrespective of the FC result? (Differential verification avoided.)
6. Disease stage. Were patients newly presenting with symptoms? Some studies had mixed groups of newly presenting and patients already known to have IBD, and we allowed up to 20% of non-new patients. Studies in patients with > 20% confirmed IBD, whether active or in remission, were excluded. Some studies clearly stated that patients were newly presenting. In others (Damms and Bischoff;[58] Garcia et al.;[59] Li et al.;[60] Licata et al.;[61] Shitrit et al.[62]), less detail was given, and we inferred that they were newly presenting from terms such as 'referred for investigation of chronic diarrhoea'. So possible answers were yes, or probably. Ideally, we would have contacted authors or excluded studies in which new presentation was not clear.
7. Were histology or endoscopy results interpreted without knowledge of the FC results? (Index test results blinded.)
8. Were the FC results interpreted without knowledge of the results of histology? (Reference standard results blinded.)
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (Relevant clinical information.)
10. Were intermediate FC results reported? (Uninterpretable results reported.)
11. Were withdrawals from the study explained? (Withdrawals explained.)

Question 6 replaced the usual QUADAS question on whether the reference test was independent of the index test, as histology, our preferred reference test, is clearly independent of calprotectin, so the usual question 6 would not help discriminate.

The term 'quality assessment' is preferred to the more traditional 'risk of bias' term because the latter, as used in systematic reviews, such as Cochrane reviews, is more associated with assessing internal validity of randomised controlled trials (RCTs). We need to assess external validity through items such as spectrum bias.

All data extractions and quality assessments were done by one author and checked by another.

The NHS Technology Adoption Centre pilot studies

The NHS Technology Adoption Centre (NTAC) sponsored two pilot studies of the implementation of FC testing:[36]

> West Northumberland Clinical Commissioning Group (CCG) used a fully quantitative test (Quantum Blue) with samples being analysed in the laboratory. It is technically possible to use this equipment as a point-of-care test in primary care, although it is thought unlikely that this would ever be economical in practice.
>
> Durham Dales CCG used a semi-quantitative point-of-care test (CalDetect, version 1, Preventis, Bensheim, Germany), with the analysis being carried out in the GP practice.
>
> In both cases there is a high cut-off value above which the patient should be referred to secondary care, and a low cut-off value below which there is a low probability of organic disease. Between the high and low cut-off values there is an intermediate range, in which case the patient should be retested. Owing to differences in the assays used there is a difference in the cut-off values used in the project sites.

Chapter 2 Results of clinical effectiveness review

The database searches retrieved 1273 references and 35 came from additional searches; there were 725 references remaining after deduplication. The flow chart is provided in *Appendix 2*. All of the 83 full-text articles assessed for eligibility were assessed independently by three authors and any differences were resolved by discussion. In the interests of brevity, studies in this section will usually be referred by name of first author and year. Full details of the baseline characteristics of all the included studies are given in *Appendix 6*.

Some issues

A number of issues, listed below, arose in this review.

Reference standard

We took histology after endoscopy to be the definitive reference standard. Some studies used other reference tests. For example, the authors of one study in a paediatric group, quite reasonably, did not consider it justifiable to use endoscopy children with normal calprotectin levels. Instead, they used a 6-month period of observation. In one study in adults, only those with high calprotectin levels had endoscopy. Those with normal levels were managed in primary care.

Note that not all CD can be reached by endoscopy. About 30% of CD in adults is ileal alone, and 50–60% is ileocolonic. But about 20% is in proximal or mid-ileum, and so not accessible by standard colonoscopy of gastroscopy. Small bowel enteroscopy is complex, expensive and available in only a few centres in the UK. So, options include video capsule camera or MRI.

Magnetic resonance imaging of the small bowel (especially after preparation with enteroclysis) is sensitive and has been shown to correlate with FC levels. Zippi *et al.*[63] reported a good correlation between MRI changes (such as wall inflammation and thickening of bowel) and FC levels.

Ultrasound has been used in several studies. Aomatsu *et al.*[64] used ultrasound to detect CD in the small bowel in children, using > 3 mm thickening of the small bowel as indicating active CD. Calprotectin levels were much higher (mean 738 μg/g) in children in clinical remission but with active lesions on ultrasound than in children in clinical remission but without activity on ultrasound (mean 18 μg/g). In this study, ultrasound was used as a reference standard for calprotectin testing, but the reverse can apply.

Canani *et al.*[65] also used ultrasound in children but found some overlap between CD and non-IBD cases. However, they described transabdominal ultrasound of bowel wall thickness to be a useful and non-invasive method in confirming IBD, especially if inflammation was localised to the ileum.

Ultrasound is useful more as screening tool, as it is not sensitive enough to assess location of IBD. Small bowel MRI with contrast follow-through is standard in children. Wireless capsule endoscopy is also used.

Tomas *et al.*[66] considered that calprotectin showed good correlation with scintigraphy with radiolabelled leucocytes, which they considered was the gold standard for measuring inflammation in the bowel, although undesirable in children because of the radiation and the need for anaesthesia, and thus not used.

Patient groups in studies

The proposed value of calprotectin testing in primary care is to help GPs make decisions on likely diagnosis, in order to decide whom to refer to specialist care for further investigation. Patients with 'red flag' signs or symptoms are referred, and so are excluded from the calprotectin pathway. So the value of

calprotectin is to guide decisions on whether to refer or not. A low calprotectin level indicates absence of inflammation, suggesting that IBS is the likeliest cause of the symptoms. A high level in someone with chronic symptoms suggests IBD, CD or UC. (FC can be raised in acute bacterial gastroenteritis but that usually resolves rapidly.)

Many studies compared 'non-organic' conditions (principally IBS in adults) with any organic condition. However, some organic conditions are not obviously inflammatory, so studies where the organic group included a mixture of conditions could make calprotectin testing look less useful. Calprotectin will therefore appear most impressive in studies that include only IBD and IBS.

Table 3 shows two things. First, the overlap in calprotectin levels between some organic conditions and IBS. Hence comparing only 'all organic' and non-organic will make calprotectin testing seem less valuable. Second, that calprotectin levels are raised in colorectal cancer, and to a lesser extent in people with larger adenomas. Adenomas are not usually regarded as being inflamed, in the sense of being infiltrated with white blood cells.

Cut-offs for calprotectin

One problem with the evidence is that many studies used only the manufacturer's recommended cut-offs. This presents a problem when it comes to assessing optimum thresholds – there is little evidence for levels other than 50 µg/g. We are grateful to Professor KD Bardhan, Dr P Basumani and Dr A Banerjee for providing unpublished data (Rotherham Hospital, 2013, personal communication) from Rotherham on different cut-offs.

There is debate about the minimum number of studies that should be used for pooling data on different cut-offs, with four being regarded as the minimum.[67] We have therefore not pooled studies if there were fewer than four at the cut-off in question but have relied on diagnostic odds ratios (DORs) as the summary statistic when there were fewer than four studies.

Spectrum

Nearly all studies come from secondary care. The secondary care studies will have a different mix of patients from those seen in primary care. (See prevalence data in later tables.) The sensitivity and specificity of testing will be the same, but the different prevalence will give different predictive values.

This may be a particular problem in comparing different tests, such as point-of-care and laboratory tests. These may appear comparable in secondary care populations, but if the calprotectin levels are much higher in those selected populations then the comparability results may not be generalisable to populations with lower calprotectin levels.

Another issue about spectrum of patients arises from another selection effect. The pilot studies of calprotectin use in primary care from the northeast of England have shown that GPs are good at diagnosing IBS. In the Durham Dales pilot, 95% of those predicted by GPs to have IBS, had it. The GPs were also good at predicting IBD – 88% (28 of 32) of patients who had high calprotectin levels, had been

TABLE 3 Data from Kok et al.[39] on FC levels (µg/g) and adenomas

Calprotectin test	Adenocarcinoma	Adenomas ≤ 1 cm	Adenomas > 1 cm	IBS
Quantum Blue	Median 215	Median 42	Median 111	Median 40
	IQR 105–300	IQR 30–105	IQR 30–264	IQR 30–69
EK-CAL	Median 274	Median 60	Median 89	Median 49
	IQR 94–442	IQR 24–108	IQR 34–217	IQR 21–99

predicted by their GPs to have IBD. So GPs may confidently diagnose IBS on clinical grounds in many patients, which implies that those who will have calprotectin testing may be a selected group. They may be more akin to those seen in secondary care – making the results from the secondary care studies more generalisable. A review by Van Roon et al.,[54] described in detail below, concluded that individual GI symptoms could not reliably distinguish between IBS and IBD, but GPs may use non-GI symptoms or signs, and 'clinical nous', to diagnose people with IBS. A recent review of IBS concluded that it could be diagnosed on clinical grounds:[3]

> The diagnosis should be reached using symptom-based clinical criteria, rather than excluding underlying organic disease by exhaustive investigation.

Choice of measure

As noted by Harbord and Whiting,[55] there is no single measure of diagnostic accuracy. They recommend that the measures most often used are sensitivity and specificity, with the trade-off between these being illustrated graphically.

In the sections that follow, we report:

- brief details of the included studies
- QUADAS quality assessment
- results
- sensitivity and specificity in paired forest plots, for all included studies
- for one study with a range of cut-off points, its own forest plot and ROC curve
- ROC curves with pooled sensitivity and specificity, and AUC
- Forest plots for the studies included in the ROC curve
- Fagan's nomograms with likelihood ratios
- tables of DORs for different cut-offs, pooled where appropriate.

Previous reviews

Five recent systematic reviews were quality assessed and summarised (*Tables 4–6*).

TABLE 4 Inflammatory bowel disease vs. non-IBD from Van Roon *et al.*[54]

IBD vs. non-IBD						
Group	Cut-off (µg/g)	Patients	Studies	Sensitivity (%)	Specificity (%)	AUC
Adults and children	50	1267	9	89	81	0.95
Adults and children	100	328	4	98	91	0.98
Adults	50	1030	6	71	80	0.94
Children	50	201	3	83	85	0.96
Children	100	231	3	98	97	0.99
CD vs. normal controls and IBS						
Adults	50	614	4	95	84	0.97
Children	50	119	2	97	79	–
Children	100	155	2	100	98	–
UC vs. normal controls and IBS						
Adults and children	50	235	2	78	78	–

TABLE 5 Characteristics and conclusions of previous reviews

Study	Inclusion criteria and methodology	Included studies	Conclusions
CEP 2010[50] *Focus*: FC for distinguishing between IBD and IBS Overall quality: medium	*Inclusion criteria* Study design: any Participants: not explicitly defined, patients with possible IBD or IBS; diagnostic procedure: laboratory and point-of-care tests for FC and other inflammatory markers Outcomes: sensitivity, specificity, PPVs, NPVs *Methodology* Search strategy: search of 11 databases (some not very relevant, such as CINAHL and BNI), studies published in the past 10 years; search terms indicated; English language only Study selection: not reported Quality assessment: not reported Data extraction: not reported Data analysis: text and tables	No. of included studies: 43 (?) – search results not described No. of participants: about 5050 Study quality: not reported Participants: not described in summary Diagnostic procedure: cut-off values for FC ranged between 18.6 µg/g and 250 µg/g	Conclusions: FC performs well in distinguishing OBD from functional bowel disease; sensitivity and specificity are > 80% in most studies (at cut-off 50 µg/g); where calculated, most PPVs and NPVs were 70–90% Recommendations for practice: none Recommendations for research: none
Jellema 2011[68] *Focus:* summary of diagnostic tests in patients with abdominal symptoms Overall quality: high	*Inclusion criteria* Study design: cohort studies, case–control studies where controls were diagnosed with IBS or in whom organic GI disease was excluded Participants: adult population consulting a physician because of non-acute GI symptoms (primary care, open-access clinics, outpatient population with prevalence of IBD of ≤ 25%); target condition was IBD but the perspective was from primary care; 'non-acute' was defined as symptoms for more than 2 weeks Diagnostic procedure: primary diagnostic studies; studies using colonoscopy, histology, barium enema and/or clinical follow-up to diagnose IBD (reference tests); index tests included: signs and symptoms, blood and faecal tests, abdominal ultrasonography (only FC considered here) Outcomes: sensitivity, specificity, data for construction of 2 × 2 table. Studies were excluded in 2 × 2 table could not be constructed	No. of included studies: Nine on FC No. of participants: 863 Study quality: Five with positive assessment on ≥ 8 of 11 quality items; range 4–10; only a minority of studies used a design relevant to primary care (and none of these was a study of calprotectin) Participants: all primary diagnosis – appeared to be newly presenting patients Diagnostic procedure: diagnostic cut-off points 15 µg/g and 170 µg/g, 10–30 mg/l	Conclusions: calprotectin showed consistent and promising findings but none of the studies was performed in primary care Authors conclusions: 'FC has excellent NPV in patients with abdominal symptoms' Recommendations for practice: none Recommendations for research: authors' conclusions' 'Before calprotectin can be used to guide clinical decisions in primary care, these markers need to be investigated by high-quality prospective studies in that specific setting'

TABLE 5 Characteristics and conclusions of previous reviews (*continued*)

Study	Inclusion criteria and methodology	Included studies	Conclusions
	Methodology		
	Search strategy: MEDLINE, EMBASE for studies published up to February 2009; search terms indicated; search of reference lists of relevant articles and reviews, etc.; languages restricted to English, Dutch, German and French		
	Study selection: selection by two independent authors; disagreements resolved by discussion; third author consulted in case of persisting disagreement		
	Quality assessment: yes, modified QUADAS tool		
	Data extraction: pre-tested forms; data extraction by two independent authors		
	Data analysis: diagnostic 2 × 2 tables, diagnostic performance measures; text and tables; distinguish between CD and UC		
Kostakis 2012[69] *Focus*: FC for diagnosis and confirming relapse in PIBD Overall quality: low	*Inclusion criteria* Study design: primary studies; case reports excluded Participants: patients aged ≤ 18 years with IBD, both newly diagnosed and previously confirmed Diagnostic procedure: measurement of FC Outcomes: sensitivity, specificity, PLR, NLR *Methodology* Search strategy: MEDLINE, EMBASE for studies published up to October 2011; search terms indicated; English language only Study selection: no details on study selection given Quality assessment: no quality assessment Data extraction: no details on data extraction given	No. of included studies: 34 No. of participants: 1345 with IBD (range 8 to 128), 1225 controls (range 0 to 509) Study quality: not reported Participants: Thirteen studies in newly diagnosed patients; nine studies in patients under treatment; 10 studies including both. No data provided on type of controls, who could be healthy controls or 'other GI disease', or have 'functional disease' not specified. Two studies in 'newly diagnosed' had no controls Diagnostic procedure: cut-off values for FC ranged between 50 and 275 µg/g	Conclusions: the FC test could be used for supporting diagnosis or confirming relapse of IBD in paediatric patients before they undergo GI endoscopy; a positive result could confirm the suspicion of either IBD diagnosis or IBD relapse (high sensitivity), but a negative result should not exclude these conditions (moderate specificity) Recommendations for practice: 50 µg/g of FC should be the cut-off point for detecting IBD Recommendations for research: none

TABLE 5 Characteristics and conclusions of previous reviews (*continued*)

Study	Inclusion criteria and methodology	Included studies	Conclusions
	Data analysis: text and tables; distinguish between IBD in general, CD, UC, assessment at first diagnosis or to assess activity/relapse		
Van Rheenen 2010[53] *Focus*: FC for investigation of suspected IBD Overall quality: medium	*Inclusion criteria* Study design: diagnostic accuracy studies Participants: the authors state that patients with IBD suspected on clinical grounds, with previously diagnosed IBD were to be excluded; however, at least one study, Bunn, was included despite most patients having previously confirmed IBD; studies with healthy controls also excluded Diagnostic procedure: stool sampling (for FC, index test) before endoscopic evaluation including histopathological verification of segmental biopsies (reference standard) Outcomes: sensitivity, specificity, PLR, NLR *Methodology* Search strategy: MEDLINE, EMBASE for studies published up to October 2009; search terms indicated; English language only; reference lists checked Study selection: first selection by one reviewer; full-text articles checked for eligibility by two independent reviewers; disagreements resolved by discussion; selection based partly on having spectrum of patients relevant to question Quality assessment: QUADAS (seven most differentiating items), no details of duplicate assessment but looks to have been done thoroughly Data extraction: items extracted were reported; no details of duplicate extraction Data analysis: meta-analysis, ROC curves; text and tables; distinguish between adults and children	No. of included trials: 13 No. of participants: 670 adults, 371 children/adolescents Trial quality: studies in children/adolescents were better quality than studies in adults; one study fulfilled all seven criteria, four fulfilled six of seven, four fulfilled five of seven, two fulfilled four of seven, and one each fulfilled three and two of seven; all studies reported FC followed by endoscopy Participants: six studies in adults, seven in children/adolescents; prevalence of IBD of between 14% and 80% (32% of adults, 61% of children/adolescents); all studies were from hospital clinics Diagnostic procedure: cut-off values for FC ranged between 24 and 150 µg/g	Conclusions: testing for FC is a useful tool for identifying patients who are most likely to need endoscopy for suspected IBD; the discriminatory power to safely exclude IBD was significantly better in studies of adults than in studies of children; at a tertiary-care level, FC can contribute important information In adults, an abnormal FC result gave 91% probability of IBD and a normal one gave a 3% probability Recommendations for practice: the authors reserved judgement about the utility of FC in primary care, given the lack of studies in primary care Recommendations for research: none

TABLE 5 Characteristics and conclusions of previous reviews (*continued*)

Study	Inclusion criteria and methodology	Included studies	Conclusions
Van Roon 2007[54] *Focus*: diagnostic precision of FC for IBD and colorectal cancer in adults and children Overall quality: high	*Inclusion criteria* Study design: diagnostic studies with a control group Participants: patients with CD, UC or CRC compared with healthy patients or those with IBS Diagnostic procedure: FC compared with histological diagnosis Outcomes: sensitivity, specificity, area under the SROC curve, DOR *Methodology* Search strategy: MEDLINE, EMBASE, Cochrane Library for studies published up to March 2006; search terms indicated; no language restrictions; reference lists checked; funnel plot suggested no publication bias Study selection: not reported Quality assessment: QUADAS, no details of duplicate assessment Data extraction: data extracted independently by two authors; in case of disagreement, consensus was reached through discussion with the senior author Data analysis: meta-analysis, ROC curves; heterogeneity assessment; text and tables; distinguish between adults and children, IBD in general, CD, UC, colorectal neoplasia	No. of included trials: 30 No. of participants: 5983 Trial quality: 19 studies rated high quality (QUADAS score > 11); range 10 to 13 Participants: 22 studies in adults, 1 in adults and children, 7 in children Diagnostic procedure: two studies assessed diagnostic precision in predicting relapse and three in examining disease activity; cut-off values for FC ranged between 18.6 and 250 µg/g Results: sensitivity analyses showed that high-quality studies (QUADAS score > 11) had higher sensitivity – 0.90 vs. 0.71 (adults 50 µg, IBD vs. no IBD) when all studies included; no different in specificity. Large studies (> 100) also gave higher sensitivity	Conclusions: FC cannot be recommended as a screening test for colorectal cancer in the general population; FC appeared to offer a good diagnostic precision in distinguishing IBD from non-IBD diagnoses with a higher precision at a cut-off of 100 µg/g; FC in patients with IBS was no different from in healthy controls FC was better for CD than UC, and better in children Recommendations for practice: none Recommendations for research: high-quality study needed investigating different cut-off points for FC
Henderson 2013[70] *Focus*: the value of FC testing in children being investigated for suspected IBD Overall quality: high	*Inclusion criteria* Study design: retrospective or prospective case–control studies Participants: children with suspected bowel inflammation (PIBD) who underwent at least colonoscopy Diagnostic procedure: FC compared with ileocolonoscopy or upper endoscopy Outcomes: sensitivity, specificity, PLR, NLR, ROC curve *Methodology* Search strategy: MEDLINE was searched up to May week 3 2012; EMBASE up to week 25 2012; PubMed, Google Scholar and	No. of included trials: 8 No. of participants: 715 (394 PIBD patients and 321 non-PIBD) Trial quality: one study each fulfilled nine, eight and six criteria, respectively; two studies fulfilled five criteria, whereas three studies fulfilled two criteria; the studies that did not fulfil most criteria had most items unclear. Only three studies had a representative spectrum of patients Participants: paediatric patients with suspected IBD. More had CD than	Conclusions: FC is a useful tool to screen children with suspected bowel inflammation; the test may lower endoscopy rates thereby benefiting both parents and children Recommendations for practice: Recommendations for research: studies to see if FC testing reduces endoscopy rates and assess cost benefits, and studies of the usefulness of FC in disease monitoring

TABLE 5 Characteristics and conclusions of previous reviews (*continued*)

Study	Inclusion criteria and methodology	Included studies	Conclusions
	The Cochrane Library were searched; search strategy available on request; reference list checked, personal collections and meeting abstracts were checked (only full-text articles were included); no language restrictions (foreign-language articles were translated using Google Translate)	UC: ratio of CD to UC about 1.5 : 1 Diagnostic procedure: in six studies, the cut-off value was 50 µg/g, whereas in two studies, the cut-off was 100 µg/g	
	Study selection: studies evaluated by two reviewers independently for eligibility, any discrepancies were resolved by discussion,		
	Quality assessment: modified version of the QUADAS tool (11 questions)		
	Data extraction: data entered into a customised database. Authors were contacted if certain parameters were uncertain mainly during the construction of 2 × 2 table		
	Data analysis: meta-analysis, 2 × 2 table, sensitivity, specificity, ROC curve, no analysis of heterogeneity		

BNI, British Nursing Index; CINAHL, Cumulative Index to Nursing and Allied Health Literature; SROC, summary receiver operating characteristic.

The 2010 review by YHEC[50] for CEP provides a good starting point, as it was done to inform the debate about the value of calprotectin in identifying people whose symptoms were due to IBS, and who therefore did not need expensive and invasive investigations, such as colonoscopy.

The YHEC review[50] sets the scene and makes many useful points, including:

- The key issue is deciding which patients should be referred for endoscopic or radiological examinations. The usual definitive diagnosis is by colonoscopy and histology, but that it invasive, unpleasant, expensive and with a risk, admittedly now small, of serious complications. There may also be long waiting lists (p. 7). One study reported that up to 40% of new GI referrals are for suspected IBS (p. 47).
- If a non-invasive test such as calprotectin could rule out IBD, patients would be spared endoscopy and might receive appropriate reassurance and treatment much earlier.
- For this to happen, calprotectin needs to have a high NPV, so that IBD can be ruled out (p. 8).
- Some patients with IBS do have biochemical evidence of inflammation (p. 5), and IBS may cover several subgroups. Some studies have reported higher calprotectin levels in patients with IBS than in healthy controls, although the differences have not always been significant (p. 20), and the levels in IBS are still well below a cut-off of 50 µg/g (p. 25).
- The Rome criteria for diagnosing IBS may be met by many patients with organic disease, resulting in misdiagnosis and failure to refer (p. 16).
- FC is a marker of intestinal inflammation, not a test for organic versus non-organic disease (p. 16).
- Most studies were from secondary care, and selection by GPs of patients likely to have organic disease may mean that results in secondary care may not be applicable to the different patient mix seen in primary care (p. 20).

TABLE 6 Quality of previous reviews

Review	Clear definition of review question (PICO)	Search strategy adequate and appropriate	Minimisation or error and bias in study selection	Appropriate quality assessment of included studies (e.g. QUADAS)	Appropriate data extraction process	Sufficient detail on primary studies	Appropriate methods used for data synthesis and comparison between studies	Conclusions reflect evidence reviewed
CEP 2010[50]	No	Yes	Not reported	Not reported	Not reported	Yes	Unclear	Yes
Jellema 2011[68]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kostakis 2012[69]	Yes	Partially	Not reported	No	Not reported	Partially	Yes	Yes
Van Rheenen 2010[53]	Yes	Partially	Partially	Yes	Yes	Yes	Yes	Yes
Van Roon 2007[54]	Yes	Yes	Not reported	Yes	Yes	Yes	Yes	Yes
Henderson 2013[70]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

PICO, participants, interventions, comparisons, outcomes.
1–4 criteria met, low quality; 5–6 criteria met, medium quality; 7–8 criteria met, high quality.

TABLE 7 Results of previous reviews

Review	Results (sensitivity is for diagnosing IBD)
CEP 2010[50]	**Sensitivity:** 63–100%
	Specificity: 37–100%
	PPV: 60–100%
	NPV: 51–100%
Jellema 2011[68]	**Sensitivity:** 84–100% in seven studies, 61% and 64% in two studies
	Specificity: 71–100%
Kostakis 2012[69]	***Newly diagnosed and untreated IBD:***
	ALL
	Sensitivity: 73.5–100% (95.8–100% for 50 µg/g as cut-off point, 73.5–100% for 100 µg/g as cut-off point)
	Specificity: 65.9–100% (65.9–92.9% for 50 µg/g as cut-off point, 69.2–100% for 100 µg/g as cut-off point)
	PLR: 2.8–34.9 (2.9–14 for 50 µg/g as cut-off point, 2.8–34.9 for 100 µg/g as cut-off point)
	NLR: 0–0.3 (0 for 50 µg/g as cut-off point, 0–0.3 for 100 µg/g as cut-off point)
	UC
	Sensitivity: 75–100%
	Specificity: 65.9–92.9%
	PLR: 2.4–14
	NLR: 0–0.4
	CD
	Sensitivity: 93.3–100%
	Specificity: 65.9–92.9%
	PLR: 2.9–14
	NLR: 0–0.1
	Already diagnosed and under treatment IBD:
	Sensitivity: 12.5–100% (100% for 50 µg/g as cut-off point, 12.5–68.2% for 100 µg/g as cut-off point)
	Specificity: 58.3–100% (58.3–80% for 50 µg/g as cut-off point, 69.2–100% for 100 µg/g as cut-off point)
	PLR: 1.1–5 (2.4–5 for 50 µg/g as cut-off point, 1.1 for 100 µg/g as cut-off point)
	NLR: 0–1 (0 for 50 µg/g as cut-off point, 0.9–1 for 100 µg/g as cut-off point)
	FC levels are much higher in patients with active IBD (newly diagnosed without treatment of under treatment with relapse) than in patients with IBD in remission, but FC levels in patients with inactive IBD are higher than those of healthy controls or patients with functional disorders or other GI diseases

TABLE 7 Results of previous reviews (*continued*)

Review	Results (sensitivity is for diagnosing IBD)
Von Rheenen 2010[53]	***Adults:*** **Sensitivity:** 93% (95% CI 85 to 97) **Specificity:** 96% (95% CI 79 to 99) **PLR:** 20 **NLR:** 0.06 ***Children/adolescents:*** **Sensitivity:** 92% (95% CI 84 to 96) **Specificity:** 76% (62% CI 79 to 86) **PLR:** 5 **NLR:** 0.1
Von Roon 2007[54]	***Adults and children (cut-off 50 µg/g):*** **Sensitivity:** 89% (95% CI 86% to 91%) **Specificity:** 81% (95% CI 78% to 84%) ***Adults and children (cut-off 100 µg/g):*** **Sensitivity:** 98% (95% CI 93% to 99%) **Specificity:** 91% (95% CI 86% to 95%) The diagnostic precision was higher in children than adults and at a cut-off of 100 vs. 50 µg/g In adults, cut-off 50 µg, UC sensitivity 0.78, specificity 0.78 CD adults, sensitivity 0.95 and specificity 0.85 at cut-off 50 µg Children with CD, sensitivity 0.97 and specificity 0.79 at cut-off 50 µg; sensitivity 1.0 and specificity 0.98 at cut-off 100 µg
Henderson 2013[70]	***Pooled sensitivity and specificity*** **Sensitivity:** 0.978 (95% CI 0.947 to 0.996) **Specificity:** 0.682 (95% CI 0.502 to 0.863) **PLR:** 3.07 **NLR:** 0.03

- The YHEC report[50] considered that high sensitivity was very important and that false positives were preferable to false negatives.
- The upper reference limit for absence of disease was suggested as 50 µg/g.
- When using point-of-care tests, borderline or elevated results should be re-examined using a quantitative method (p. 22).
- Calprotectin testing was much better than blood tests, CRP and ESR, with NPVs 89%, 68% and 69%, respectively (p. 27). The best blood test was CRP but it was effective in only 53% of patients (p. 48). In cost-effectiveness analysis, calprotectin dominated CRP and ESR (p. 55), giving more correct diagnoses at less cost (p. 55, tables 32 and 33).

Further details of the YHEC review[50] and other reviews are given in *Table 5*. Reporting of several aspects of the review was scanty. However, it should be noted that the YHEC remit[50] was restricted and did not include doing a full systematic review to standards such as those of the Cochrane reviews.

The YHEC cost-effectiveness analysis[50] was refined and extended by Mascialino *et al.* (Thermo Fisher Scientific 2013, personal communication and conference poster) by the inclusion of another arm, which had patients who had indeterminate calprotectin results on first test. This arm was populated with unpublished Swedish data. They also used a wider range of data on sensitivity and specificity of calprotectin, CRP and ESR than the single paper by Tibble *et al.*[49] used by YHEC.[50] The conclusions were the same as those of YHEC[50] – calprotectin dominates ESR and CRP.

The review by Van Rheenen *et al.*[53] appeared to match our main interest, as it was reported to be about the value of calprotectin in the investigation of suspected IBD, with a view to determining whether it reduced the number of unnecessary endoscopies. It also appeared to be a high-quality review. However, not all of the included studies were of newly presenting patients. Bunn *et al.*[71] had more patients with confirmed IBD than new patients. Kolho *et al.*[72] enrolled a group of newly presenting patients but only 30 of the 132 stool samples were taken at presentation, with others being taken after treatment, as long as 72 weeks later. So the patient group was correct, but the timing of testing was not always suitable for our purposes.

One advantage of the Van Rheenen review[53] for our purposes was that, overall, only 32% of adults with symptoms were found to have IBD. That proportion may be more similar to the mix of patients seen in primary care than some studies from specialist care. A disadvantage is that only two studies in adults excluded patients with rectal bleeding. Such patients would normally be referred for GI investigation on 'red flag' grounds and so are outwith the remit of this review. However, bleeding seems to be quite common in people with IBS. For example, Otten *et al.*[73] report that 26% of the group confirmed as having IBS, had rectal bleeding. Other studies in the Van Rheenen review[53] mentioned rectal bleeding but did not give proportions.[74–76]

Van Rheenen *et al.*[53] reported that the pooled results gave sensitivity of 93% (95% CI 0.85% to 0.97%) and specificity of 0.96% (0.79% to 0.99%). Screening by calprotectin would reduce the number of adults requiring endoscopy by 67% but they estimated that 3% without IBD would have endoscopy and 2% with IBD would not have endoscopy and so have the disease missed.

However, they appeared to have pooled results at all cut-off levels, so that they pooled Schroeder *et al.*,[77] which used a cut-off of 24 µg/g, with Limburg *et al.*,[78] which used a cut-off of 100 µg/g. Such pooling does not seem appropriate.

The adult results reflect the high proportion with IBS. The results in children differ because only about 7% had IBS and 61% had IBD. Van Rheenen *et al.*[53] estimated that the number requiring endoscopy would be reduced by only 35%, with 9% of those without IBD having endoscopy, and 5% with IBD being missed.

Van Rheenen *et al.*[53] noted that most studies were from secondary care, and provide a Fagan plot so that results for a population more representative of that seen in primary care can be estimated. From this, they expect that given a primary care expected prevalence of 5% with IBD, the NPV would be over 99.8%, good for ruling out IBD. However, the PPV falls to 55%. (The cut-off level is not clear, since results are described as normal or not normal.)

Van Rheenen *et al.*[53] concluded that calprotectin is a useful test for identifying those most likely to need endoscopy.

An earlier review, by Van Roon *et al.*[54] had a broader remit, examining the value of calprotectin in the diagnosis of both IBD and colorectal cancer. It was a high-quality review. The approach was less suited to our purposes, as they included studies with healthy controls, and others in patients with known IBD. Some studies did not include people with IBS. Nevertheless, some useful findings were that:

- The sensitivity of CRP was low, ranging from 35% to 40%.
- The sensitivity of ESR was also low, 18% to 52%.
- For IBD, a cut-off of 100 µg/g gave slightly better precision than 50 µg/g, with AUCs of 0.98 and 0.95.
- Calprotectin at a cut-off of 50 µg/g performed well for differentiating between those with IBS and healthy controls, and those with IBD, AUC 0.97, with slightly higher precision at cut-off of 100 µg/g.
- Sensitivity for CD was higher than for UC (CD 0.95 in adults and 0.97 in children at 50 µg/g cut-offs, and UC 0.78).
- Levels of calprotectin in people with IBS were similar to those in healthy controls.
- Calprotectin could not be recommended as a screening test for colorectal cancer.
- A sensitivity analysis excluding lower-quality studies improved the sensitivity without affecting specificity, as did excluding smaller studies.

Von Roon *et al.*[54] also pooled results but, more correctly, pooling only studies using the same cut-offs. The pooling did not include the grouping we would have found most useful. They pooled IBD compared with 'not IBD', and CD versus a mix of healthy controls and IBS. And most of the studies they included were not in newly presenting patients. The data shown in *Table 4* come from their Table 3.

Note the suggestion that calprotectin may be less sensitive in UC than CD.

Von Roon *et al.*[54] noted some weaknesses in the evidence, including spectrum bias, commenting that:

> *FC has a good diagnostic precision for separating IBD from non-IBD diagnoses overall. Whilst this finding is likely to hold true in patients with severe IBD, it may not necessarily translate to a clinical setting where the patient has a low pre-test probability of IBD, i.e. where a clinician is attempting to differentiate patients with functional abdominal pain syndromes or IBS from IBD patients with mild 'functional-like' symptoms.*

Jellema *et al.*[68] set out to do a systematic review on the diagnosis of IBD in primary care, in adults only. Their intention appears to have been to exclude studies in patients with established IBD. In order to increase relevance to primary care, they excluded studies in which the prevalence of IBD was more than 25%, although as they point out, even that would be a high prevalence for a primary care population. (Though, as noted above, we need to take into account the difference between the prevalence of IBD in the whole primary care population, and the prevalence in those selected by GPs for referral to specialist care.)

Unfortunately, few of their 24 included studies were carried out in primary care – only three partly in that setting. It was a high-quality review. No pooling of results was done. Useful findings included:

- Symptoms associated with IBD (diarrhoea, abdominal pain, blood in stools, weight loss) provided individually poor sensitivity and specificity.
- Among blood and faecal tests, calprotectin performed best.
- The performance of CRP was very variable, with sensitivity ranging from 0.55 to 1.0 depending on cut-offs; specificity ranging from 0.42 to 0.90. ESR was similar.

Jellema *et al.*[68] had reservations about applying results from specialist care to primary care:

> *In a setting with low disease prevalence, the same combination of sensitivity and specificity will lead to much lower positive predictive values compared with a setting with a high disease prevalence.*

Kostakis *et al.*[69] reviewed the evidence on FC in PIBD. Few details of methods were given, so the quality score was low. No data were given by type of control – which could be healthy children or 'other GI disease'. They included some studies with no controls. They concluded that the cut-off should be 50 μg/g rather than 100 μg/g, on the basis of slightly higher sensitivity (95.8–100% vs. 87–100%, after excluding an outlier study) but similar specificity (68–93% vs. 69–94%). No pooling of results was done.

The most recent systematic review comes from Henderson *et al.*,[70] and was of paediatric studies. It was a high-quality review, enhanced by the contacting of authors for further information. This meant that they could include a study (Perminow *et al.*;[79] which we did not) after they obtained unpublished details. The selection was rigorous, with children required to have at least colonoscopy. This meant excluding a study (Van de Vijver *et al.*[80]) in which children with negative FCs did not have colonoscopy but were instead followed up for 6 months. As will be reported below, we were less rigorous and allowed this to be included.

Henderson *et al.*[70] included eight studies with a total of 715 subjects. Quality was assessed using a modified QUADAS checklist, with no studies achieving full marks, and with spectrum bias being one problem, attributed to selection bias amongst referrals to tertiary centres. Most studies used a cut-off of 50 μg/g. The authors concluded that FC testing had high sensitivity of almost 98%, with reasonable specificity of 68%. PLR was 3.07, NLR 0.03. They noted that FC testing was inexpensive (their local cost being about £28 including labour costs). This compares with the cost of day case endoscopic assessment in children of £1500, and the additional costs of small bowel imaging.

Summary
Some reviews are now out of date. The most recent ones (YHEC;[50] Jellema *et al.*;[68] Henderson *et al.*[70]) all conclude that FC testing is very useful. Henderson *et al.*[70] focuses only on use in children but is right up to date, and very high quality.

The tests

Table 8 shows the calprotectin tests included, and the studies of each included in the following sections, below: *Studies of calprotectin in the differentiation of inflammatory bowel disease and irritable bowel syndrome*; *Studies of calprotectin: organic versus irritable bowel syndrome*; *Studies of calprotectin: inflammatory bowel disease versus non-inflammatory bowel disease*; and *Studies of calprotectin: organic versus non-organic bowel disease*. Note that the numbers of studies apply only to those that we could include in our meta-analyses. There are other studies of these tests, and indeed we include some elsewhere in this report.

The comparisons

The decision problem concerns the use of calprotectin to help distinguish between inflammatory and non-inflammatory bowel conditions. For GPs, this is part of distinguishing between patients who need to be referred to secondary care and those who can be managed in primary care. However, in practice, the distinction in adults is usually between patients at the more troublesome end of the IBS spectrum and IBD; we start with that in the next section (see *Studies of calprotectin in the differentiation of inflammatory bowel disease and irritable bowel syndrome*, below). In adult medicine, this is the most important comparison.

TABLE 8 Evidence base for the calprotectin tests

Name of test	Type of test	Evidence base
Nycomed Pharma	ELISA	IBS vs. IBD: one study, El-Badry 2010[81]
		IBD vs. non-IBD: two studies – Limberg 2000;[78] Sidler 2008[82]
		Organic vs. IBS: none
		Organic vs. non-organic: none
Immundiagnostik ELISA kit	ELISA	IBS vs. IBD: Basumani 2012,[83] unpublished; Schroder 2007[77]
		IBD vs. non-IBD: none
		Organic vs. IBS: Basumani 2012[83]
		Organic vs. non-organic: none
EK-CAL	ELISA	IBS vs. IBD: none
		IBD vs. non-IBD: Damms 2008[58]
		Organic vs. IBS: none
		Organic vs. non-organic: Manz 2012;[84] Kok 2012;[39] Burri 2013[85]
Calprest	ELISA	IBS vs. IBD: none
		IBD vs. non-IBD: five studies – Fagerberg 2008;[66] Diamanti 2010;[86] Tomas 2007;[66] Canani 2006;[65] Licata 2012[62]
		Organic vs. IBS: Carroccio 2003[87]
		Organic vs. non-organic: Tomas 2007;[66] Shitrit 2007;[62] Garcia 2006[59]
Calpro Calprotectin ELISA test (ALP)	ELISA	IBS vs. IBD: Otten 2009;[73] Schoepfer 2008;[75] Li 2006[60]
		IBD vs. non-IBD: Van der Vijver 2012;[80] Henderson 2012[30]
		Organic vs. IBS: none
		Organic vs. non-organic: none
Not known	ELISA	IBS vs. IBD: Bharathi 2005[88]
		IBD vs. non-IBD: Ashorn 2009[89]
Quantum Blue	POCT	IBS vs. IBD: none
		IBD vs. non-IBD: none
		Organic vs. IBS: none
		Organic vs. non-organic: Kok 2012[39]
Prevent ID Caldetect	POCT	IBS vs. IBD: Otten 2009[73]
		IBD vs. non-IBD: none
		Organic vs. IBS: none
		Organic vs. non-organic: Lee 2013[90]
Prevista (no longer available)	POCT	IBS vs. IBD: none
		IBD vs. non-IBD: Damms 2008[58]
		Organic vs. IBS: none
		Organic vs. non-organic: none
EliA platform	EliA	None

In *Studies of calprotectin: organic versus irritable bowel syndrome*, below, we look at another way of distinguishing between patients who should be referred, and those with IBS, in two studies that compare 'organic' with IBS. In adult medicine there are other organic causes that can cause symptoms such as colorectal neoplasia.

Note that we are assuming that in the situations in which calprotectin would be used, coeliac disease has already been detected or ruled out by blood testing. Coeliac disease is a bowel disease characterised by inflammation but would not have high calprotectin levels because the inflammation is mediated by lymphocytes, not neutrophils. In children with coeliac disease, calprotectin may be mildly elevated.

In paediatrics, studies aim to distinguish between IBD and non-IBD, as IBS is much less common. Some adult studies also make this comparison, but most studies of IBD versus non-IBD come from paediatric gastroenterology. These are dealt with in *Studies of calprotectin: inflammatory bowel disease versus non-irritable bowel syndrome*, below.

In *Studies of calprotectin: organic versus non-organic bowel disease*, below, we include 'organic versus non-organic'. This comparison is less relevant because the organic group can contain a mixture of inflammatory and non-inflammatory conditions. We deal with this group in less detail.

Studies of calprotectin in the differentiation of inflammatory bowel disease and irritable bowel syndrome

We included seven studies[60,73,75,77,81,83,88] in this group, shown in the paired forest plots below (see *Figure 3*). One of these studies (Basumani *et al.*[83] and P Basumani, Rotherham Hospital, 2012, personal communication) is not yet published. One study (Bharathi *et al.*),[88] available only as an abstract, gave no detail of clinical setting. Only three studies[73,83,88] gave data at cut-offs of other than 50 µg/g, and one did not provide enough data to calculate sensitivity. As expected, low thresholds gave high sensitivity for IBD but poor specificity. The studies were in adults.

All used ELISA tests, and one[73] also used a point-of-care test.

Note that numbers in the tables reflect total numbers in each study, and not all may be relevant for our purposes. Numbers in forest plots (see *Figures 3*, *9, 12 and 19*) will sometimes be smaller than numbers in the studies. For example, some studies included 'healthy controls', who are not relevant to this review.

Table 9 gives brief details of the IBD versus IBS studies, *Table 10* shows their QUADAS assessments, and *Table 11* their results.

Figure 3 provides dual forest plots for the IBD versus IBS studies. *Figure 4* provides dual forest plots for the Basumani-only study,[80] nicely illustrating the trade-off between sensitivity and specificity, and *Figure 5* provides the ROC curve for the Basumani data.[80] (One point is missing because points 1 and 2 are the same.)

Figure 6 gives the ROC curve at 50 µg/g.

The most useful study was that of Basumani *et al.*[83] (see *Figure 4*), because it provided data at six cut-offs, as shown for clarity below. At the lower levels, as expected, sensitivity is high, but specificity low.

The summary point shows the summary sensitivity, and the confidence contour shows the CI or region for the summary point.

TABLE 9 Outline of studies comparing IBD vs. IBS

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Basumani 2012[83]	119	New referrals with chronic diarrhoea	District General Hospital, Yorkshire	To assess value of FC as screening tool to avoid colonoscopy	Histology	No FC FC < 8 No histology
Schroder 2007[77]	76	Diarrhoea for more than 4 weeks, unknown cause unknown	Hospital	To assess utility of FC to detect inflammation in patients with ?IBD, ?IBS	Colonoscopy and biopsy	Previous investigations, GI bleeding, polyps, pregnancy
Otten 2008[73]	114	Consecutive patients referred with lower abdominal symptoms, referred to endoscopy unit	Endoscopy unit, The Netherlands	To evaluate POCT FC and lactoferrin tests for assessing inflammation; and differentiating IBS and IBD	Colonoscopy and biopsy	Age < 18 years; previous colon surgery; iron deficiency
Schoepfer 2008[75]	94	Outpatients and inpatient	Gastroenterology Department, University Hospital, Switzerland	To assess accuracy FC and lactoferrin to detect inflammation in patients with ?IBS, ?IBD	Colonoscopy including terminal ileum and biopsies	Incomplete colonoscopy, microscopic colitis, no FC sample, infections, polyps, aspirin or NSAIDs, etc.
El-Bady 2010[81]	29	GI symptoms for at least 6 months, and endoscopy necessary to exclude organic pathology	Internal Medicine Department, Cairo	To evaluate FC at different cut-offs for differentiation functional and organic disorders	Colonoscopy into ileum with biopsies	NSAIDS, aspirin, anticoagulants, arthritis and other diseases affecting FC
Li 2006[60]	240	Outpatients and inpatients with IBS or IBD, healthy controls; patients followed up after polyp removal with no recurrence	Hospital, Peking	To assess FC in differential diagnosis of IBS and exclude organic diseases	Colonoscopy with biopsy in IBD group	No upper GI symptoms; adenomas; severe other disease
Bharathi 2005[88]	58	Patients presenting with abdominal pain or loose stools	Not reported	To assess NPV of FC for excluding bowel pathology in young patients with ?IBS	Various – endoscopy, ultrasound	Not reported

TABLE 10 QUADAS quality assessment of studies comparing IBD vs. IBS

Quality criterion	Basumani 2012[83]	Schroder 2007[77]	Otten 2008[73]	Schoepfer 2008[77]	El-Badry 2010[81]	Li 2006[60]	Bharathi 2005[88]
Spectrum	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Reference standard	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acceptable delay?	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear
Whole sample verified?	Yes	Yes	No	Yes	Yes	Yes	No
Same reference standard	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes
Newly diagnosed?	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Blinded reference testing?	Unclear	Yes	Yes	Yes	Unclear	Unclear	Unclear
Index results blinded?	Unclear	Yes	Yes	Yes	Unclear	Unclear	Unclear
Same clinical data	Unclear	Yes	Yes	Yes	Unclear	Yes	Unclear
Intermediate results reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Withdrawals explained?	Yes	Yes	Yes	Yes	Yes	Unclear	Yes

The prediction contour outlines the prediction region for the true sensitivity and specificity in a future study. (For details see Harbord and Whiting.)[55]

Figure 7 shows the pooled forest plots for sensitivity and specificity at 50 µg/g and *Figure 8* shows the nomogram with LRs.

It should be noted that some experts advise caution in the use of the I^2 test for indicating heterogeneity in reviews of diagnostic test accuracy and suggest that they should not be routinely used (see Cochrane Handbook for systematic reviews of diagnostic accuracy v1.0 December 2012, paragraph 10.4.3).[56]

Table 12 shows DORs for IBD versus IBS.

Conclusions of section 'Studies of calprotectin in the differentiation of inflammatory bowel disease and irritable bowel syndrome'
Calprotectin testing appears very useful for differentiating between IBS and IBD.

Almost all sensitivities are high, the outlier being Otten *et al.*[73] with the 60 µg/g cut-off using a POCT. As expected, and shown best by the Basumani *et al.*[83] data in *Figure 5*, there is a trade-off between sensitivity and specificity.

The only point-of-care test in the group is the PreventID, which performed well at the 15 µg/g cut-off but not so well at the 60 µg/g cut-off, though it is curious that its specificity should be so high at the lower cut-off. The POCT with a 15 µg/g cut-off had higher specificity (95%) than the ELISA at 50 µg/g (87%).

The variability amongst sensitivities was much less than amongst specificities. Heterogeneity was moderate for sensitivity (I^2 = 37%) but high for specificity (94%). (However, see earlier note about the I^2 test in diagnostic reviews.) Even using the same PhiCal ELISA test (Calpro, Lysaker, Norway) with the 50 µg/g cut-off, sensitivities ranged from 83% to 96%.

TABLE 11 Results of studies comparing IBD vs. IBS

Study	Cut-off value (µg/g)	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)	Accuracy	Disease prevalence, % (95% CI)
Basumani 2012[83]	<8	0.20 (0.11 to 0.32)	1.00 (0.93 to 1.00)	2.04 (1.67 to 2.50)	0	0.56	10.91 (5.77 to 18.28)
Basumani 2012[83]	<25	0.20 (0.11 to 0.32)	1.00 (0.93 to 1.00)	2.04 (1.67 to 2.50)	0	0.56	10.91 (5.77 to 18.28)
Basumani 2012[83]	<50	0.24 (0.13 to 0.37)	1.00 (0.94 to 1.00)	2.51 (1.97 to 3.21)	0	0.65	10.91 (5.77 to 18.28)
Basumani 2012[83]	<75	0.31 (0.16 to 0.48)	0.99 (0.93 to 1.00)	3.59 (2.46 to 5.25)	0	0.76	10.91 (5.77 to 18.28)
Basumani 2012[83]	<100	0.38 (0.21 to 0.58)	0.99 (0.93 to 1.00)	4.99 (3.18 to 7.83)	0.10 (0.02 to 0.67)	0.83	10.91 (5.77 to 18.28)
Basumani 2012[83]	<150	0.42 (0.22 to 0.63)	0.98 (0.92 to 1.00)	5.83 (3.38 to 10.08)	0.19 (0.05 to 0.69)	0.85	10.91 (5.77 to 18.28)
Bharathi 2005[88]	60	0.00 (0.00 to 0.27)	1.00 (0.92 to 1.00)			0.79	0.00 (0.00 to 6.16)
El-Badry 2010[81]	50	1.00 (0.81 to 1.00)	0.75 (0.43 to 0.95)		0.15 (0.05 to 0.43)	0.90	68.97 (49.17 to 84.72)
El-Badry 2010[81]	100						
Li 2006[60]	50	0.95 (0.86 to 0.99)	0.93 (0.84 to 0.98)	18.67 (6.18 to 56.63)	0.07 (0.03 to 0.18)	0.94	50.00 (40.74 to 59.26)
Otten 2008[73]	≥15	0.82 (0.63 to 0.94)	1.00 (0.96 to 1.00)	18.2 (7.7 to 42.7)	0	0.96	20.18 (13.24 to 28.72)
Otten 2008[73]	≥60	0.88 (0.60 to 0.98)	0.91 (0.83 to 0.96)	27.7 (6.7 to 113.3)	0.4 (0.2 to 0.7)	0.90	20.18 (13.24 to 28.72)
Otten 2008[73]	>50	0.65 (0.47 to 0.81)	0.99 (0.93 to 1.00)	7.25 (4.25 to 12.38)	0.05 (0.01 to 0.34)	0.89	20.18 (13.24 to 28.72)
Schoepfer 2008[75]	50	1.00 (0.93 to 1.00)	0.73 (0.57 to 0.86)		0.17 (0.10 to 0.29)	0.88	68.09 (57.67 to 77.33)
Schroder 2007[77]	15	1.00 (0.92 to 1.00)	0.91 (0.76 to 0.98)		0.07 (0.02 to 0.20)	0.96	59.21 (47.33 to 70.35)

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Basumani 2012–1[83]	12	48	0	50	8.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.51 (0.41 to 0.61)		
Schroder 2007[77]	42	0	3	31	15.0	PhiCal	Adult	Secondary	ELISA	0.93 (0.82 to 0.99)	1.00 (0.89 to 1.00)		
Otten 2008a[83]	23	5	0	86	15.0	Prevent ID	Adult	Secondary	POCT	1.00 (0.85 to 1.00)	0.95 (0.88 to 0.98)		
Basumani 2012–2[83]	12	48	0	50	25.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.51 (0.41 to 0.61)		
Basumani 2012–3[83]	12	39	0	59	50.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.60 (0.50 to 0.70)		
Otten 2008c[83]	22	12	1	79	50.0	PhiCal	Adult	Secondary	ELISA	0.96 (0.78 to 1.00)	0.87 (0.78 to 0.93)		
Schoepfer 2008[75]	53	0	11	30	50.0	PhiCal	Adult	Secondary	ELISA	0.83 (0.71 to 0.91)	1.00 (0.88 to 1.00)		
El-Badry 2010[81]	17	0	3	9	50.0	PhiCal	Adult	Secondary	ELISA	0.85 (0.62 to 0.97)	1.00 (0.66 to 1.00)		
Li 2006[60]	56	3	4	57	50.0	PhiCal	Adult	Secondary	ELISA	0.93 (0.84 to 0.98)	0.95 (0.86 to 0.99)		
Otten 2008b[73]	14	2	9	89	60.0	Prevent ID	Adult	Secondary	POCT	0.61 (0.39 to 0.80)	0.98 (0.92 to 1.00)		
Bharathi 2005[88]	0	12	0	46	60.0	PhiCal	Not known	Not known	ELISA	Not estimable	0.79 (0.67 to 0.89)		
Basumani 2012–4[83]	11	25	1	73	75.0	Immundiagnostik	Adult	Secondary	ELISA	0.92 (0.62 to 1.00)	0.74 (0.65 to 0.83)		
Basumani 2012–5[83]	11	18	1	80	100.0	Immundiagnostik	Adult	Secondary	ELISA	0.92 (0.62 to 1.00)	0.82 (0.73 to 0.89)		
Basumani 2012–6[83]	10	14	2	84	150.0	Immundiagnostik	Adult	Secondary	ELISA	0.83 (0.52 to 0.98)	0.86 (0.77 to 0.92)		

FIGURE 3 Inflammatory bowel disease vs. IBS.

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Basumani 2012–1[83]	12	48	0	50	8.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.51 (0.41 to 0.61)		
Basumani 2012–2[83]	12	48	0	50	25.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.51 (0.41 to 0.61)		
Basumani 2012–3[83]	12	39	0	59	50.0	Immundiagnostik	Adult	Secondary	ELISA	1.00 (0.74 to 1.00)	0.60 (0.50 to 0.70)		
Basumani 2012–4[83]	11	25	1	73	75.0	Immundiagnostik	Adult	Secondary	ELISA	0.92 (0.62 to 1.00)	0.74 (0.65 to 0.83)		
Basumani 2012–5[83]	11	18	1	80	100.0	Immundiagnostik	Adult	Secondary	ELISA	0.92 (0.62 to 1.00)	0.82 (0.73 to 0.89)		
Basumani 2012–6[83]	10	14	2	84	150.0	Immundiagnostik	Adult	Secondary	ELISA	0.83 (0.52 to 0.98)	0.86 (0.77 to 0.92)		

FIGURE 4 Inflammatory bowel disease vs. IBS: Basumani (2012)[83] data only.

FIGURE 5 Inflammatory bowel disease vs. IBS, ROC curve: Basumani (2012)[83] data only.

FIGURE 6 Summary receiver-operating characteristic curve for FC in the diagnosis of bowel diseases: IBD vs. IBS at a cut-off level of 50 μg/g.

Figure 6 provides the summary: sensitivity 93% and specificity 94%, for ELISA tests, at a 50 μg/g cut-off. These are based on five studies.[60,73,75,81,83] There was only one study for the 100 μg/g cut-off.[73] With an AUC of 0.97 at 50 μg/g, there is little room for improvement.

The only study using a POCT[73] performed well at cut-off 15 μg/g with sensitivity 100% and specificity 95%. At 60 μg/g, sensitivity was only 61%, which is unlikely to be acceptable given the importance of not missing people with IBD.

All studies were on adults.

On this evidence base, it may be unwise to recommend any ELISA cut-off other than 50 μg/g.
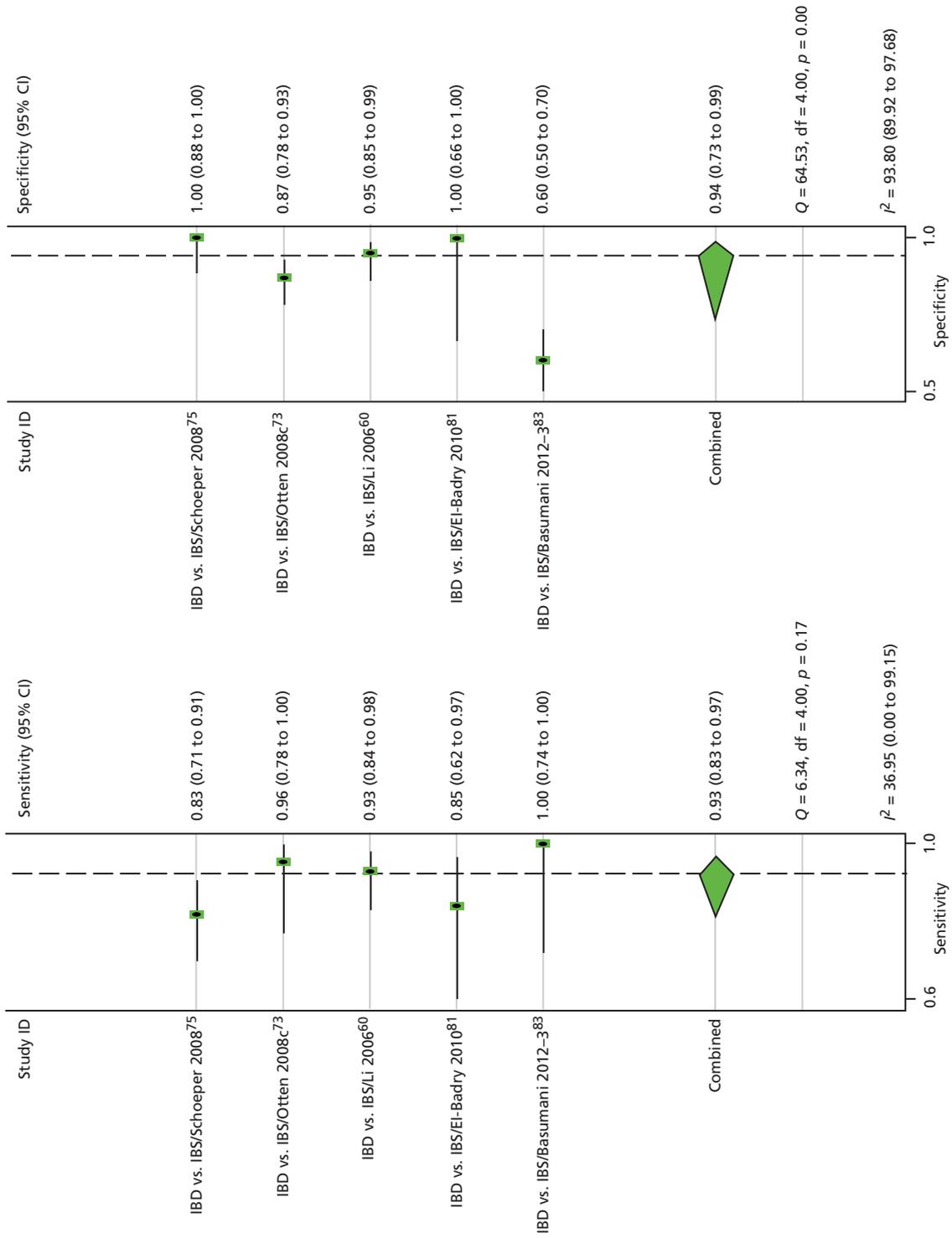
FIGURE 7 Forest plots of pooled sensitivity and specificity of FC in the diagnosis of bowel diseases: IBD vs. IBS at a cut-off level of 50 µg/g.
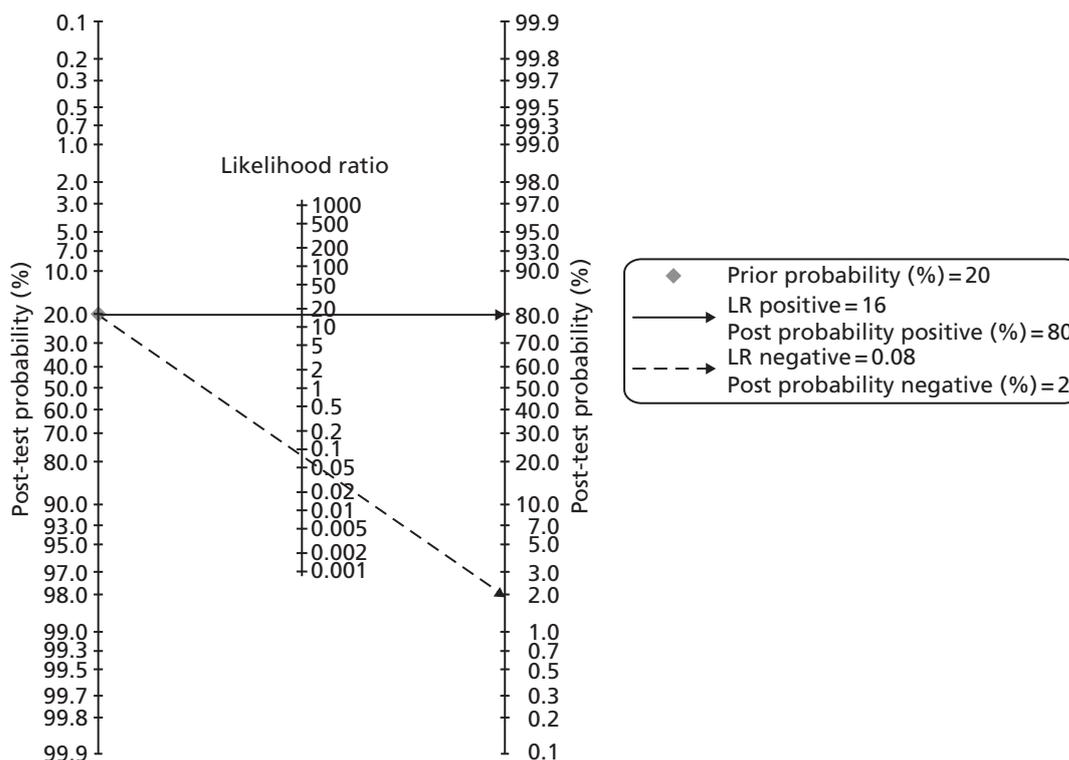
FIGURE 8 The use of the Fagan's nomogram (a straight line through the pre-test probability of 20% and the LR − of 0.20 yields a post-test probability of about 2%): IBD vs. IBS at a cut-off level of 50 µg/g.

TABLE 12 Diagnostic odds ratios: IBD vs. IBS

No. of patients	No. of studies	Cut-off level (µg/g)	DOR (95% CI)
119	1	8	21 (1 to 344)
76	1	15	10 (4 to 27)
119	1	25	21 (1 to 344)
596	5	50	202 (47 to 868)
119	1	75	32 (4 to 1400)
119	1	100	49 (6 to 2129)
119	1	150	30 (5 to 295)

Note the large CIs around all of the DORs.

Studies of calprotectin: organic versus irritable bowel syndrome

The term *organic* covers a range of conditions, and the range varies among studies (see *Tables 13* and *22*). Some of these conditions would not normally be regarded as inflammatory. Inflammation implies the presence of white blood cells, and one would not expect these in lesions such as colonic polyps. However, FC is often raised in patients with larger polyps (as shown later in *Figure 20*).

Table 13 gives details of the two studies in this group; *Table 14* gives their QUADAS assessments, and *Table 15* the results. *Figure 9* shows forest plots for both and *Figure 10* for the Basumani data only. *Figure 11* shows the ROC curves and *Table 16* has DORs at different cut-off levels.

TABLE 13 Outline of studies comparing organic vs. IBS

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Basumani 2012[83]	119	New referrals with chronic diarrhoea	District General Hospital, Yorkshire	To assess value of FC as screening tool to avoid colonoscopy	Histology	No FC FC < 8 No histology
Carroccio 2003[87]	120	Chronic diarrhoea for more than 4 weeks, with or without abdominal pain; unknown origin	Outpatient clinics of the University Hospital and of the Paediatric Division of 'Di Cristina' Hospital, Italy	To assess value of FC in identifying organic causes of chronic diarrhoea	All patients evaluated by the Rome criteria for IBS and haematology and chemistry tests. Adults under age of 40 years – sigmoidoscopy or colonoscopy with biopsy; children with positive occult blood in the stool or with serum indices of inflammation – colonoscopy and biopsy	Previous investigation: GI bleeding; familial adenomatous polyposis and hereditary non-polyposis; colorectal cancer syndrome; pregnancy

TABLE 14 Quality Assessment of Diagnostic Accuracy Studies quality assessment of studies comparing organic vs. IBS

Quality criterion	Basumani 2012[83]	Carroccio 2003[87]
Spectrum	Yes	Yes
Reference standard	Yes	Yes
Acceptable delay?	Unclear	Yes
Whole sample verified?	Yes	No
Same reference standard	Yes	Yes
Newly diagnosed?	Yes	Yes
Blinded reference testing?	Unclear	Yes
Index results blinded?	Unclear	Yes
Same clinical data	Unclear	Yes
Intermediate results reported?	Yes	Yes
Withdrawals explained?	Yes	Yes

So FC may flag up the presence of conditions other than IBD, such as some colorectal cancers and large adenomas, but results are more variable than with IBD. Therefore, in studies with a mix of organic conditions, calprotectin may not appear as reliable. However, this should not detract from its good performance in detecting IBD and excluding IBS.

The low sensitivity in the Carroccio study[87] may be partly due to their case mix, which is related to their institution's role as a referral centre for food intolerances. Their organic group included many (about one-third) with coeliac disease who had negative calprotectin tests.

TABLE 15 Results of studies comparing organic vs. IBS

Study	Cut-off value	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)	Accuracy	Disease prevalence, % (95% CI)
Basumani 2012[83]	<8 µg/g	0.29 (0.19 to 0.42)	0.98 (0.90 to 1.00)	1.94 (1.55 to 2.43)	0.09 (0.01 to 0.64)	0.59	17.65 (11.27 to 25.70)
Basumani 2012[83]	<25 µg/g	0.30 (0.20 to 0.43)	0.98 (0.90 to 1.00)	2.03 (1.61 to 2.56)	0.09 (0.01 to 0.61)	0.61	17.65 (11.27 to 25.70)
Basumani 2012[83]	<50 µg/g	0.33 (0.21 to 0.46)	0.97 (0.89 to 1.00)	2.27 (1.72 to 3.01)	0.16 (0.04 to 0.60)	0.66	17.65 (11.27 to 25.70)
Basumani 2012[83]	<75 µg/g	0.38 (0.23 to 0.54)	0.92 (0.84 to 0.97)	2.80 (1.82 to 4.32)	0.38 (0.19 to 0.78)	0.74	17.65 (11.27 to 25.70)
Basumani 2012[83]	<100 µg/g	0.44 (0.26 to 0.62)	0.92 (0.84 to 0.97)	3.63 (2.17 to 6.08)	0.41 (0.22 to 0.75)	0.79	17.65 (11.27 to 25.70)
Basumani 2012[83]	<150 µg/g	0.52 (0.31 to 0.72)	0.91 (0.84 to 0.96)	5.06 (2.70 to 9.47)	0.43 (0.25 to 0.75)	0.83	17.65 (11.27 to 25.70)
Carroccio 2003[87] (all patients)	50 µg/g	0.83 (0.70 to 0.92)	0.83 (0.70 to 0.92)	4.04 (2.17 to 7.53)	4.04 (2.17 to 7.53)	0.74	54.17 (44.83 to 63.29)
Carroccio 2003[87] (adults)	50 µg/g	0.70 (0.50 to 0.86)	0.74 (0.59 to 0.86)	3.17 (1.61 to 6.23)	0.46 (0.28 to 0.75)	0.73	42.86 (31.09 to 55.25)
Carroccio 2003[87] (children)	50 µg/g	0.96 (0.80 to 1.00)	0.58 (0.37 to 0.78)	10.71 (1.59 to 72.00)	0.31 (0.18 to 0.53)	0.78	70.00 (55.39 to 82.14)
Carroccio 2003[87] (all patients)	100 µg/g	0.88 (0.73 to 0.97)	0.59 (0.48 to 0.70)	6.35 (2.38 to 16.90)	0.58 (0.46 to 0.74)	0.68	54.17 (44.83 to 63.29)
Carroccio 2003[87] (adults)	100 µg/g	0.81 (0.54 to 0.96)	0.69 (0.55 to 0.81)	5.78 (1.81 to 18.48)	0.61 (0.44 to 0.85)	0.71	42.86 (31.09 to 55.25)
Carroccio 2003[87] (children)	100 µg/g	0.95 (0.74 to 1.00)	0.45 (0.27 to 0.64)	7.71 (1.13 to 52.66)	0.52 (0.36 to 0.75)	0.64	70.00 (55.39 to 82.14)

TABLE 16 Diagnostic odds ratios: organic vs. IBS

No. of patients	No. of studies	Cut-off level (µg/g)	DOR (95% CI)
119	1	8	20.8 (3.0 to 880.2)
119	1	25	22.6 (3.3 to 955.0)
239	2	50	3.3 (2.2 to 4.7)
119	1	75	7.3 (2.3 to 25.1)
239	2	100	2.7 (2.0 to 3.6)
119	1	150	9.8 (3.0 to 31.9)

Studies of calprotectin: inflammatory bowel disease versus non-inflammatory bowel disease

There were 11 studies in this group: eight in paediatric patients and three in adults. All used ELISA tests and one (Damms and Bischoff[58]) also used the Prevista POCT.

Details of the studies are shown in *Tables 17–19*: their QUADAS assessments are provided in *Table 18* and the results in *Table 19*. It should be borne in mind that symptoms of IBD in children may be 'subtle and atypical' (Sidler *et al.*[82]) rather than the typical diarrhoea, abdominal pain and weight loss. Impaired growth can be one presentation.

At a cut-off of 50 µg/g, the overall results pooled for IBD versus IBS, show very high sensitivity (99%: 95% CI 95% to 100%) (see *Figure 13*) but moderate specificity (74%), probably because there are organic conditions with raised calprotectin levels in the non-IBD group. There is moderate heterogeneity for sensitivity but high for specificity (see *Figure 14*).

At a cut-off of 100 µg/g, sensitivity falls to 94% (95% CI 86 to 98%) but specificity improves to 82% (95% CI 67% to 91%).

Henderson *et al.*[30] report results from a relatively large group of children, by linking referrals to the regional paediatric gastroenterology service (5600) with laboratory calprotectin results (4155 results) and endoscopy records, to create a cohort of 190 children investigated for possible IBD, who all had calprotectin and full endoscopy records. Ninety-one were shown to have IBD, of whom 62 had CD, 21 UC and the other eight unclassified IBD. The pre-test probability of IBD was 0.48. They classed calprotectin results as:

- normal ≤ 50 µg/g
- possible inflammation 51–100 µg/g
- GI inflammation 101–200 µg/g
- active GI inflammation > 200 µg/g.

and comment that in practice, they find the cut-off of 200 µg/g as being the most useful for likely diagnosis of IBD.

They provide results for four thresholds for positivity as follows in *Table 20*.

These results nicely show the trade-off between sensitivity and specificity.

(They also provide data for a cut-off of > 800 µg/g but sensitivity drops too much for that to be useful.)

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Basumani 2012a[83]	20	48	1	50	8.0	Immundiagnostik	Adult	Secondary	ELISA	0.95 (0.76 to 1.00)	0.51 (0.41 to 0.61)		
Basumani 2012b[83]	20	46	1	52	25.0	Immundiagnostik	Adult	Secondary	ELISA	0.95 (0.76 to 1.00)	0.53 (0.43 to 0.63)		
Basumani 2012c[83]	19	39	2	59	50.0	Immundiagnostik	Adult	Secondary	ELISA	0.90 (0.70 to 0.99)	0.60 (0.50 to 0.70)		
Carroccio 2003c[87]	25	1	10	14	50.0	Calprest	Paediatric	Secondary	ELISA	0.71 (0.54 to 0.85)	0.93 (0.68 to 1.00)		
Carroccio 2003a[87]	43	9	22	46	50.0	Calprest	Adult and paediatric	Secondary	ELISA	0.66 (0.53 to 0.77)	0.84 (0.71 to 0.92)		
Carroccio 2003b[87]	19	8	11	32	50.0	Calprest	Adult	Secondary	ELISA	0.63 (0.44 to 0.80)	0.80 (0.64 to 0.91)		
Basumani 2012d[83]	15	25	6	73	75.0	Immundiagnostik	Adult	Secondary	ELISA	0.71 (0.48 to 0.89)	0.74 (0.65 to 0.83)		
Basumani 2012e[83]	14	18	7	80	100.0	Immundiagnostik	Adult	Secondary	ELISA	0.67 (0.43 to 0.85)	0.82 (0.73 to 0.89)		
Carroccio 2003d[87]	30	4	35	51	100.0	Calprest	Adult and paediatric	Secondary	ELISA	0.46 (0.34 to 0.59)	0.93 (0.82 to 0.98)		
Carroccio 2003f[87]	18	1	17	14	100.0	Calprest	Paediatric	Secondary	ELISA	0.51 (0.34 to 0.69)	0.93 (0.68 to 1.00)		
Carroccio 2003e[87]	13	3	17	37	100.0	Calprest	Adult	Secondary	ELISA	0.43 (0.25 to 0.63)	0.93 (0.80 to 0.98)		
Basumani 2012f[83]	13	14	8	84	150.0	Immundiagnostik	Adult	Secondary	ELISA	0.62 (0.38 to 0.82)	0.86 (0.77 to 0.92)		

FIGURE 9 Organic vs. IBS.

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Basumani 2012a[83]	20	48	1	50	8.0	Immundiagnostik	Adult	Secondary	ELISA	0.95 (0.76 to 1.00)	0.51 (0.41 to 0.61)		
Basumani 2012b[83]	20	46	1	52	25.0	Immundiagnostik	Adult	Secondary	ELISA	0.95 (0.76 to 1.00)	0.53 (0.43 to 0.63)		
Basumani 2012c[83]	19	39	2	59	50.0	Immundiagnostik	Adult	Secondary	ELISA	0.90 (0.70 to 0.99)	0.60 (0.50 to 0.70)		
Basumani 2012d[83]	15	25	6	73	75.0	Immundiagnostik	Adult	Secondary	ELISA	0.71 (0.48 to 0.89)	0.74 (0.65 to 0.83)		
Basumani 2012e[83]	14	18	7	80	100.0	Immundiagnostik	Adult	Secondary	ELISA	0.67 (0.43 to 0.85)	0.82 (0.73 to 0.89)		
Basumani 2012f[83]	13	14	8	84	150.0	Immundiagnostik	Adult	Secondary	ELISA	0.62 (0.38 to 0.82)	0.86 (0.77 to 0.92)		

FIGURE 10 Organic vs. IBS: Basumani (2012)[83] data only.

FIGURE 11 Organic vs. IBS, ROC curve: Basumani (2012)[83] data only.

Henderson et al.[30] also provide data that shows the different mix of non-IBD conditions in children. The non-IBD conditions included IBS (about one-third of cases), non-specific colitis, post-infectious enteropathy, cow's milk or wheat intolerance, pinworms and allergic enteropathy.

Ashorn et al.[89] included three serological markers, all of which reflect immune response to commensal intestinal bacteria:

- anti-*Saccharomyces cerevisiae* antibodies (ASCA)
- OmpW – antibodies against an outer membrane protein of *Bacteroides caccae*
- Antibodies against *I2* from *Pseudomonas fluorescens*.

The sensitivity of these was much poorer than that of calprotectin overall in IBD but higher in CD (67%) than in UC (14%).

Canani et al.[65] also examined the use of serum markers in children. They found sensitivities of 41% for CRP, 52% for ESR, 78% for ASCA/perinuclear anti-neutrophil cytoplasmic antibodies (pANCA), and 93% for calprotectin.

Fagerberg et al.[91] reported data on faecal occult blood, which was less useful than calprotectin (cut-off 50 µg/g). Sensitivity, specificity, and PPV and NPV for calprotectin were 95%, 93%, 95% and 93%, and for faecal occult blood, 58%, 91%, 92% and 56%, respectively. ESR and CRP had poor sensitivities of 41% and 36%, and NPVs of 52% and 50%. All children with IBD had calprotectin levels of > 50 µg/g. The 95% sensitivity arose because it was expressed in terms of inflammation, not IBD, and one child had a normal calprotectin level (15 µg/g) but non-specific proctitis. One child with no mucosal inflammation identified had a calprotectin level of 65 µg/g.

Fagerberg et al.[91] note that 60% of the children in their study had colonic inflammation, which they consider to be typical of paediatric groups being investigated – a contrast from adult groups with their much lower proportion of IBD, due to the commonness of IBS.

Shaoul et al.[92] (not included in our meta-analysis) provide a report on calprotectin levels in children (age range 8–17 years) with untreated CD (a case series with no non-IBD comparison group, hence an exclusion for our purposes). The title of their paper is 'limitations of FC' but this appears to be chosen for two reasons. First, they note a lack of correlation with the Paediatric Crohn's Disease Activity Index (PCDAI).

TABLE 17 Outline of studies comparing IBD vs. non-IBD

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Ashorn 2009[89]	55	Suspicion of IBD	Hospital in Finland	To examine the association of FC with serological markers in children and adolescents with IBD	Upper and lower endoscopies with biopsy	Not reported
Canani 2006[65]	45	Suspicion of IBD	Paediatric gastroenterology unit in Italy	To assess the diagnostic accuracy of FC and other tests independently and in combination; to assess value of FC as screening tool for IBD in children	Presence or absence of previously reported clinical, radiographic, endoscopic and histopathological criteria. Bowel wall ultrasonography within 24 hours of admission	Patients with symptoms or signs strongly suggestive of IBD, such as abdominal mass
Damms 2008[58]	84	Patients aged >18 years undergoing colonoscopy for GI disorders or for colorectal cancer screening – medical check-up	Gastroenterological departments of three hospitals and three outpatient gastroenterology clinics in Germany	To assess the diagnostic accuracy of the new calprotectin rapid test compared with ELISA in detecting colonoscopy proven IBD and malignancies	Colonoscopy; for colorectal cancer screening – medical check-up	Known extraintestinal inflammatory disease; NSAIDs; anticoagulants
Diamanti 2010[86]	197	Recurrent abdominal pain and altered bowel habit	Gastroenterology and Nutrition Unit of Hospital, Rome, Italy	To assess the diagnostic precision and value as a screening tool of FC compared with histology	Colonoscopy including intubation of terminal ileum with biopsy	Not reported
Fagerberg 2005[91]	36	Children with GI symptoms who were scheduled for colonoscopy to rule out IBD	Hospitals in Stockholm, and Vasteras (Sweden)	To determine if FC can be used as a diagnostic test of colonic inflammation to identify those children who require colonoscopy	Complete ileocolonoscopy with biopsy	Had no bacterial gastroenteritis detectable by faecal culture or serology and did not have any other chronic inflammatory disease
Henderson 2012[30]	190	Patients undergoing endoscopy	The paediatric gastroenterology department at the Royal Hospital for Sick Children in Edinburgh, UK	To describe the differences in FC levels between IBD types and non-IBD disease categories	IBD patients: standard clinical, histological and radiological findings Non-IBD (control) patients: upper and lower endoscopy	Insufficient stool sample; aged <1 or >18 years; >6 months' delay between sample and endoscopy; previous endoscopy; FC sample after endoscopy; known diagnoses of GI diseases

TABLE 17 Outline of studies comparing IBD vs. non-IBD (*continued*)

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Licata 2012[61]	346	Consecutive patients referred with chronic (≥4 week) non-bloody diarrhoea of unknown origin	Gastroenterology outpatient department at the University of Palermo, Italy	To assess the diagnostic performance of FC as a stool-screening biomarker for organic intestinal disease	Colonoscopy with biopsies	GI bleeding, known malignancies of bowel; colonic surgery; recent respiratory or UT infection; pregnancy, alcohol abuse, NSAIDs
Limburg 2000[78]	110	Patients referred for colonoscopy with chronic diarrhoea (≥4 weeks) of unknown origin or chronic colitis of unknown cause	The Mayo Clinic (Rochester, MN)	To assess and compare calprotectin and Hb as stool screening biomarkers for colorectal inflammation	Colonoscopy; biopsies taken when clinically indicated	Abnormalities on GI radiographs; GI bleeding; GI endoscopy within the preceding 2 weeks, epistaxis within the preceding 1 week, active menstruation, known colorectal neoplasia, familial adenomatous polyposis and hereditary non-polyposis colorectal cancer syndrome
Sidler 2008[82]	61	Children aged between 2 and 18 years referred for further investigation with GI symptoms suggestive of an OBD	Gastroenterology outpatient clinic at Sydney Children's Hospital, Sydney, Australia	To define the appropriate roles for faecal S100A12 and calprotectin in the initial investigations of children with suspected IBD	Upper GI endoscopy and complete ileocolonoscopy with biopsy	Previous diagnosis of OBD; infectious gastroenteritis; use of NSAIDs, antibiotics or corticosteroids preceding 2 weeks
Tomas 2007[66]	43	Patients referred for further investigation with GI symptoms	Paediatric gastroenterology unit of university hospital in Spain	To evaluate FC in paediatric patients with signs and symptoms suggestive of IBD	Clinical criteria, laboratory, image and endoscopic test results	Not reported
Van de Vijver 2012[80]	117	Children aged between 6 and 18 years referred for further investigation with abdominal symptoms and clinical suspicion of IBD	Paediatric outpatient clinics of six general hospitals and one tertiary care hospital in the northern region of the Netherlands	To determine a diagnostic strategy to minimise the no. of patients with negative endoscopy results without missing any cases of IBD	Some patients had endoscopy; others – stool tests for bacteria, ova and parasites, gastroscopy, different imaging and dietary assessment. Complete resolution after 6 months follow-up	Younger children (who have higher normal values of FC)

Hb, haemoglobin.

TABLE 18 Quality assessment of studies (QUADAS) comparing IBD vs. non-IBD

Quality criterion	Ashorn 2009[89]	Canani 2006[65]	Damms 2008[58]	Diamanti 2010[86]	Fagerberg 2005[91]	Henderson 2012[30]	Licata 2012[61]	Limburg 2000[78]	Sidler 2008[82]	Tomas 2007[66]	Van de Vijver 2012[80]
Spectrum	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes
Reference standard	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acceptable delay?	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Unclear	Yes
Whole sample verified?	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes	Yes	No
Same reference standard	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Newly diagnosed?	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Blinded reference testing?	Unclear	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Unclear	Unclear	Yes
Index results blinded?	Unclear	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Yes
Same clinical data	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intermediate results reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Withdrawals explained?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

TABLE 19 Results for studies comparing IBD vs. non-IBD

Study	Cut-off value (µg/g)	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)	Accuracy	Disease prevalence, % (95% CI)
Ashorn 2009[89]	≥100	0.96 (0.89 to 1.00)	0.67 (0.38 to 0.88)	9.75 (1.50 to 63.37)	0.13 (0.05 to 0.29)	0.89	80.00 (67.03 to 89.57)
Canani 2006[65]	95.3	0.93 (0.76 to 0.99)	0.89 (0.65 to 0.99)	8.33 (2.25 to 30.92)	0.08 (0.02 to 0.32)	0.91	92.59 (75.71 to 99.09)
Damms 2008[58] (ELISA)	50	0.79 (0.60 to 0.88)	1.00 (0.93 to 1.00)	4.71 (2.96 to 7.50)	0	0.83	21.43 (13.22 to 31.74)
Damms 2005[58] (POCT)	50	0.55 (0.36 to 0.74)	0.96 (0.87 to 1.00)	4.51 (2.70 to 7.54)	0.14 (0.04 to 0.51)	0.82	21.43 (13.22 to 31.74)
Diamanti 2010[86]	100	0.82 (0.75 to 0.88)	1.0 (0.93 to 1.0)	3.08 (2.24 to 4.22)	0	0.87	59.39 (52.18 to 66.31)
Diamanti 2010[86]	160	0.88 (0.81 to 0.93)	1.00 (0.94 to 1.00)	5.00 (3.23 to 7.75)	0	0.92	59.39 (52.18 to 66.31)
Fagerberg 2005[91]	50	0.96 (0.77 to 1.00)	0.93 (0.66 to 1.00)	13.36 (2.02 to 88.54)	0.05 (0.01 to 0.33)	0.94	61.11 (43.46 to 76.86)
Henderson 2012[30]	>50	0.62 (0.53 to 0.70)	0.96 (0.85 to 0.99)	1.8 (.15 to 2.1)	0.05 (0.01 to 0.20)	0.70	47.89 (40.61 to 55.25)
Henderson 2012[30]	>100	0.68 (0.59 to 0.76)	0.95 (0.86 to 0.99)	2.3 (1.8 to 3.0)	0.06 (0.02 to 0.17)	0.77	47.89 (40.61 to 55.25)
Henderson 2012[30]	>200	0.77 (0.67 to 0.84)	0.92 (0.84 to 0.97)	3.6 (2.5 to 5.0)	0.09 (0.04 to 0.20)	0.83	47.89 (40.61 to 55.25)
Henderson 2012[30]	>300	0.83 (0.74 to 0.90)	0.89 (0.81 to 0.95)	5.2 (3.3 to 8.0)	0.13 (0.07 to 0.24)	0.86	47.89 (40.61 to 55.25)
Henderson 2012[30]	>800	0.93 (0.84 to 0.98)	0.79 (0.71 to 0.86)	14.5 (6.1 to 34.4)	0.29 (0.21 to 0.41)	0.84	47.89 (40.61 to 55.25)
Licata 2012[61]	150	0.82 (0.74 to 0.88)	0.84 (0.78 to 0.88)	6.40 (4.35 to 9.44)	0.28 (0.21 to 0.37)	0.83	41.04 (35.81 to 46.43)
Limburg 2000[78]	100	0.63 (0.46 to 0.78)	0.93 (0.85 to 98)	4.79 (2.89 to 7.93)	0.21 (0.09 to 0.47)	0.83	26.36 (18.42 to 35.62)
Sidler 2008[82]	50	0.76 (0.60 to 0.88)	1.00 (0.83 to 1.00)	3.00 (1.81 to 4.98)	0	0.84	50.82 (37.70 to 63.86)
Sidler 2008[82]	100	0.97 (0.82 to 1.00)	0.91 (0.75 to 0.98)	27.10 (3.93 to 186.78)	0.10 (0.03 to 0.29)	0.93	50.82 (37.70 to 63.86)
Tomas 2007[66]	50	0.94 (0.70 to 1.00)	1.00 (0.74 to 1.00)	13.00 (1.98 to 85.46)	0	0.96	53.57 (33.87 to 72.49)
Tomas 2007[66]	100	0.93 (0.68 to 1.00)	0.92 (0.63 to 1.00)	12.13 (1.98 to 80.15)	0.07 (0.01 to 0.48)	0.93	53.57 (33.87 to 72.49)
Tomas 2007[66]	150	1.00 (0.75 to 1.00)	0.87 (0.60 to 0.99)		0.13 (0.04 to 0.48)	0.93	53.57 (33.87 to 72.49)
Tomas 2007[66]	200	1.00 (0.74 to 1.00)	0.81 (0.54 to 0.96)		0.20 (0.07 to 0.55)	0.89	53.57 (33.87 to 72.49)
Van de Vijver 2012[80]	50	0.68 (0.55 to 0.79)	1.00 (0.94 to 1.00)	3.8 (2.6 to 5.5)	0	0.83	35.9 (27.24 to 45.29)
Van de Vijver 2012[80] (excluding GI)	50	0.78 (0.64 to 0.88)	1.00 (0.93 to 1.00)	5.17 (3.11 to 8.59)	0	0.88	40.38 (30.87 to 50.46)

TABLE 20 Measures of diagnostic accuracy for increasing FC levels in Henderson *et al.* (2012)[30]

FC cut-off (µg/g)	Sensitivity	Specificity	NPV
> 50	0.98	0.44	0.96
> 100	0.97	0.50	0.95
> 200	0.93	0.74	0.92
> 300	0.89	0.83	0.83

However, this may reflect the limitations of clinical activity scores rather than of calprotectin. Second, in their group of 60 children with CD, three had normal calprotectin levels. Two of these had 'minimal' findings on endoscopy, and the other had moderate changes. Interestingly, two of the children had internal fistulae at diagnosis. They also report CRP and ESR results: 8 of 10 patients with normal CRP had raised calprotectin, as did 9 of 10 with normal ESR.

Sidler *et al.*[82] compared FC with faecal S100A12, CRP and ESR. S100A12 is a protein from the same S100 family as calprotectin (which is a complex of S100A8 and S100A9). S100A12 performed better than calprotectin because of specificity. Sensitivities were similar at 100% for calprotectin and 97% for S100A12, but specificities were 67% and 97%, respectively. NPVs were 97% for S100A12 and 100% for calprotectin (data from table – text says NPV for calprotectin was 95%). ESR had 74% sensitivity and CRP 81%. Serum S100A12 had sensitivity of only 22%. The 67% specificity for calprotectin is considerably lower than in other studies such as Fagerberg *et al.*[91] (93%).

All but one of the 31 patients with IBD had CD. The low specificity of calprotectin was due to raised levels in some children without IBD. These children had various conditions, including *Helicobacter pylori* infection, a duodenal ulcer and reflux oesophagitis, but the paper does not say which child had the raised calprotectin.

Bremner *et al.*[93] (not included in meta-analysis) report a study of calprotectin in children, but most were not newly diagnosed and so it was an exclusion for our purposes. However, they noted that some children without bowel inflammation had raised (> 50 µg/g) calprotectin. Three had functional constipation on laxative treatment, and one had normal findings but a family history of IBD. The latter raises the possibility that calprotectin may be raised before there is clinical evidence of IBD. As in adults, calprotectin is not raised in eosinophilic, lymphocytic or non-specific colitis.[94]

In *Figure 12*, specificity is rather more variable than sensitivity, and CIs also vary. The precision of both depends on patient numbers. For example, in studies having a higher proportion of patients with disease than without, estimates of sensitivity will be more precise than those of specificity. This is best illustrated by the Diamanti study.[86]

For the 50 µg/g threshold, *Figure 13* has the ROC curve, *Figure 14* the pooled forest plot, and *Figure 15* the nomogram and LRs. *Figures 16–18* provide the same for the threshold of 100 µg/g.

Table 21 shows DORs at different cut-offs for the IBD versus non-IBD comparison.

Summary inflammatory bowel disease versus non-inflammatory bowel disease

In these mostly paediatric studies, the overall results pooled for IBD versus IBS, show very high sensitivity (99%: 95% CI 95% to 100%) (see *Figure 13*), but moderate specificity (74%) at a cut-off of 50 µg/g. At a cut-off of 100 µg/g, sensitivity falls to 94% (95% CI 86% to 98%), but specificity improves to 82% (95% CI 67% to 91%).

Calprotectin is therefore a valuable test in children with suspected IBD, and will allow most with non-IBD conditions to avoid invasive investigations, in particular colonoscopy.

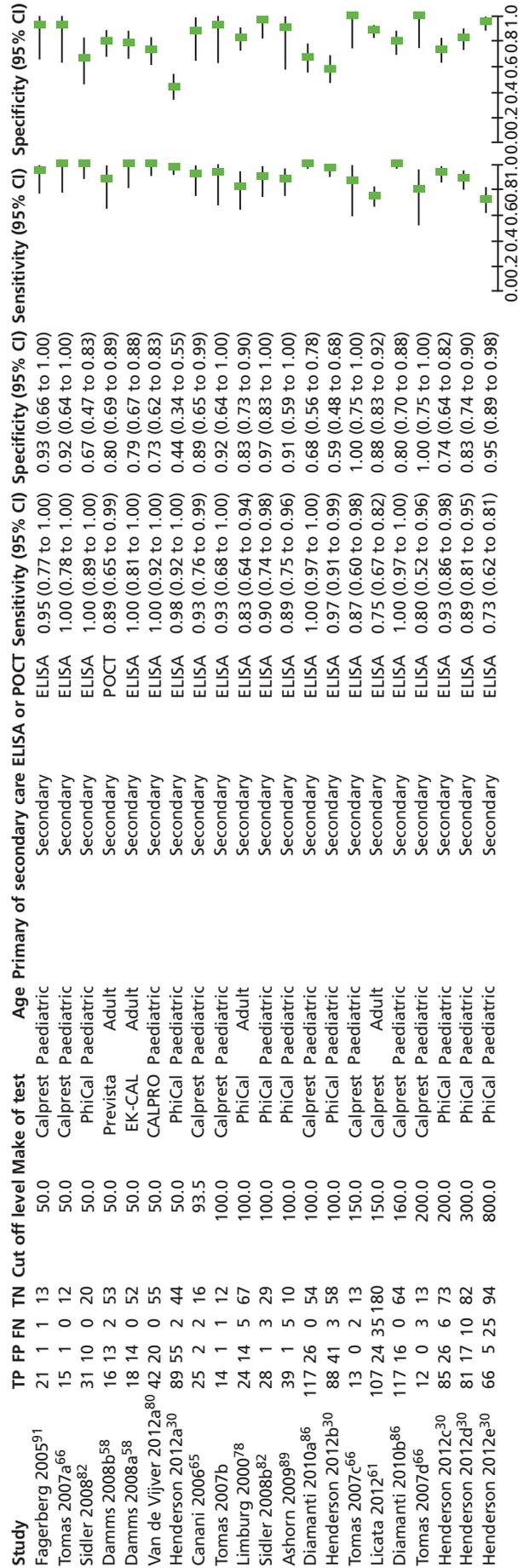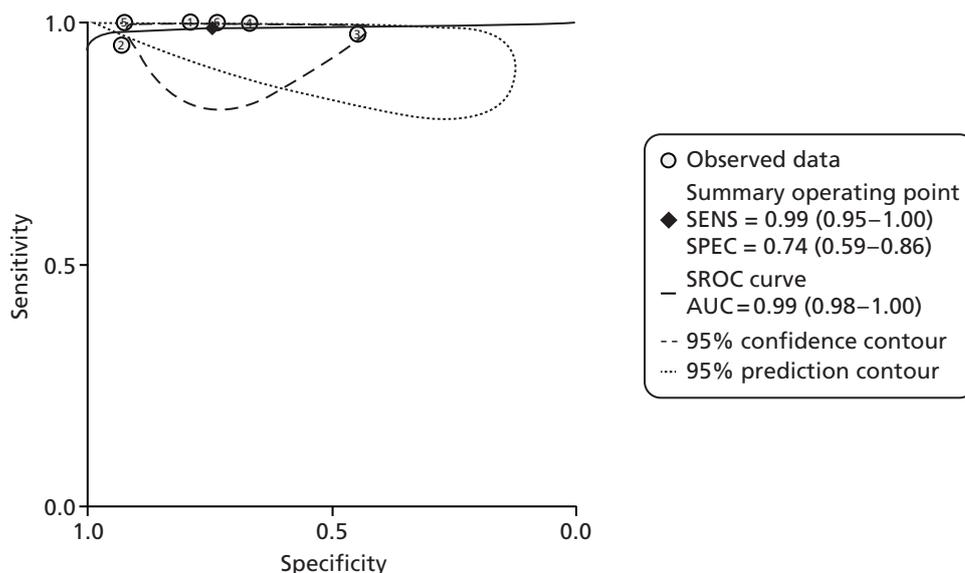| Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Fagerberg 2005[91]	21	1	1	13	50.0	Calprest	Paediatric	Secondary	ELISA	0.95 (0.77 to 1.00)	0.93 (0.66 to 1.00)		
Tomas 2007a[66]	15	1	0	12	50.0	Calprest	Paediatric	Secondary	ELISA	1.00 (0.78 to 1.00)	0.92 (0.64 to 1.00)		
Sidler 2008[82]	31	10	0	20	50.0	PhiCal	Paediatric	Secondary	ELISA	1.00 (0.89 to 1.00)	0.67 (0.47 to 0.83)		
Damms 2008b[58]	16	13	2	53	50.0	Prevista	Adult	Secondary	POCT	0.89 (0.65 to 0.99)	0.80 (0.69 to 0.89)		
Damms 2008a[58]	18	14	0	52	50.0	EK-CAL	Adult	Secondary	ELISA	1.00 (0.81 to 1.00)	0.79 (0.67 to 0.88)		
Van de Vijver 2012a[80]	42	20	0	55	50.0	CALPRO	Paediatric	Secondary	ELISA	1.00 (0.92 to 1.00)	0.73 (0.62 to 0.83)		
Henderson 2012a[30]	89	55	2	44	50.0	PhiCal	Paediatric	Secondary	ELISA	0.98 (0.92 to 1.00)	0.44 (0.34 to 0.55)		
Canani 2006[65]	25	2	2	16	93.5	Calprest	Paediatric	Secondary	ELISA	0.93 (0.76 to 0.99)	0.89 (0.65 to 0.99)		
Tomas 2007b	14	1	1	12	100.0	Calprest	Paediatric	Secondary	ELISA	0.93 (0.68 to 1.00)	0.92 (0.64 to 1.00)		
Limburg 2000[78]	24	14	5	67	100.0	PhiCal	Adult	Secondary	ELISA	0.83 (0.64 to 0.94)	0.83 (0.73 to 0.90)		
Sidler 2008b[82]	28	1	3	29	100.0	PhiCal	Paediatric	Secondary	ELISA	0.90 (0.74 to 0.98)	0.97 (0.83 to 1.00)		
Ashorn 2009[89]	39	1	5	10	100.0	PhiCal	Paediatric	Secondary	ELISA	0.89 (0.75 to 0.96)	0.91 (0.59 to 1.00)		
Diamanti 2010a[86]	117	26	0	54	100.0	Calprest	Paediatric	Secondary	ELISA	1.00 (0.97 to 1.00)	0.68 (0.56 to 0.78)		
Henderson 2012b[30]	88	41	3	58	100.0	PhiCal	Paediatric	Secondary	ELISA	0.97 (0.91 to 0.99)	0.59 (0.48 to 0.68)		
Tomas 2007c[66]	13	0	2	13	150.0	Calprest	Paediatric	Secondary	ELISA	0.87 (0.60 to 0.98)	1.00 (0.75 to 1.00)		
Licata 2012[61]	107	24	35	180	150.0	Calprest	Adult	Secondary	ELISA	0.75 (0.67 to 0.82)	0.88 (0.83 to 0.92)		
Diamanti 2010b[86]	117	16	0	64	160.0	Calprest	Paediatric	Secondary	ELISA	1.00 (0.97 to 1.00)	0.80 (0.70 to 0.88)		
Tomas 2007d[66]	12	0	3	13	200.0	Calprest	Paediatric	Secondary	ELISA	0.80 (0.52 to 0.96)	1.00 (0.75 to 1.00)		
Henderson 2012c[30]	85	26	6	73	200.0	PhiCal	Paediatric	Secondary	ELISA	0.93 (0.86 to 0.98)	0.74 (0.64 to 0.82)		
Henderson 2012d[30]	81	17	10	82	300.0	PhiCal	Paediatric	Secondary	ELISA	0.89 (0.81 to 0.95)	0.83 (0.74 to 0.90)		
Henderson 2012e[30]	66	5	25	94	800.0	PhiCal	Paediatric	Secondary	ELISA	0.73 (0.62 to 0.81)	0.95 (0.89 to 0.98)		

FIGURE 12 Inflammatory bowel disease vs. non-IBD.

FIGURE 13 Summary receiver operating characteristic curve for FC in the diagnosis of bowel diseases: IBD vs. non-IBD at a cut-off level of 50 μg/g.

Studies of calprotectin: organic versus non-organic bowel disease

Table 22 gives details of the studies, *Table 23* their QUADAS assessments and *Table 24* the results.

Figure 19 shows the dual forest plots. *Figure 20* gives the result for the Kok *et al.*[39] study only, to show the effect of excluding small adenomas, on the grounds that these are not associated with inflammation.

Table 25 shows the DORs.

The organic diseases that give rise to raised calprotectin include NSAID enteropathy, diverticular disease, polyps, and coeliac disease, but calprotectin can be normal in the presence of some of these, including polyps and diverticulosis.[44]

Ranges

It is worth noting that notwithstanding the generally good predictive value of calprotectin for differentiating IBD and IBS in adults, and IBD and non-IBD in children, the range of results can be wide, with some low levels in patients with IBD and raised levels in people with IBS.

Tables 26 and 27 give some examples of both full ranges and IQRs. In some studies, the ranges do not overlap, whereas in others they do. For example, in El-Badry *et al.*,[81] the value of FC in patients with IBD ranged between 98 and 637 μg/g, which does not overlap with the value of FC in patients with IBS (14–65 μg/g). In all other studies, the range of FC in patients with IBD overlapped with the range of FC in patients with IBS. In some studies, such as Li *et al.*[60] and Schroder *et al.*,[77] the range of FC level in patients with IBD was wide with the lowest value being 15 μg/g and the highest being 2574 μg/g.

The range of results in studies comparing IBD and non-IBD in children was similar to that found in studies comparing IBD and IBS in adults. In some studies (Canini *et al.*;[65] Diamanti *et al.*;[86] Sidler *et al.*[82]), the ranges overlapped, in others they did not. It should also be noted that in some patients with IBS, FC levels were high, considerably more than the manufacturer's cut-off levels (Canini *et al.*;[65] Diamanti *et al.*;[86] Sidler *et al.*[82]).

FIGURE 14 Forest plots of pooled sensitivity and specificity of FC in the diagnosis of bowel diseases: IBD vs. non-IBD at a cut-off level of 50 µg/g.

FIGURE 15 The use of the Fagan's nomogram (a straight line through the pre-test probability of 20% and the LR− of 0.20 yields a post-test probability of about 2%): IBD vs. non-IBD at a cut-off level of 50 μg/g.
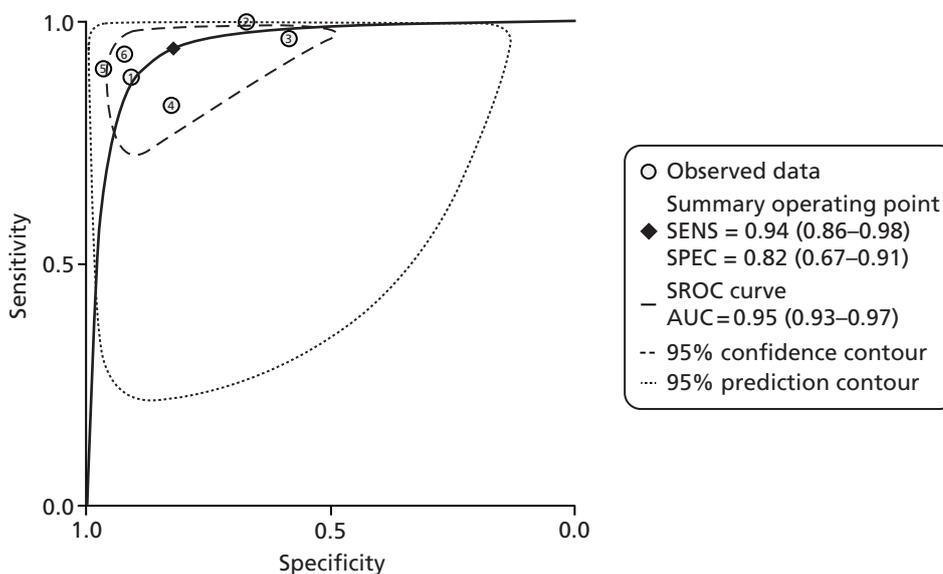
FIGURE 16 Summary receiver-operating characteristic (SROC) curve for FC in the diagnosis of bowel diseases: IBD vs. non-IBD at a cut-off level of 100 μg/g.
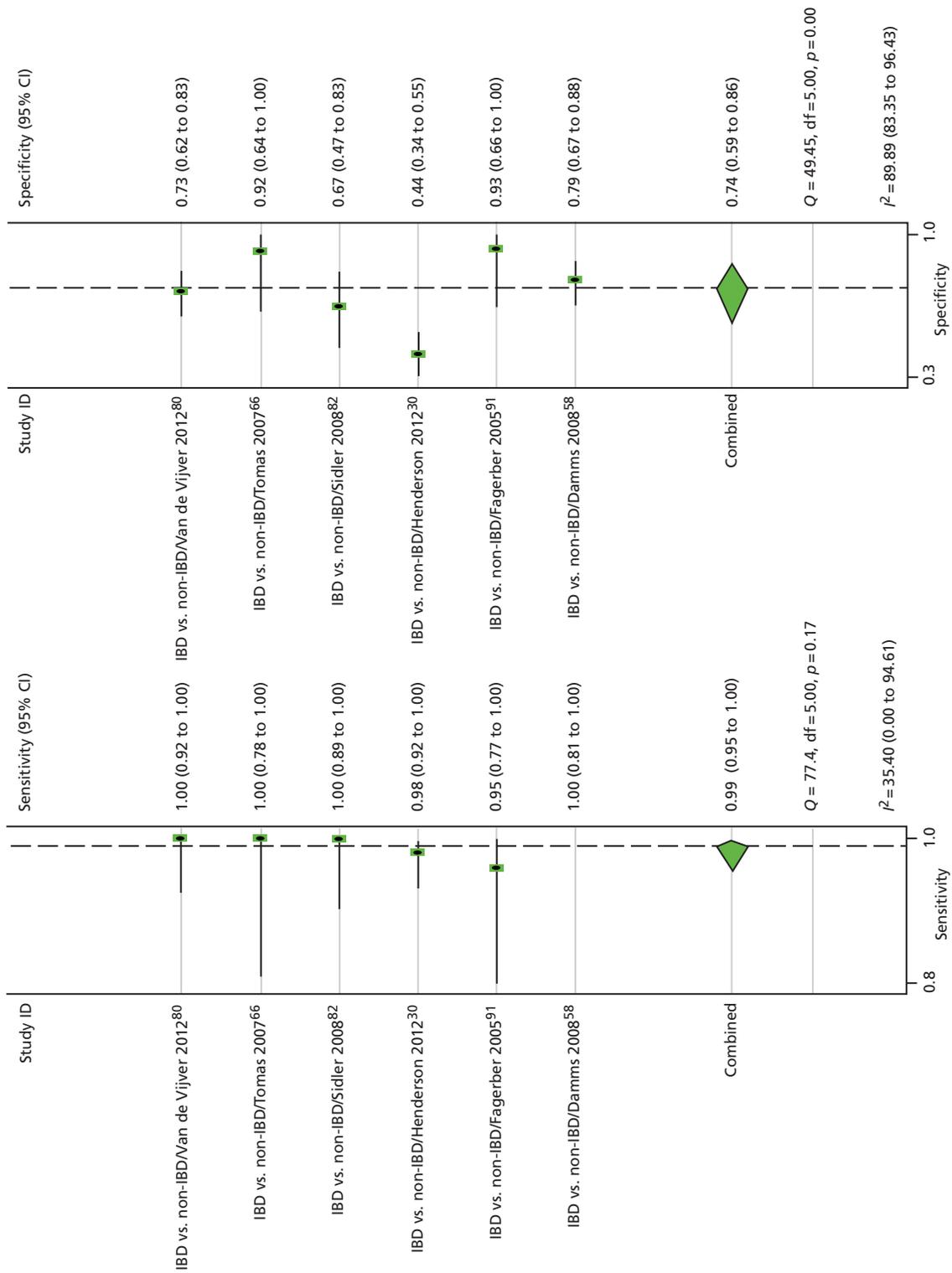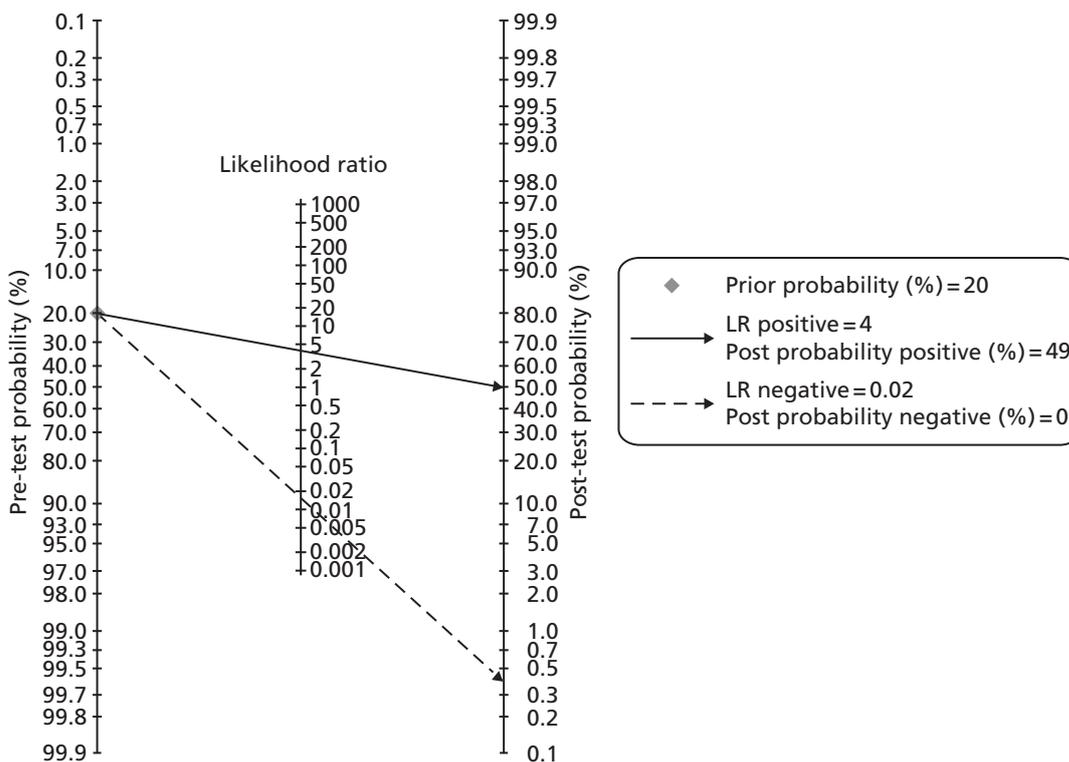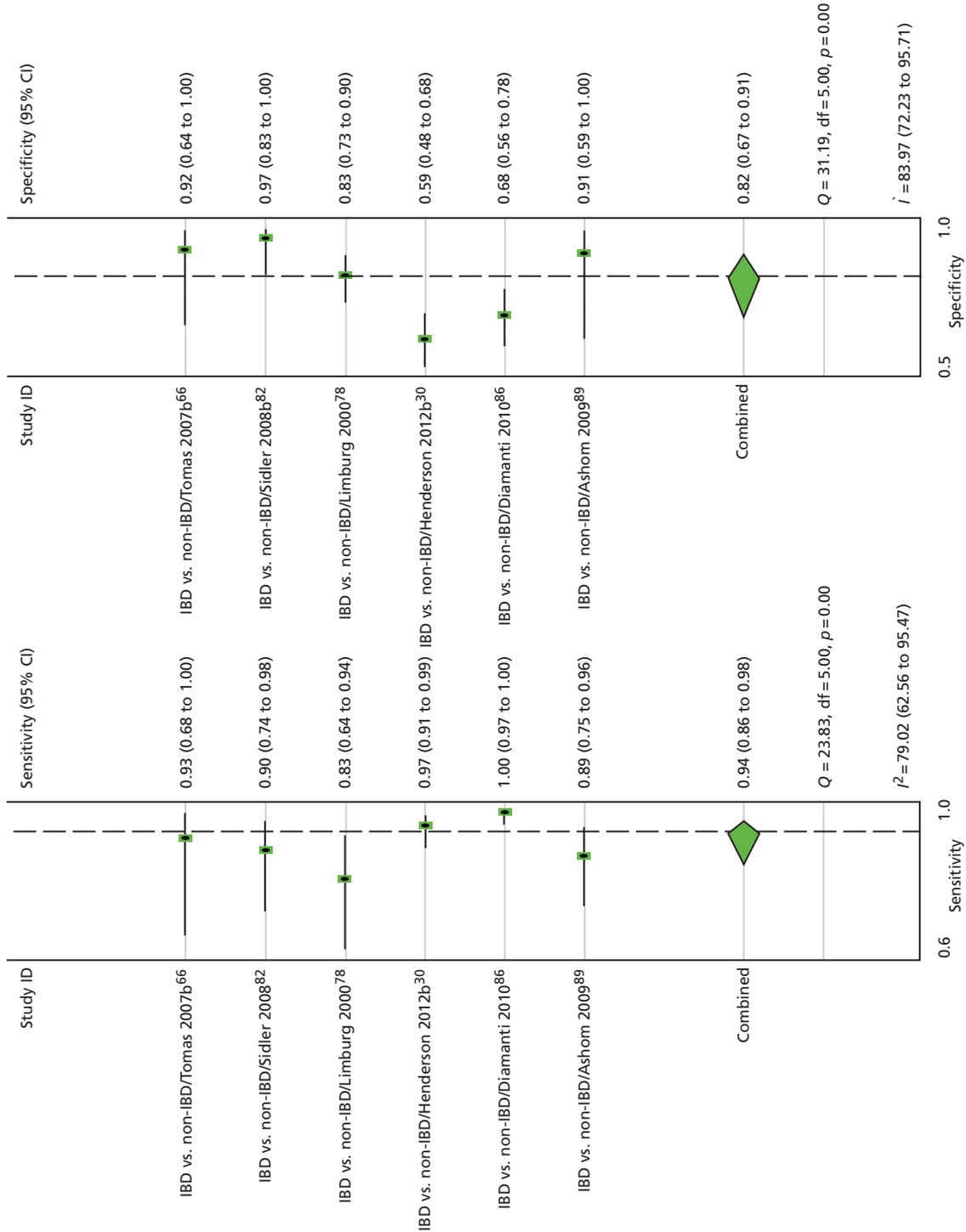
FIGURE 17 Forest plots of pooled sensitivity and specificity of FC in the diagnosis of bowel diseases: IBD vs. non-IBD at a cut-off level of 100 µg/g.
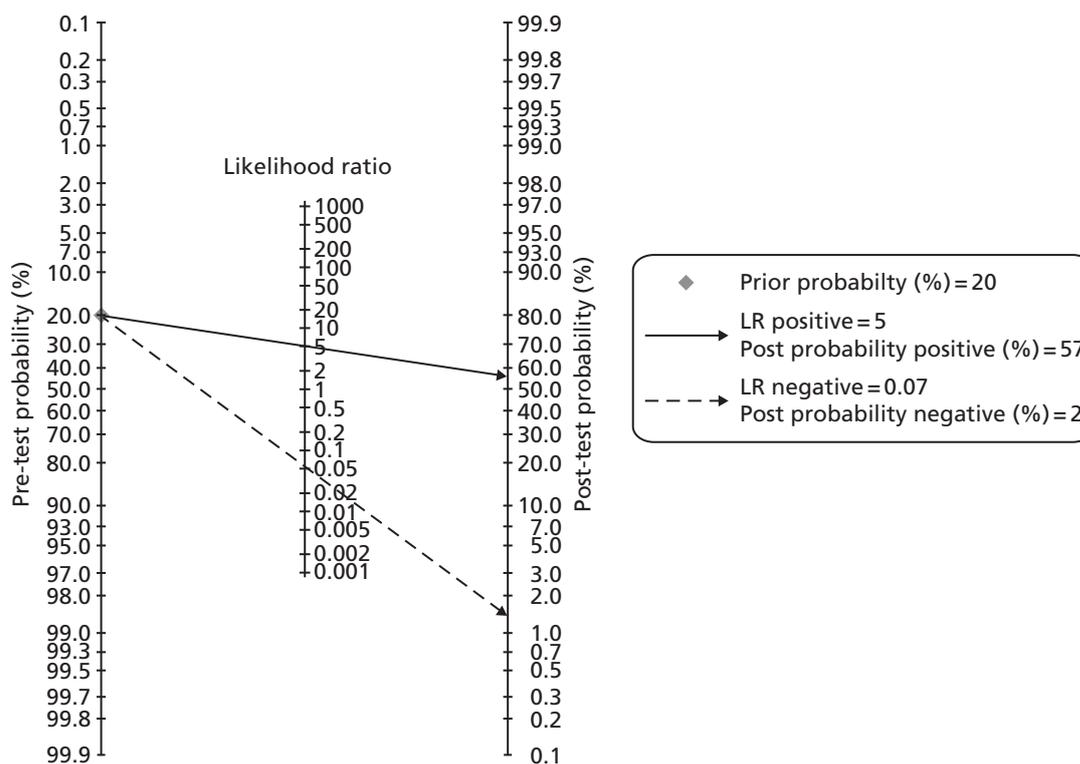
FIGURE 18 The use of the Fagan's nomogram (a straight line through the pre-test probability of 20% and the LR− of 0.20 yields a post-test probability of about 2%): IBD vs. non-IBD at a cut-off level of 100 µg/g.

TABLE 21 Diagnostic odds ratios: IBD vs. non-IBD studies

No. of patients	No. of studies	Cut-off level	DOR (95% CI)
531	6	50	246 (44 to 1376)
45	1	93.5	100 (10.2 to 1250.0)
656	6	100	79 (31 to 202)
389	2	150	5.1 (3.8 to 6.9)
197	1	160	114 (7.2 to 1804)
233	2	200	8.5 (4.5 to 15.8)
190	1	300	39.1 (15.8 to 99.9)
190	1	800	49.6 (17.2 to 169.8)

Choice of test

Given that the focus of this review is in the performance of calprotectin tests for distinguishing patients who need to be referred from those who do not, we need to assess the comparative performance of different tests at levels representative of patients in that situation. These will consist of patients with IBS, most with normal or low calprotectins; and patients with IBD, some of who will have lowish calprotectin levels (50–200 µg/g), and some of whom will have florid inflammation and high levels. Given the low to high range of results, it may be unsafe to extrapolate from studies that compare different tests in groups of patients with much higher levels, such as when calprotectin is used to identify relapse or to monitor treatment.

We therefore gave preference to studies in newly presenting patients.

TABLE 22 Outline of studies comparing OBD vs. non-OBD

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Burri 2013[85]	405	Patients undergoing endoscopy of the GI tract for abdominal discomfort	Department of Gastroenterology of the University Hospital Basel, Switzerland	To compare three different assays in their ability to identify patients with organic intestinal disease	Endoscopy, esophagogastroduodenoscopy and histology	Younger than 18 years
Dolwani 2004[48]	63	Consecutive patients undergoing small bowel BaFT examination	Gastroenterology outpatient clinic, University of Wales Hospital	To compare the utility of a single FC estimation to BaFT in exclusion of intestinal inflammation	Rigid sigmoidoscopy and stool cultures	Known malignancy, on NSAIDs, coeliac diseases, etc.
Garcia 2006[59]	190	Consecutive individuals who underwent colonoscopy for medical indications	Hospital Universitario Reina Sofia Córdoba, Spain	To assess the usefulness of FC to predict the presence of pathological colonoscopy	Colonoscopy, clinical criteria, endoscopic and histological findings	Severe cardiopulmonary disease, kidney or liver disease, coeliac disease, known malignancy
Kok 2012[39]	382	Patients consulting their GPs for persistent lower-abdominal complaints	Data from the CEDAR study in 170 general practices in two regions of the Netherlands	To quantify the diagnostic accuracy of three biomarker tests for the inclusion or exclusion of OBD	Endoscopy, and biopsies if required	<18 years old, previously diagnosed with OBD, or positive on triple faeces test
Lee 2013[90]	122	Patients presenting with chronic diarrhoea without a pre-existing diagnosis of IBD	Durham Dales, England, primary and secondary care	To determine whether the manufacturer's cut-off levels for referral are useful in diagnosis in patients with chronic diarrhoea	Not clear	Not reported
Manz 2012[94]	538	Patients undergoing endoscopy of the GI tract for abdominal discomfort	Department of Gastroenterology of the University Hospital Basel in Switzerland	To prospectively investigate the value of FC as a biological marker for the diagnosis of intestinal organic disease in symptomatic patients	Endoscopy and biopsies as decided by the endoscopist	Age <18 years

Study	No. of patients	Recruits	Setting	Aim	Reference test	Exclusions
Shitrit 2007[62]	69	Patients referred to the department of gastroenterology for colonoscopic examination of various indications, including screening	Department of Gastroenterology, Shaare Zedek Medical Center, Israel	To assess the predictive value of FC in organic colonic disease	Colonoscopy and histopathology	Intake of NSAIDs and/or antibiotics during the previous 3 months, concomitant serious illness
Tibble 2002[49]	602	Patients referred to a gastroenterology outpatient department by GPs	Gastroenterology outpatient department of a teaching hospital in South London	To determine if the use of FC and IP are useful in differentiating between patients with organic and nonorganic disease	One or more invasive diagnostic imaging procedures, appropriate to their symptoms	Previously diagnosis of IBD, colorectal carcinoma, and other serious diseases
Tomas 2007[66]	43	Referred by GPs; all patients had clinical symptoms suggestive of organic intestinal disease or IBS that had not responded to therapy	Patients referred to the paediatric GI unit of a hospital in Mallorca, Spain	To evaluate FC in paediatric patients with signs and symptoms suggestive of IBD	Clinical criteria, laboratory, image and endoscopic test results	Not reported
Turvill 2012[46]	630	New patient referrals from primary care, aged 16–60 years, with intestinal symptoms	The Department of Gastroenterology, York Hospital, York, UK	To determine the NPV of a normal FC in excluding organic intestinal disease in patients with intestinal symptoms	Colonoscopy, supportive histology, barium meal, Computed tomographic enterography and capsule endoscopy	Patients with fast-track colorectal symptoms

BaFT, barium follow-through; IP, intestinal permeability.

TABLE 23 Quality assessment of studies (QUADAS) comparing OBD vs. non-OBD

Quality criterion	Burri 2013[85]	Dolwani 2004[48]	Garcia 2006[59]	Kok 2012[39]	Lee 2013[90]	Manz 2012[84]	Shitrit 2007[62]	Tibble 2002[49]	Tomas 2007[66]	Turvill 2012[46]
Spectrum?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes
Reference standard?	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Acceptable delay?	Unclear	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Unclear	Unclear
Whole sample verified?	Unclear	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Same reference standard	No	Yes	Yes	Yes	Unclear	No	Yes	Yes	Unclear	Yes
Newly diagnosed?	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Yes	Yes	Yes
Blinded reference testing?	Yes	Yes	Unclear	Yes	No	Yes	Yes	Unclear	Unclear	Unclear
Index results blinded?	Yes	Yes	Unclear	Yes	Yes	No	Unclear	Unclear	Unclear	Unclear
Same clinical data	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intermediate results reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear
Withdrawals explained?	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Unclear

TABLE 24 Table of results comparing OBD vs. non-OBD

Study	Cut-off value (µg/g)	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)	Accuracy	Disease prevalence, % (95% CI)
Burri 2013[85] EK-CAL	50	0.90 (0.83 to 0.94)	0.89 (0.85 to 0.92)	15.78 (9.23 to 27.00)	0.23 (0.17 to 0.31)	0.89	35.31 (30.65 to 40.18)
Burri 2013[85] PhiCal	22.4	0.84 (0.75 to 0.91)	0.84 (0.78 to 0.87)	9.67 (5.93 to 15.77)	0.36 (0.28 to 0.46)	0.84	35.18 (30.25 to 40.35)
Burri 2013[85] PhiCal	50	0.93 (0.84 to 0.98)	0.76 (0.71 to 0.81)	25.33 (9.40 to 68.30)	0.58 (0.49 to 0.67)	0.79	35.18 (30.25 to 40.35)
Dolwani 2004[48]	60	0.60 (0.39 to 0.79)	1.00 (0.91 to 1.00)	4.80 (2.77 to 8.33)	0	0.84	23.81 (13.98 to 36.21)
Garcia 2006[59]	217	0.74 (0.63 to 0.83)	0.90 (0.82 to 0.95)	4.52 (3.06 to 6.66)	0.19 (0.11 to 0.32)	0.83	38.42 (31.47 to 45.74)
Kok 2012,[39] including only large adenomas as OBD, POCT (Quantum Blue)	50	0.24 (0.19 to 0.31)	0.92 (0.87 to 0.95)	1.7 (1.4 to 2.0)	0.5 (0.3 to 0.7)	0.58	16.23 (12.68 to 20.32)
Kok 2012,[39] large adenomas only, ELISA	50	0.23 (0.18 to 0.28)	0.93 (0.88 to 0.96)	1.5 (1.3 to 1.7)	0.4 (0.2 to 0.7)	0.51	16.23 (12.68 to 20.32)
Kok 2012,[39] all adenomas, POC	50	0.32 (0.26 to 0.39)	0.81 (0.74 to 0.86)	1.4 (1.1 to 1.7)	0.7 (0.5 to 0.9)	0.56	25.92 (21.59 to 30.62)
Kok 2012,[39] all adenomas, ELISA	50	0.33 (0.27 to 0.39)	0.84 (0.77 to 0.89)	1.4 (1.2 to 1.6)	0.6 (0.4 to 0.8)	0.54	25.92 (21.59 to 30.62)
Lee 2013[90]	60	0.37 (0.24 to 0.52)	0.94 (0.86 to 0.98)	2.56 (1.82 to 3.59)	0.26 (0.1 to 0.63)	0.70	18.9 (12.3 to 26.9)
Manz 2012,[84] overall	50	0.87 (0.81 to 0.92)	0.84 (0.80 to 0.88)	10.36 (6.93 to 15.50)	0.29 (0.23 to 0.36)	0.85	39.41 (35.25 to 43.68)
Manz 2012,[84] overall	10	0.55 (0.50 to 0.60)	0.93 (0.88 to 0.96)	1.88 (1.68 to 2.10)	0.12 (0.07 to 0.21)	0.67	39.41 (35.25 to 43.68)

continued

TABLE 24 Table of results comparing OBD vs. non-OBD (*continued*)

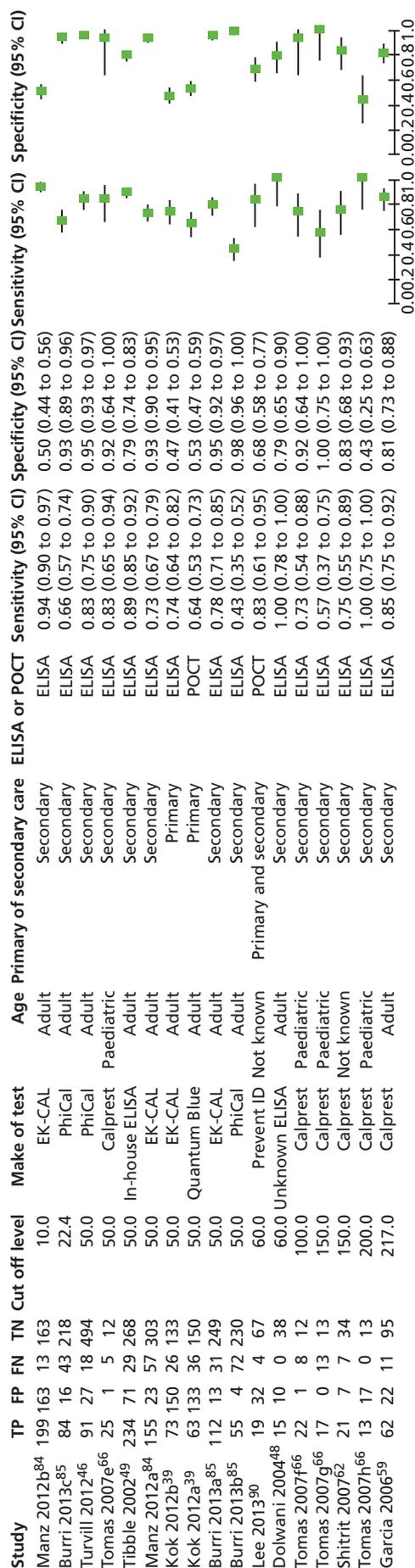| Study	Cut-off value (μg/g)	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)	Accuracy	Disease prevalence, % (95% CI)
Manz 2012,[84] upper GI	48	0.68	0.75	3.23	0.5		
Manz 2012,[84] Lower GI	50	0.93	0.82	10.6	0.17		
Shitrit 2007[62]	150	0.75 (0.55 to 0.89)	0.83 (0.68 to 0.93)	4.39 (2.16 to 8.91)	0.30 (0.16 to 0.58)	0.80	40.58 (28.91 to 53.08)
Tibble 2002[49]	10 mg/l = 50 μg/g	0.77 (0.72 to 0.81)	0.90 (0.86 to 0.93)	4.25 (3.44 to 5.25)	0.14 (0.10 to 0.20)	0.83	43.69 (39.68 to 47.76)
Tomas 2007[66]	50	0.96 (0.80 to 1.00)	0.71 (0.44 to 0.90)	10.83 (1.64 to 71.7)	0.18 (0.08 to 0.41)	0.86	69.77 (53.90 to 82.80)
Tomas 2007[66]	100	0.96 (0.78 to 1.00)	0.60 (0.36 to 0.81)	9.53 (1.43 to 63.45)	0.29 (0.16 to 0.53)	0.79	69.77 (53.90 to 82.80)
Tomas 2007[66]	150	1.00 (0.80 to 1.00)	0.50 (0.30 to 0.70)		0.43 (0.29 to 0.65)	0.70	69.77 (53.90 to 82.80)
Tomas 2007[66]	200	1.00 (0.75 to 1.00)	0.43 (0.25 to 0.63)		0.57 (0.41 to 0.77)	0.60	69.77 (53.90 to 82.80)
Turvill 2012[46]	< 50	0.77 (0.68 to 0.84)	0.96 (0.95 to 0.98)	16.11 (11.05 to 23.48)	0.17 (0.11 to 0.27)	0.93	17.30 (14.43 to 20.49)
Turvill 2012[46]	< 60	0.81	0.96				
Turvill 2012[46]	< 75	0.86	0.93				
Turvill 2012[46]	< 100	0.91	0.91				

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)
Manz 2012b[84]	199	163	13	163	10.0	EK-CAL	Adult	Secondary	ELISA	0.94 (0.90 to 0.97)	0.50 (0.44 to 0.56)
Burri 2013c[85]	84	16	43	218	22.4	PhiCal	Adult	Secondary	ELISA	0.66 (0.57 to 0.74)	0.93 (0.89 to 0.96)
Turvill 2012[46]	91	27	18	494	50.0	PhiCal	Adult	Secondary	ELISA	0.83 (0.75 to 0.90)	0.95 (0.93 to 0.97)
Tomas 2007e[66]	25	1	5	12	50.0	Calprest	Paediatric	Secondary	ELISA	0.83 (0.65 to 0.94)	0.92 (0.64 to 1.00)
Tibble 2002[49]	234	71	29	268	50.0	In-house ELISA	Adult	Secondary	ELISA	0.89 (0.85 to 0.92)	0.79 (0.74 to 0.83)
Manz 2012a[84]	155	23	57	303	50.0	EK-CAL	Adult	Secondary	ELISA	0.73 (0.67 to 0.79)	0.93 (0.90 to 0.95)
Kok 2012b[39]	73	150	26	133	50.0	EK-CAL	Adult	Primary	ELISA	0.74 (0.64 to 0.82)	0.47 (0.41 to 0.53)
Kok 2012a[39]	63	133	36	150	50.0	Quantum Blue	Adult	Primary	POCT	0.64 (0.53 to 0.73)	0.53 (0.47 to 0.59)
Burri 2013a[85]	112	13	31	249	50.0	EK-CAL	Adult	Secondary	ELISA	0.78 (0.71 to 0.85)	0.95 (0.92 to 0.97)
Burri 2013b[85]	55	4	72	230	50.0	PhiCal	Adult	Secondary	ELISA	0.43 (0.35 to 0.52)	0.98 (0.96 to 1.00)
Lee 2013[90]	19	32	4	67	60.0	Prevent ID	Not known	Primary and secondary	POCT	0.83 (0.61 to 0.95)	0.68 (0.58 to 0.77)
Dolwani 2004[48]	15	10	0	38	60.0	Unknown ELISA	Adult	Secondary	ELISA	1.00 (0.78 to 1.00)	0.79 (0.65 to 0.90)
Tomas 2007f[66]	22	1	8	12	100.0	Calprest	Paediatric	Secondary	ELISA	0.73 (0.54 to 0.88)	0.92 (0.64 to 1.00)
Tomas 2007g[66]	17	0	13	13	150.0	Calprest	Paediatric	Secondary	ELISA	0.57 (0.37 to 0.75)	1.00 (0.75 to 1.00)
Shitrit 2007[62]	21	7	7	34	150.0	Calprest	Not known	Secondary	ELISA	0.75 (0.55 to 0.89)	0.83 (0.68 to 0.93)
Tomas 2007h[66]	13	17	0	13	200.0	Calprest	Paediatric	Secondary	ELISA	1.00 (0.75 to 1.00)	0.43 (0.25 to 0.63)
Garcia 2006[59]	62	22	11	95	217.0	Calprest	Adult	Secondary	ELISA	0.85 (0.75 to 0.92)	0.81 (0.73 to 0.88)

FIGURE 19 Organic bowel disease vs. non-OBD.

Study	TP	FP	FN	TN	Cut off level	Make of test	Age	Primary of secondary care	ELISA or POCT	Sensitivity (95% CI)	Specificity (95% CI)
Kok 2012c[39]	47	147	15	173	50.0	Quantum Blue	Adult	Primary	ELISA	0.76 (0.63 to 0.86)	0.54 (0.48 to 0.60)
Kok 2012d[39]	51	176	11	144	50.0	EK-CAL	Adult	Primary	ELISA	0.82 (0.70 to 0.91)	0.45 (0.39 to 0.51)

FIGURE 20 Organic bowel disease (excluding adenomas of ≤ 1 cm) vs. non-OBD.

TABLE 25 Organic bowel disease vs. non-OBD

No. of patients	Cut-off level	No. of studies	DOR (95% CI)
538	10	1	15.3 (8.3 to 30.3)
405	22.4	1	26.6 (13.7 to 53.0)
3005	50	7	33 (13 to 81)
638	60	1	46.5 (2.9 to 743.6)
43	100	1	33 (3.5 to 1472.1)
112	150	2	2.8 (1.9 to 4.0)
43	200	1	12.2 (0.8 to 190.9)
190	217	1	24.3 (10.4 to 58.8)

TABLE 26 Adults: IBD vs. IBS

	Range		IQR	
Study	IBD	IBS	IBD	IBS
Carroccio 2003[87]	180 to 400	10 to 210		
El-Badry 2010[81]	98 to 637	14 to 65		
Kok 2012,[39] ELISA			55 to 1200	21 to 99
Kok 2012,[39] Quantum Blue			64 to 300	30 to 69
Li 2006[60]	23 to 2574	1 to 73	120 to 1118	6 to 27
Schroder 2007[77]	15 to 2553	0 to 24		

TABLE 27 Children: IBD vs. non-IBD

	Range		IQR	
Study	IBD	Non-IBD	IBD	Non-IBD
Ashorn 2009[89]	90 to 2250 (CD)	0 to 90		
	105 to 900 (IC)			
	5 to 2600 (UC)			
Canani 2006[65]	150 to 800 (CD)	0 to 160		
	200 to 1100 (UC)			
Diamanti 2010[86]	162 to 9500	15 to 400		
Fagerberg 2005[91]	213 to 440	7 to 28		
Sidler 2008[82]	52 to 12000	19 to 201		
Tomas 2007[66]			322 to 2967	36 to 193

Studies comparing faecal calprotectin tests: newly presenting patients

Four studies (Damms and Bischoff;[58] Kok et al.;[39] Otten et al.;[73] Burri et al.[85]) reported studies in which more than one type of commercial test kit was used, as shown in *Table 28*. Burri et al.[85] compared two different ELISAs, and the other three studies used a rapid test and an ELISA in the same patients. All of these studies were in adult patients.

Otten 2008

Otten et al.[73] tested the correlation between the point-of-care test (PreventID CalDetect cut-off of ≥ 15 µg/g compared with an ELISA (PhiCal) cut-off of > 50 µg/g on 114 samples for distinguishing between IBD and IBS. The correlation between the two tests gave a Cohen's kappa of 0.69.

CalDetect had a sensitivity of 1.00 (95% CI 0.85 to 1.00) and a specificity of 0.95 (0.88 to 0.98) compared with the PhiCal, which gave sensitivity of 0.96 (0.78 to 1.00) and specificity of 0.87 (0.78 to 0.93).

TABLE 28 Comparison of FC tests

Study ID	No. of samples compared	Diseases being diagnosed	Test 1	Test 2	Measurements comparing tests
Burri 2012[85]	361	OBD vs. non-OBD	PhiCal ELISA > 22.4 µg/g: sensitivity = 0.66 (0.57 to 0.74), specificity = 0.93 (0.89 to 0.96), AUC = 0.842	EK-CAL, Bühlman ELISA > 51 µg/g: sensitivity = 0.78 (0.71 to 0.85), specificity = 0.95 (0.92 to 0.97), AUC = 0.918	Correlation rho = 0.702, mean (SD) of difference in measurements = 30.9 (198.0) µg/g, difference in AUC, $p < 0.001$
Kok 2012[39]	382	OBD (including all adenomas as OBD) vs. non-OBD	Quantum Blue POCT 50 µg/g: sensitivity = 0.64 (0.54 to 0.72), specificity = 0.53 (0.48 to 0.59), AUC = 0.66 (0.60 to 0.72)	EK-CAL, Bühlman ELISA 50 µg/g: sensitivity = 0.74 (0.64 to 0.82), specificity = 0.47 (0.41–0.53), AUC = 0.65 (0.59 to 0.72)	ICC = 0.88 (0.85 to 0.90) [kappa = 0.66 (0.59 to 0.73)]
Kok 2012[39]	382	OBD (including advanced > 1 cm adenomas as OBD) vs. non-OBD	Quantum Blue POCT 50 µg/g: sensitivity = 0.76 (0.64 to 0.85), specificity = 0.54 (0.48 to 0.59), AUC = 0.75 (0.67 to 0.72)	EK-CAL, Bühlman ELISA 50 µg/g: sensitivity = 0.82 (0.70 to 0.91), specificity = 0.45 (0.39 to 0.51), AUC = 0.73 (0.66 to 0.81)	
Otten 2008[73]	114	IBD vs. IBS	PreventID CalDetect POCT 15 µg/g: sensitivity = 0.96 (0.78 to 1.00), specificity = 0.95 (0.88 to 0.98)	PhiCal ELISA 50 µg/g: sensitivity = 0.96 (0.78 to 1.00), specificity = 0.87 (0.78 to 0.93)	Cohen's kappa = 0.69
Damms 2008[58]	140	IBD vs. non-IBD	Prevista Rapid POCT: sensitivity = 0.89 (0.65 to 0.98) specificity = 0.80 (0.69 to 0.89), AUC = 0.896	EK-CAL Bühlman ELISA: sensitivity = 1.00 (0.82 to 1.00), specificity = 0.79 (0.67 to 0.98), AUC = 0.955	Correlation of line intensity of rapid with ELISA test value r = 0.862

Therefore, the rapid test at 15 μg/g was more sensitive and specific than the ELISA at 50 μg/g but neither difference was statistically significant.

Damms 2008

Damms and Bischoff[58] compared the Prevista POC and Bühlmann ELISA tests, both at a 50 μg/g cut-off, for detecting IBD, in 140 patients. The Bühlmann ELISA kit gave a sensitivity of 1.00 (0.82 to 1.00), specificity = 0.79 (0.67 to 0.88), area under the ROC (AUC) curve = 0.955. The Prevista rapid test sensitivity was 0.89 (0.65 to 0.98), and specificity was 0.80 (0.69 to 0.89), AUC = 0.896. So there were no significant differences in performance.

Kok 2012

Kok et al.[39] compared the Quantum Blue point-of-care test with the EK-CAL ELISA in OBD (which included all adenomas as OBD) versus non-OBD in 382 primary care patients, both at 50 μg/g cut-off levels. The agreement between the calprotectin POC and ELISA test was good [intraclass correlation coefficient (ICC) 0.88 (0.85–0.90), kappa = 0.66 (0.59–0.73)].

Note that the mean age was 60 years, so there was a higher prevalence of neoplasia, both benign and malignant, than would be expected in the age groups at which IBS was being separated from IBD. Of those investigated, 26% had OBD, of which 19% had colorectal cancer, and 54% had adenomas. Only 19% had IBD, 7% had UC and 2% had CD.

The point-of-care test had an AUC = 0.66 (95% CI 0.60 to 0.72), similar to the ELISA AUC = 0.65 (95% CI 0.59 to 0.72). The sensitivity of the point-of-care test was lower than the ELISA, with values of 0.64 (0.54 to 0.72) and 0.74 (0.64 to 0.82), respectively, but POCT had a higher specificity 0.53 (0.48 to 0.59) vs. 0.47 (0.41 to 0.53). As with the previous studies, these results overlap.

When small (≤ 1 cm) adenomas were excluded from the OBD category, the performance of both tests improved, with increased AUCs, specificity and sensitivity. The AUCs are still similar between POC and ELISA at 0.75 (0.67 to 0.72) and 0.73 (0.66 to 0.81), respectively. The point-of-care test still has a lower sensitivity = 0.76 (0.64 to 0.85) than the ELISA sensitivity = 0.82 (0.70 to 0.91), but a higher specificity = 0.54 (0.48 to 0.59) vs. specificity = 0.46 (0.41 to 0.51), respectively.

Of the 19 patients with IBD, the point-of-care test identified 15 and the ELISA 16, both at 50 μg/g.

Burri 2013

Burri et al.[85] performed a post-hoc analysis of a prospective study (Manz et al.,[84] also included in this study) to compare two ELISA tests. The EK-CAL is monoclonal, and the PhiCAL is polyclonal.

The cut-off values used were > 51 μg/g for EK-CAL (Bühlman ELISA) and > 22.4 μg/g for PhiCal. These were optimal cut-off values calculated from ROC analysis. The manufacturers' cut-offs were both 50 μg/g.

Calprotectin concentrations measured by EK-CAL correlated better with PhiCal (Spearman's rho = 0.702, $p < 0.001$) than with IBD-Scan (for lactoferrin) (rho = 0.592, $p < 0.001$). The mean (standard deviation, SD) of the difference between the measurements of FC using EK-CAL and PhiCal was 30.9 (198.0) μg/g.

The AUC (0.918) for EK-CAL was significantly better than for PhiCal (AUC = 0.842, $p < 0.001$) (from text – Figure 3 has slightly different AUCs).

EK-CAL ELISA at a cut-off of > 51 μg/g had a sensitivity of 0.78 (0.71 to 0.85) and specificity of 0.95 (0.92 to 0.97) compared with PhiCal at a cut-off of > 22.4 μg/g [sensitivity = 0.66 (0.57 to 0.74), specificity = 0.93 (0.89 to 0.96)].

So the monoclonal EK-CAL performed slightly better than the PhiCAL.

Hence in these studies in newly presenting patients, there is little difference in performance between the point-of-care tests and the ELISA tests.

Studies comparing faecal calprotectin tests in studies not in newly presenting patients

These are less useful for our purposes because they may reflect comparative reliability at much higher FC levels in active IBD or, conversely, lower levels in those in remission.

Caveat Many of the comparative studies are not yet published in full so details are often sparse. Missing details include those relating to sponsorship.

Studies comparing point-of-care testing and enzyme-linked immunosorbent assay tests

Kolho 2012

Kolho *et al.*[72] compares the Quantum Blue (Bühlmann) with ELISA, presumably EK-CAL, in 132 (134?) stool samples from 56 paediatric patients with CD, median age 13 years, range 1–18 (*Table 29*). Faecal samples were obtained from 10 European paediatric gastroenterology units, from patients taking part in European Growth, Relapse and Outcomes With Therapy (GROWTH) CD study.

Thirty of the faecal samples were obtained at the time of diagnosis, and the others, 8–72 weeks after starting treatment. The same stool samples were used for both tests.

Median FC value was significantly higher using the QB [317 mg/g (IQR 81–830; range 0–1862) versus 172 mg/g (IQR 50–840; range < 30–1656, respectively, $p = 0.001$) compared with the ELISA assay. (Note the high levels, a reminder of the need for caution when extrapolating from the results of studies not in newly presenting patients.)

The correlation between tests was better in values of < 300 µg/g, when dilutions were not required, which is the range most relevant to this review. Correlation between tests was high: Spearman's rho = 0.94, $p < 0.001$. There was more scatter at higher levels (from figure 1), but this was seen mainly in values of > 600 µg/g.

Interestingly, the authors comment that there was no difference in relative test performance between samples at first presentation and from follow-up (p. 437) but note that all patients had CD.

TABLE 29 Comparison of POCT and ELISA tests in Kolho *et al.*[72]

Test	Median FC levels (µg/g)	IQR	Range	Correlation – Spearman's rho	ICC analysis	Agreement between tests at 100 µg/g cut-off	Agreement between tests at 150 µg/g cut-off
ELISA	172	50–840	< 30 to 1656	0.94 $p < 0.001$	0.97 (95% CI 0.95 to 0.98)	87%, kappa = 0.87 (95% CI 0.60 to 0.84)	87%, kappa = 0.87
Quantum Blue	317	81–830	0 to 1862				

Wassell 2012

Wassell *et al.*[95] compared Quantum Blue POCT against PhiCal ELISA in 47 samples sent to the laboratory for 'routine calprotectin analysis'. They tested three extractions of the same stool in three patients and found considerable variation: results varied from – 31.3% to 31.5%.

Both manufacturers recommend a < 50 μg/g cut-off as upper limit of normal. With that cut-off, 4 of the 47 patients had results that fell on different sides of the cut-off. Two patients were positive by ELISA but negative by POCT and two were negative by ELISA but positive by POCT.

The authors, from Bristol, concluded that Quantum Blue was suitable for excluding IBD. They suggested that the POCT could be used in GI clinics to give immediate results, or in smaller laboratories that do not have sufficient throughput to justify an ELISA system.

Dolci 2012

Dolci and Panteghini[96] (published as a letter to the editor) compared the Quantum Blue point-of-care test with the Calprest (Eurospital) ELISA assay in stool specimens from 67 consecutive patients with suspected IBD, and found a 92.5% (95% CI 83% to 98%) agreement (*Table 30*). POCT was done on fresh samples. Samples for the ELISA test were frozen, thawed and tested within 2 weeks of collection.

Note that the cut-off used for the Quantum Blue POCT was much higher than that of the established ELISA method. Five patients showed discrepant results, four being positive only with ELISA (two borderline results, 94 and 98 μg/g stool) and one positive only with POCT.

Coorevits 2012

Coorevits *et al.*[97] (abstract only) compared Quantum Blue POCT with the Bühlmann ELISA in 128 samples, in patients aged 16–72 years.

Cut-off values used were:

- *negative for IBD* < 50 μg/g faeces (as suggested by manufacturer)
- *positive for IBD* > 200 μg/g faeces
- *intermediate zone* 50–200 μg/g faeces, result uncertain.

Coorevits *et al.*[97] noted that FC values up to 210 μg/g faeces have been described in patients with IBS, and used an ELISA cut-off of > 200 μg/g as indicative of IBD. They had 50 patients with results above this level and 83 below.

They found good correlation ($R^2 = 0.89$) between the tests. After applying different cut-offs to the POCT and assessing the numbers of discordant results between POCT and ELISA, they concluded that 30 μg/g for ruling out inflammation, and 110 μg/g for confirming it, appeared to be the most suitable cut-offs for the POCT. This left a grey zone of 30 to 110 μg. This gave 89.4% (127/142) agreement with the ELISA and 10.6% (15/142) mismatches.

TABLE 30 Comparison of POCT and ELISA tests in Dolci and Panteghini[96]

Test	Positive	Negative	Total
ELISA (cut-off 90 μg/g)	20	47	67
POCT (cut-off 200 μg/g)	17	50	67

TABLE 31 Comparison of tests in Hessells *et al.*[98]

| Test | Cut-off level (µg/g) | | Correct classification | Sensitivity | Specificity | PPV | NPV |
	Quantum Blue	TRFIA					
Quantum Blue	30	50	0.77	0.96	0.69	0.55	0.98
	40	50	0.86	0.92	0.84	0.69	0.96
	50	50	0.88	0.88	0.84	0.68	0.94
	60	50	0.85	0.79	0.87	0.7	0.91
	Cut-off level (µg/g) – ELISA CalDetect						
CalDetect	15	50	0.65	0.96	0.53	0.44	0.97
	60	50	0.78	0.88	0.74	0.57	0.94

Studies comparing different point-of-care tests

Hessells 2012

Hessells *et al.*[98] compared two rapid tests, Quantum Blue and PreventID CalDetect, using the laboratory quantitative time-resolved fluorimetric immunoassay (TRFIA) as the gold standard, using a cut-off of 50 µg/g (*Table 31*). The PreventID is a rapid semiquantitative test, with lines: negative, < 15 µg/g, has two lines; positive, > 60 µg/g, has four lines, and indeterminate, 15–60 µg/g, has three lines.

The TRFIA test is reported to have some advantages over ELISA (better precision, wider range, greater sensitivity) but need not be considered further here – its role is simply as gold standard comparator for the two rapid tests.

The patient group was a mixture of new referrals with suspected IBD (*n* = 40), and suspected relapses (*n* = 45) referred to a Dutch gastroenterology unit. Performance was assessed at four cut-off levels for Quantum Blue (30, 40, 50 and 60 µg/g) and two cut-off levels for CalDetect (15 and 60 µg/g) (the lowest and highest detection levels). The same samples were used for rapid and TRFIA testing. A TRFIA level of 50 µg/g was used as the golden standard test performance.

Optimal cut-off levels were 40 µg/g for the Quantum Blue test (NPV 0.96, sensitivity 0.92, specificity 0.69) and 15 µg/g for the Caldetect test (NPV 0.97, sensitivity 0.96, specificity 0.44).

The correlation between the rapid tests and TRFIA was good for both tests (kappa test, *p* < 0.0001), but significantly better for Quantum Blue (kappa 0.77; 95% CI 0.64 to 0.90) than for CalDetect (kappa 0.46; 95% CI 0.32 to 0.60).

The authors concluded both tests performed well, but that the Quantum Blue test was superior (at cut-off of 40 µg/g) to the PreventID CalDetect in reducing the number of colonoscopies. Because of its high NPV, the number of colonoscopies might be reduced by 62%. The Quantum Blue can be used with a POC reader giving a quantitative result.

Studies comparing point-of-care testing with enzyme-linked immunosorbent assay

Sydora 2012

Sydora *et al.*[99] also compared the Quantum Blue with a standard calprotectin ELISA method (from Alpco Immunoassays, Salem, NH, USA – probably Bühlmann). The participants included patients with UC, CD or IBS, and volunteers with no known intestinal problems.

The IBD patients group had significantly higher calprotectin levels than IBS patients and healthy controls ($p = 0.01$). There was no difference in calprotectin concentrations between IBS patients and controls. Some patients with IBD who had undergone recent surgery had calprotectin levels similar to controls and patients with IBS.

Results were available in 8 hours from the ELISA method but in 30 minutes from Quantum Blue. However, the ELISA had a much wider range. Quantum Blue has a minimum measurement value of 30 µg/g and a maximum of 300 µg/g. This would not be a problem for the NICE decision group of patients, where the focus is in distinguishing between normal and raised levels. So as soon as the level is abnormal, referral is triggered and the height does not matter at this stage.

Sydora et al.[99] concluded that:

- With Quantum Blue, a cut-off at 150 µg/g distinguishes healthy control subjects and IBS patients from those with active IBD with a specificity of 100% (after excluding IBD patients who had undergone recent surgery).
- ELISA testing gives the same specificity at a cut-off of 230 µg/g.
- The desk-top Quantum Blue is as accurate as the ELISA in distinguishing between inflammatory and non-inflammatory intestinal disorders but can do so in 30 minutes compared with 8 hours for the ELISA.

Vestergaard 2008

Vestergaard et al.[100] compared the semiquantitative PreventID Caldetect with the PhiCal ELISA in 95 samples from 82 patients and 13 healthy volunteers with no history or symptoms of bowel disease. The patients had IBD (27 CD, 15 UC, 3 indeterminate colitis); chronic diarrhoea (24); abdominal pain (6) or other reasons. The age range of the patients was 2–86 years. Their results are shown in *Table 32*.

Correlation was good but 18 patients had a positive calprotectin ELISA test but were negative with the rapid test. The authors used the recommended ELISA cut-off of 50 mg calprotectin/kg stool.

Vestergaard et al.[100] regarded a calprotectin concentration by the rapid test of < 15 µg/g as reliable for excluding IBD. With calprotectin concentrations of > 15 µg/g, they recommend checking the POCT result by quantitative measurement. So the PreventID could be a useful screening test to rule out inflammation.

Shastri 2009

Shastri et al.[101] compared the Immundiagnostik ELISA with the PreventID CalDetect, with a cut-off of 15 ng/ml for both tests, in 823 patients. The ELISA had slightly better sensitivity and specificity than the POCT (e.g. sensitivity for CD 96% vs. 93%, specificity 89% vs. 83%) but these were not significantly different. For NPV the different was greater (88% vs. 79% for CD; 83% vs. 76% for UC), but, again, CIs overlapped.

Shastri et al.[101] report that the POCT can be done in 5 minutes, which is less than most reports.

TABLE 32 Comparison of PreventID with ELISA

	Sensitivity, % ELISA cut-off 50 mg calprotectin/kg stool	Specificity, % ELISA cut-off 50 mg calprotectin/kg stool	PPV (%)	NPV (%)
Rapid test, cut-off 60 µg/g	66 (52–79)	100 (53–80)	100 (90–100)	72 (60–83)
Rapid test, cut-off 15 µg/g	96 (87–100)	70 (55–83)	79 (67–88)	94 (81–99)

Labaere 2013

Labaere et al.[102] compared eight different assays for calprotectin: four ELISAs, three POC (Quantum Blue, Eurospital Calfast, Biotest Certest) and the automated immunoassay from Phadia (details available only from a recent meeting abstract) and poster. They compared the tests for both distinguishing IBD from non-IBD, and for monitoring IBD, and also compared results with endoscopic and histological findings. They reported that sensitivity (82–83%) and specificity (84–89%) were similar among the assays. For distinguishing IBD from non-IBD, they concluded that the best tests were Quantum Blue, Phadia and Calprolab, with ratios of median IBD to non-IBD of 14, 12 and 10. They conclude that 'All calprotectin assays showed acceptable and comparable clinical performance for diagnosis of IBD.'

They also conclude that the quantitative POCTs could replace the ELISA tests. They had reservations about the comparative merits of the tests for monitoring disease activity but that is not relevant to this review.

Summary

The overall message from these studies is that the point-of-care tests are about as good as the ELISA tests.

Studies comparing enzyme-linked immunosorbent assay tests

Whitehead 2013

Whitehead et al.[103] compared three ELISA tests: Immundiagnostik, Bühlmann and Eurospital. All assays performed satisfactorily but the Bühlmann test gave higher results. They suggest that each laboratory determines its own reference range.

Loitsch 2010 (meeting abstract)

The tests compared were Immundiagnostik (Calp-ID) and Bühlmann EK-CAL (Calp-Bu)[104] (*Table 33*).

The patients were a mixture of those with active IBD, those with IBD in remission, and those with IBS. There were 108 patients with IBD (77 active and 31 in remission), and 96 with IBS. Loitsch et al.[104] used the manufacturer's cut-off values. The sensitivities, specificities and accuracy of the Calp-ID and the Bühlmann EK-CAL were as shown. Note that the specificity appears to be the same in all three groups but that other parameters vary.

The specificity of Calp-ID is higher and sensitivity is lower, but the overall accuracy is higher. Note the much lower specificity for both tests. The authors concluded that both tests provide a reliable and non-invasive way of differentiating IBD from IBS but that the Calp-ID was the more accurate test.

Tomkins 2012 (meeting abstract)

Tomkins et al.[105] from Coventry compared two ELISAs: Immundiagnostik PhiCal version 1 (PhiCal 1) and the Bühlmann EK-CAL in 62 patients, of which 38 had IBD or other organic pathology (age range

TABLE 33 Comparison of tests in Loitsch et al.[104]

Comparison	Specificity (%)		Sensitivity (%)		Accuracy (%)	
	Calp-Bu	Calp-ID	Calp-Bu	Calp-ID	Calp-Bu	Calp-ID
IBS (*n* = 96) vs. active IBD (*n* = 77)	63.40	79.20	97.40	93.50	78.60	82.10
IBS (*n* = 96) vs. active colitis (*n* = 77)	63.40	79.20	100.00	97.80	75.30	85.20
IBS (*n* = 96) vs. active CD (*n* = 41)	63.40	79.20	95.40	90.20	72.90	82.50

15–49 years, mean 36 years), and 24 had IBS (age range 20–48 years, mean 36 years) (*Table 34*). All participants had a colonoscopy with biopsy.

Hence using 50 µg/g as the cut-off, PhiCal 1 performed slightly better than EK-CAL but NPVs were similar.

Results from different FC methods are not directly comparable, despite widespread adoption of single cut-offs.

General practitioner assessment and referral: implications for modelling

Adults

As noted previously, we lack published data on the use of calprotectin testing in primary care. However, we have the results from the NTAC pilots,[36] and these provide data on referral patterns by GPs in the UK (assuming that those in the North East are representative).

The Durham Dales pilot provides data on GP referrals with no calprotectin testing, and the effect that testing would have. The data allow us to explore what might happen if calprotectin testing is made available.

The test used was the POCT PreventID, which divides people into three groups:

- negative < 15 µg/g
- positive > 60 µg/g
- intermediate ≥ 15 µg/g but ≤ 60 µg/g.

General practitioners made diagnoses based on clinical assessment without knowledge of the calprotectin results. They referred those that they thought might have IBD, and managed those that they thought had IBS in primary care.

A final consultant diagnosis was made, based on calprotectin test results and clinical data, including endoscopy. The clinical data came from GP and outpatient data, where patients were referred, or just from GP data, when patients were not referred. Note that those diagnosed as IBS (and not referred) did not have colonoscopy so it is not possible to completely exclude false negatives. These would have IBD but appear clinically to be IBS and have negative calprotectin results. Such false negatives are unlikely given the high sensitivity (100% – see *Figure 3*) of calprotectin in this POCT at the 15 µg/g cut-off, but not impossible. (The Durham Dales pilot could not be used in our main assessment because of the lack of a definitive reference test.)

For assessing the sensitivity and specificity of GP assessment, there are two options using the Durham Dales pilot data:

1. use calprotectin as reference test
2. use final consultant diagnosis.

TABLE 34 Comparison of tests in Tomkins *et al.*[105]

Test	Cut-off value (µg/g)	Sensitivity, % (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Bühlmann EK-CAL	50	86 (42 to 99)	60 (33 to 83)	50 (22 to 78)	90 (54 to 99)
PhiCal 1	50	78 (40 to 96)	92 (60 to 100)	88 (47 to 99)	86 (56 to 97)

If we compare GP diagnosis with calprotectin levels, and assume that a positive calprotectin test implied possible IBD and an indication to refer, then we have a 2 × 2 table as shown by *Table 35*.

So sensitivity 28/34 = 82% and specificity 79/83 = 95% where 'positive' = a positive FC test. If we exclude the two indeterminates (who would be re-tested, rather than referred), sensitivity is 88%. Of the 83 diagnosed as IBS, only four had high calprotectin levels, a 5% error rate giving NPV of 95%. Note that the four are not false negative in the sense of being missed IBD but in the sense of being 'false non-referrals'. Not all would have IBD. So without FC testing, GPs would not refer 4 of 32 patients with high calprotectin.

Given that the above comparison is of GP, versus calprotectin, the final consultant diagnosis is more useful for our purposes. *Table 36* compares the GP diagnosis (without knowledge of calprotectin result) and the consultant diagnosis [with knowledge of calprotectin result and of endoscopy where performed. Note that far more patients (33) had endoscopy than were found to have IBD].

Numbers are slightly less than in the previous table because some patients do not appear to have been followed up. No data are given in the YHEC report on the presumed diagnosis or calprotectin results in five missing cases.[36] The sixth was found to have cancer.

These results show that the GPs referred all those diagnosed as IBD, giving a 'whole pathway' sensitivity of 100% (if we assume there were no false-negative IBDs as discussed above.) 'Whole pathway' combines GP assessment, calprotectin testing and consultant opinion based on clinical data that included endoscopy (mainly colonoscopy but some flexible sigmoidoscopies).

However this is achieved at a specificity of 79% for GP assessment without calprotectin testing. Without calprotectin testing, GPs refer a group of whom around 25% have IBD (7 of 29) and 75% have IBS. This matches results from routine care – that > 60% of colonoscopies in young people are normal.

This implies that if GPs had access to calprotectin testing, they might be able to reduce referrals by a considerable amount – about three-quarters. The Durham Dales data suggest that GPs refer about a quarter of patients presenting to them with GI symptoms that could be due to IBS or IBD. The number of patients in the study is quite small, but that proportion is similar to the figure of 29% reported in the BSG guideline on IBS, which increases confidence in the pilot data.[106]

TABLE 35 General practitioner diagnosis compared with calprotectin level

	FC positive	FC negative	Total
GP IBD	28	4	32
GP IBS	6 (4 high, 2 indeterminate)	79	85
Total	34	83	117

TABLE 36 General practitioner diagnosis compared with final consultant diagnosis

	Consultant IBD	Consultant IBS	Total
GP IBD	7	22	29
GP IBS	0	82	82
Total	7	104	111

The prevalence of IBD in the whole population was 6.3% (standard error 2.3%) but amongst those referred it was almost 25% (6.3 of 25%).

In the pilot, a GP decision to refer set a patient on a pathway that could lead to colonoscopy and possible other invasive investigations. This decision would not be taken lightly. However, if FC testing is introduced, we might expect that GPs would consider testing in a wider patient group then they would consider for referral. They refer only about 25% of those that present to them with these symptoms. We can create a scenario analysis assuming that if calprotectin testing becomes available then GPs will test twice as many as they would have referred in the absence of FC testing.

We also note from the Durham pilot that if GPs thought that a patient had IBS, they were right at least 95% of the time because only 5% of those they thought had IBS had high calprotectin and needed referred (NPV 95%). These 'false non-referrals' could theoretically include some with IBD. In our scenario analysis, we assume that all patients with IBD will be in the larger group (50% of all patients with symptoms, so 222 patients) that will have calprotectin testing. If we assume that 50% of patients with symptoms will be tested, we get figures as shown in *Table 36*. All of the 6.25% of patients with IBD are tested, and assuming that the POCT sensitivity of 100%, no patients with IBD would be missed.

If we used an ELISA test, with a sensitivity of 93% (from meta-analysis, we would miss 0.44%, or 0.49 patients, in the numbers in this group).

The extra group consists of those regarded by the GP as less likely to have IBD than the 25% (because the GP did not refer them), and the GP is really doing the test to confirm IBS. The false positive rate among the additional 25% tested, will therefore be much less than in the 25% referred. One option is to assume that there will be no new false positives.

So figures change to those shown in *Table 37a*.

The prevalence of IBD in the tested group is half that in the referred group – about 12.5%. As all of those with IBD are tested, there are no false-negatives if we assume sensitivity of calprotectin testing to be 100%. Specificity is 90%. If we assumed that there would be more false positives, specificity would be 80% if we double the false positives to 44%, and 85% if we increased them to 33.

If the calprotectin test was the average ELISA with sensitivity 93% and specificity 94%, the figures in the *Table 37a* would change to those shown in *Table 37b*.

Only 9% would be referred owing to the greater specificity of ELISA, but 0.49 patients would be missed.

If we assume that only patients with raised calprotectin are referred, and that calprotectin is 100% sensitive for detecting newly presenting (and hence active) IBD, then with calprotectin testing, GPs will refer about 9% (20/222) compared with the 25% referred when they have no calprotectin testing available – a drop of around three-quarters. However, not all of the calprotectin false positives would be referred if GPs, aware of the imperfect specificity of the test, used clinical judgement and a repeat test with the more specific (94%) ELISA test before referral. That would reduce number referred to about 20 (approximately seven true positives and 13 false positives) or 9% – a drop of over 60%.

So for modelling purposes, using the PreventID test, we can use a prevalence of IBD of 6.3% (among all people with symptoms), and in the absence of FC testing, a sensitivity of GP referral of IBD of 100%, and 79% specificity.

Using the North European data from Shivananda *et al.*,[29] we would expect in this adult group, a ratio of UC to CD of 3 : 2. (Incidence of UC 12.9 in 15- to 44-age group, based on 539 cases; of CD 8.7, based on 365 cases.)

TABLE 37a Expected numbers if 50% of presenting patients are tested with FC

	IBD	IBS	Total
GP + FC IBD	7	22	29
GP + FC IBS	0	193	193
Total	7	215	222

TABLE 37b Expected numbers if 50% of presenting patients are tested with FC: ELISA test

	IBD	IBS	Total
GP + FC IBD	6.51	13.49	20
GP + FC IBS	0.49	201.51	202
Total	7	215	222

Note that there are some weaknesses in the above arguments:

1. The 50% is a rather arbitrary assumption. We have reasonably assumed that more patients with symptoms would have calprotectin testing than were referred when testing was not available but we cannot say if 50% is correct. Given that GPs are good at diagnosing IBS, we would not expect 100% to be tested.
2. Our base-case assumption is that doubling the number tested would not increase the number of false positives. As the extra 25% tested would have less severe symptoms than the first 25% (referred), it seems reasonable to rule out a doubling of false positives. However, assuming no increase may be too optimistic.
3. The 100% sensitivity for the point-of-care test is based on only one study[73] with not very large numbers, and needs to be replicated in a larger study. The mean ELISA sensitivity was 93%. However, GPs would not simply rely on the test results alone, knowing that sensitivity was not perfect, and some of the false negatives on ELISA testing might be referred on clinical nous.

Children

Modelling requires different assumptions in children. Based on the recent UK study by Henderson *et al.*,[30] 48% of referred cases (91/190) had IBD. The ratio of CD to UC is much higher, 2.3 : 1.
The potential reduction in colonoscopies is therefore greater.

Chapter 3 Economics

In the following, all costs have been converted to 2011 prices using the Personal Social Services Research Unit (PSSRU) Hospital and Community Services Costs (HCSC) price index.[107] Any foreign currency amounts have been converted to pound sterling at the contemporaneous April 2005 exchange rate, with these amounts then being converted to 2011 prices using the PSSRU HCSC price index. Where the base year is not given within the paper, it has been assumed to be the year of publication. The original amounts are given in square brackets.

A review of the cost-effectiveness literature for FC testing is presented. The cost-effectiveness studies of this review inform the de novo economic model structure but do not directly contribute to its inputs. The summary of them is intended to outline what cost-effectiveness work has previously been undertaken in the area, mainly by way of background. This is followed by a review of studies of QoL that may suitable for inclusion in a cost–utility analysis of FC testing, health-related QoL (HRQoL) for three conditions having to be considered: IBS, CD and UC. Given the centrality of colonoscopy to the question in hand, a brief review of the adverse events associated with colonoscopy is then presented. A relatively simple cost consequence model of FC testing is then presented, augmented by some considerations around the loss of utility among false negatives during their period of incorrect treatment. This is followed by a full cost–utility model, much of the structure of this being drawn from the modelling for CG61: Diagnosis and Management of Irritable Bowel Syndrome in Primary care;[35] the modelling for CG152: Crohn's Disease: Management in Adults, Children and Young People;[108] and the modelling for the current draft of the UC guideline: Ulcerative Colitis: Management in Adults, Children and Young People.[109]

Faecal calprotectin tests economic literature

Hornung and Anwar[110] analysed the results of the 40 patients who had FC testing between January 2009 and April 2010. This appears to be all the patients tested with FC within the North Tees and Hartlepool NHS Foundation Trust. No detail is given of which calprotectin method was used. Patients were split into those with IBS-like symptoms of unknown cause in whom IBD needed to be ruled out ($n = 22$) and those with known IBD ($n = 18$). Nine per cent (two people) of the first group and 61% of the second group had a high (level not stated) calprotectin result. But Hornung and Anwar[110] note that in the first group FC testing did not result in a change in treatment for any patient, compared with 12 of the 18 IBD patients having their treatment changed as a result of the FC result. As a consequence, it appears that of the eight colonoscopies avoided, none was in the group of patients with IBS like symptoms of unknown cause in whom IBD needed to be ruled out. However, it is reported that 13 of the 17 colonoscopies in the newly presenting group were normal, although it is not clear if the four with abnormal findings had IBD, nor if they included the two newly presenting patients with high calprotectin.

Mindemark and Larrson[111] undertook a cost minimisation analysis comparing the diagnostic pathway of FC testing followed by colonoscopy with direct referral to colonoscopy, the aim being to rule out IBD. For FC two cut-offs were used: 50 and 100 μg/g. The study data were drawn from a retrospective analysis of 3639 Swedish patients. Test costs were £24.57 [€29] for FC, £576.20 [€680] for colonoscopy in adults and £1152.40 [€1360] for colonoscopy in paediatric patients. One-third of patients were paediatric, with a further 13% being aged over 65 years. In the paediatric group 54% had a FC of < 50 μg/g, whereas 71% were < 100 μg/g. In those aged 18–65 years, 52% had a FC of < 50 μg/g, whereas 68% had < 100 μg/g, with the respective percentages for those over 65 years being 30% and 51%. If a threshold of 50 μg/g was used to exclude organic disease, around 50% of colonoscopies could be avoided. If a cut-off of 100 μg/g was used, 67% could be avoided.

The direct costs of the diagnostic strategies were £2,791,680 [€3,294,600] for direct referral to colonoscopy, £1,461,369 [€1,724,611] for FC with a 50 μg/g cut-off and £985,409 [€1,162,931] for FC

with a 100 µg/g cut-off. The study does not consider false negatives and the costs of them subsequently re-presenting, nor does it specify what assumptions were made about sensitivity and specificity – it appears that perfect distinction between inflammatory and non-inflammatory is assumed.

One useful point made is that a reduction in colonoscopy for distinguishing between inflammatory and functional conditions would help resolve waiting list pressures for colonoscopy for other reasons. This is relevant to the UK following the roll-out of colorectal cancer screening. Calprotectin testing may not save money, as the capacity released might be used for other purposes, but it would reduce or avoid the need to expand services to cope with, for example, follow-up after colorectal screening.

Goldfarb *et al.*[112] compared wireless capsule endoscopy with colonoscopy coupled with a small bowel follow through with a barium swallow for the diagnosis of CD. They note that although 75% of patients with CD have small bowel involvement, more than one-third have disease limited to the small bowel. The parameters populating the decision tree model are not entirely clear but it also included a perforation rate for colonoscopies of 0.03% and retention of the wireless capsule in 0.75% of cases. Costs were based upon Medicare reimbursement rates, with the conclusion that wireless capsule endoscopy had a diagnostic yield of 70% compared with 54% for colonoscopy with small bowel follow-through (SBFT), while also saving an average £197 [US$291].

The YHEC FC testing report[50] for CEP provides the most comprehensive review to date of the economics of FC testing compared with testing with ESR and CRP (*Table 38*). It considers the primary care patient population, presenting with symptoms suggestive of IBS but no 'red flag' symptoms. Those testing positive are assumed to be referred to secondary care for colonoscopy. Those testing negative are treated as patients with IBS, initially with dietary and lifestyle advice. Among those testing negative, for those with IBS, 50% do not respond to dietary and lifestyle advice, whereas, for those with IBD, 100% do not respond to dietary and lifestyle advice. Non-responders seek further advice and medication from their GP after 2 months. Among non-responders, for those with IBS, 5% do not respond to further medication, whereas, for those with IBD, 100% do not respond to further medication. These are then referred on to secondary care for further investigation. For the base case, all investigations are assumed to be colonoscopy with no sigmoidoscopy, although both are assumed to have 100% sensitivity and 100% specificity.

For the base case comparing ESR/CRP against FC, FC was found to be dominant owing to FC correctly diagnosing more patients with IBS and IBD at lower cost. Note that within the model all false negatives eventually re-present and, as a consequence, all IBD patients are eventually correctly diagnosed.

A second comparison comparing FC ELISA with FC POCT found that FC ELISA was more expensive overall. Although FC ELISA diagnosed more IBD patients correctly, owing to its poorer specificity it also resulted in

TABLE 38 York Health Economics Consortium report[50] base-case parameter values

Presenting population					
Prevalence of IBD in presenting population, %					10
IBS patients uncontrolled on dietary advice alone, %					50
Of whom uncontrolled on medication and requiring further investigation, %					5

Test characteristics					
	ESR/CRP	**FC**	**FC ELISA**	**FC POCT**	**Colonoscopy**
Sensitivity, %	35	90	96	61	100
Specificity, %	73	80	87	98	100
Cost, £	4.64	25.00	25.00	27.68	544

more patients with IBS being incorrectly sent for colonoscopies. This was the source of the additional costs under FC ELISA. This underlines the importance of the specificity of the tests, particularly given the relatively low prevalence of IBD in the presenting patient population.

Results were sensitive to the prevalence of IBD in the presenting population, sensitivities and specificities and the costs of the tests.

Mascialino et al.[113] (sponsored by Thermo Fischer Scientific) augment the YHEC model[50] with a third branch in the decision tree model for FC for indeterminate results of between 50 and 250 µg/g. Those with FC levels of > 250 µg/g follow the YHEC positive result branch,[50] and those with FC levels of < 50 µg/g follow the YHEC negative result branch, whereas those with an indeterminate result receive a second test. Unfortunately, possibly owing to being only a poster presentation, quite how the new indeterminate branch is populated is unclear. The tables of the poster still report only sensitivities and specificities. The overall conclusions mirror those of the YHEC report,[50] only with more correct diagnoses and larger cost savings. But it is unclear how these have been arrived at owing to the lack of detail about how the third branch of the model has been populated.

In another conference abstract, Mascialino et al. (2013)[51] report that in modelling, calprotectin dominates CRP and ESR being more accurate and less costly, with an estimate of £100 lower cost per patient in the UK.

Dubinsky et al.[114] modelled the cost-effectiveness of three main alternative diagnostic strategies in the US context using serological markers: ASCA for CD, and pANCA for UC. They developed a decision tree model comparing immediate referral for colonoscopy, with the possibility of barium upper GI investigation and a SBFT, with two diagnostic testing strategies: a primary ASCA and PANCA assay, with subsequent referral to colonoscopy, and a sensitive primary assay followed by a more specific second confirmatory assay with subsequent referral to colonoscopy (Table 39). Costs were taken from the Medicare fees schedule. Those with negative results could return for further testing after 2–3 months, with 50% of true negatives with IBS representing, and presumably 100% of false negatives with IBD representing. Unusually, based upon expert opinion, Dubinsky et al.[114] do not assume 100% sensitivity and 100% specificity for colonoscopy.

Cost-effectiveness over a 1-year period was measured as the cost per correct diagnosis, this encompassing both correct diagnoses of IBD and correct diagnoses of IBS. Given this definition of effectiveness, the sequential testing strategy resulted in more correct diagnoses that both direct referral and a single primary test before referral: 97.90% accuracy compared with 95.95% and 96.95%, respectively. The sequential testing was also cheaper than both direct referral and a single primary test before referral: £1511 (US$1641) compared with £2015 (US$2189) and £1740 (US$1890), respectively. As a consequence, the sequential testing was found to dominate both direct referral and a single primary test before referral.

TABLE 39 Dubinsky model[114] inputs

Presenting population			
Prevalence of IBD in presenting population, %			20
IBS patients with persistent symptoms, %			50
Test characteristics	**Primary assay**	**Second assay**	**Colonoscopy**
Sensitivity, %	80	65	95
Specificity, %	50	65	95
Cost	£48 (US$54)	£48 (US$54)	£1730 (US$1880)

The results for sequential testing are of interest, although their relevance is limited to a degree by effectiveness being measured in terms of correct diagnoses, and so not distinguishing between correct diagnoses of IBS and correct diagnoses of IBD. The relevance of the results is limited by ASCA and pANCA not being regarded as comparators in the NICE scope, but it does illustrate the possible benefits of sequencing tests.

Quality of life

Health-related quality of life studies have been summarised if they provide direct estimates of utilities, or provide supporting data on either the differences in QoL between any of the conditions under consideration or data on the QoL related to symptom severity. Patient characteristics in terms of age, sex, disease severity and disease duration are not presented in the text but to the extent they are available are presented in the summary tables. The justification for the values selected for use in the de novo modelling is presented below (see *Cost–utility modelling*).

Irritable bowel syndrome: quality-of-life studies

Akehurst *et al.*[7] (sponsored by Novartis) undertook a survey of patients with IBS in the UK primary care setting. A sample of 161 patients with IBS was selected from GP lists based upon the Rome I Criteria, with an additional 213 control patients being selected. Controls were matched for age, sex and social characteristics by the patient's GP. The SF-36, EQ-5D and Irritable Bowel Syndrome Quality of Life (IBS-QOL) questionnaire were administered at baseline and subsequently at 3 months. Patients with IBS reported average baseline SF-36 values that were statistically significantly worse than those reported by the control group for every dimension of the SF-36. Similarly, for the EQ-5D the mean baseline score reported by those with IBS of 67.5 was statistically significantly lower than that of the control group. Unfortunately, it is not clear what EQ-5D rating algorithm was used for this, with it being described as the '*EQ-5D derived score*', and there is no reference to the UK social tariff. The mean baseline EQ-5D rating scale, presumably the EQ-5D visual analogue scale (VAS), reported by those with IBS of 64.2, was also statistically significantly worse than the 80.3 of the control group. Parallel statistically significant differences were also reported at 3 months, although the mean QoL values between baseline and 3 months were not statistically different.

In what was apparently a follow-up study to Akehurst *et al.*,[7] Ricci *et al.* (sponsored by Novartis)[115] compared the HRQoL of 305 IBS patients selected from GP lists with 330 controls. Unfortunately, Ricci *et al.*[115] is only available as an abstract but notes a statistically significant relationship between the severity of IBS and patients' reported VAS scores. No further detail on this is provided.

Bernklev *et al.*[116] also report SF-36 scores among IBS patients and compares these with French population norms, finding IBS to significantly adversely affect scores, but no overall QoL values are reported.

Within a broader paper comparing the performance of the EQ-5D and SF-36 across seven patient groups, Brazier *et al.*[117] apply the EQ-5D and SF-36 to 161 UK IBS patients recruited from primary care. All patients were observed twice. The EQ-5D was evaluated using the UK social tariff, whereas the SF-36 was valued using the Short Form questionnaire-6 Dimensions (SF-6D) algorithm. Based upon 314 responses, the mean EQ-5D index was 0.662 compared with 0.666 for the SF-6D. It appears probable, given the similarity of baseline characteristics and results with those of Akehurst *et al.*[7] (and that Brazier was a named author of the Akehurst *et al.*[7] paper) that the IBS patient group and responses were the same for both papers.

Bracco *et al.*[118] (sponsored by Novartis), in an economic evaluation of tegaserod compared with placebo for IBS, evaluated the EQ-5D responses using the UK social tariff of IBS patients: 247 receiving tegaserod and 238 receiving placebo. The adjusted average baseline utility was 0.726. At week 4, these had

improved to 0.795 in the tegaserod group and 0.759 in the placebo group, but by week 12 had fallen back slightly to 0.792 and 0.747, respectively.

Dibonaventura et al.[119] (sponsored by Novartis) compared SF-6D utilities among 109, 83 and 204 patients with IBS-C in the UK, France and Italy respectively with matched controls in the UK, France and Italy. Respondents were recruited through the National Health and Wellness Survey, a self-administered internet based survey. UK patients with IBS-C reported a mean utility of 0.65 compared with 0.71 for their matched controls, the corresponding figures being 0.63 and 0.71 for the French sample and 0.66 and 0.70 for the Italian sample.

Pare et al.[120] (sponsored by Novartis) reported the UK social tariff EQ-5D index among 1555 Canadian primary care patients. The patient group recruited had mainly IBS-C or IBS-alternating type (IBS-A), apparently owing to a desire for the results to be relevant to patients eligible for tegaserod. The mean EQ-5D index was 0.64.

Puhan et al.[121] applied time trade-off (TTO), standard gamble (SG) and the SF-36 to 96 Canadian patients with IBS. Patients were identified through either medical records of the gastroenterology clinic at McMaster Health Sciences Centre or through 10 local gastroenterologists. The SF-36 was valued using the SF-6D transformation. This resulted in three mean estimates for HRQoL: 0.84 for SG, 0.76 for TTO and 0.85 for SF-6D.

Spiegel et al.[5] (sponsored by Takeda) administered the EQ-5D among 257 American IBS patients, with these split into three groups of those with constipation IBS-C, those with diarrhoea IBS-D and those with mixed, IBS-M.[5] Note that the severity of disease within the sample was mixed, with 16% having mild disease, 32% having moderate disease and 55% having severe disease. Patients were followed up at 3 months. It is unclear what algorithm was used to construct the EQ-5D utilities. There was no statistically significant difference in the mean utilities of 0.76 for those with IBS-C, 0.76 for IBS-D and 0.73 for IBS-M. There was a statistically significant difference between the mean utilities of 0.70 for those with severe disease compared with 0.80 for those with non-severe disease, and between the mean utilities of 0.78 for those experiencing considerable relief of symptoms at 3 months compared with 0.73 without considerable relief of symptoms at 3 months.

Wang et al.[4] administered the EQ-5D among 198 people with IBS and 251 people without IBS. These were recruited from those attending the National Foundation for Digestive Diseases Symposium, a free public education symposium held at the Raffles Hotel in Singapore. The EQ-5D was evaluated using the UK social tariff. The mean utility among those with IBS was 0.739, which was statistically significantly lower than the 0.849 of those without IBS.

Inflammatory bowel disease: quality-of-life studies

Konig et al.[122] recruited 121 outpatients and 31 inpatients with IBD from German hospitals, with 123 having CD and 29 having UC. Both groups had a mean of around two active phases in the past year, with 60% of Crohn's patients in remission compared with 70% of UC patients. A total of 79% of outpatients were in remission compared with only 7% of inpatients. The German version of the EQ-5D was administered. The EQ-5D utility index was calculated using the German population mapping function, which asked respondents to rate EQ-5D health states using the EQ-5D VAS. Konig et al.[122] found that 30% of outpatients and 19% of inpatients classed themselves as having health state 11111, i.e. having no problems in all EQ-5D dimensions. Sixty-four per cent of outpatients and 45% of inpatients classed themselves as having no problems in four dimensions, with only one dimension being classed as having some problems. A histogram of the EQ-5D VAS shows a steady increase in the proportion of patients classifying themselves over the range 0–80 and then tailing off again, but the EQ-5D index shows very few respondents classifying themselves as having a utility of < 50. But note that the German EQ-5D index algorithm apparently tends to value health states more highly than the UK social tariff, with this applying with particular force to worse health states. The mean EQ-5D index was 0.875 for those in remission compared with 0.627 for those with active disease, whereas for outpatients it was 0.803 compared with 0.619 for inpatients.

Leidl et al.[123] administered the EQ-5D among 270 patients with CD and 232 patients with UC. Patients were recruited from the German Patients' association for inflammatory diseases, and were split into slight [Crohn's Disease Activity Index (CDAI) 0–3], moderate (CDAI 4–7) and severe (CDAI > 7) subgroups. Leidl et al.[123] also applied both the German and the UK social tariff in their valuations. The mean value for both tariffs were broadly similar, with the exception of CD patients with severe disease who were assigned a QoL of around 0.50 using the German tariff, but 0.33 using the UK social tariff. The mean values applying the UK social tariff, taken from the graph, for mild, moderate and severe disease were 0.87, 0.67 and 0.33 among patients with CD, and 0.91, 0.73 and 0.67 for UC patients.

Stark et al.[124] contacted a random sample of 724 patients with CD and 723 patients with UC from the German IBD association (DCCV), the largest voluntary support organisation of IBD patients in Germany. Thirty-seven per cent agreed to participate and 36% completed the EQ-5D at baseline: 270 patients with CD and 253 patients with UC. Those with inactive, slight, moderate and severe disease were 57.1%, 33.2%, 9.3%, and 0.4%, respectively, among patients with CD, with 57.1% being in remission, compared with 62.1%, 26.3%, 9.4% and 2.2% for UC patients with 62.1%, being in remission. At baseline the mean utility among patients with CD using the UK social tariff was 0.77 overall, with 0.89 for those in remission and 0.61 for those with active disease. Using the German tariff it was 0.86 overall, with 0.95 for those in remission and 0.75 for those with active disease. At baseline, the mean utility among UC patients using the UK social tariff was 0.84 overall, with 0.91 for those in remission and 0.71 for those with active disease. Using the German tariff it was 0.92 overall, with 0.96 for those in remission and 0.84 for those with active disease. The study also re-administered the EQ-5D at 4 weeks, but the resulting data are used to assess only construct validity and are of limited interest for current purposes.

Turunen et al.[125] mailed 550 Finnish 'paediatric' IBD patients and 1650 age- and sex-matched controls from the same municipality a bespoke questionnaire. Sixty-seven per cent of the patient group and 37% of the control group responded. The IBD patients had been previously identified for another study, through chart review of two major Finnish hospitals. Unfortunately, this resulted in a mean age among responders of 21 years. The questionnaire posed four generic questions on physical, emotional, social and overall QoL, with these being rated on a VAS of range 1–7. The main result of interest is that there were no major differences in mean responses between those with CD and those with UC.

Casellas et al.[126] administered the Inflammatory Bowel Disease Questionnaire (IBDQ) and the EQ-5D among 1156 Spanish IBD patients, 628 with CD and 528 with UC. These were composed of both inpatients and outpatients: 141 and 487 respectively for CD, and 108 and 420 for UC. Among patients with CD 268 were in relapse while 360 were in remission, while among UC patients 212 were in relapse and 316 were in remission. It appears that the valuation of the EQ-5D used the Spanish valuation set as reported in Badia et al..[127] Within a multivariate regression analysis of the IBDQ, Cassellas et al.[126] found that the underlying condition was not statistically significant, with a t-statistic of only − 0.067. The 25th percentile, median and 75th percentile EQ-5D preference values were estimated. For those with CD in remission these were 0.70, 0.80 and 1.00, whereas for those with mild disease they were 0.50, 0.72 and 0.80, and for those with moderate to severe disease they were 0.50, 0.60 and 0.70. For those with UC in remission, these were 0.80, 1.00 and 1.00, whereas for those with mild disease they were 0.50, 0.72 and 0.80, and for those with moderate to severe disease they were 0.50, 0.50 and 0.70.

Bernklev et al.[116] administered the SF-36 among 166 Norwegian patients with CD and 348 Norwegian patients with UC. All patients with IBD or possible IBD in four areas of south-eastern Norway had been identified 5 years previously. At the 5-year follow-up, 200 patients with CD and 454 patients with UC remained diagnosed with IBD, with 166 and 348 of these, respectively, giving their consent to participate in the study. Patients with CD had lower mean scores in all dimensions compared with patients with UC, but the paper does not appear to report whether these were significantly different or not. Both patients with CD and patients with UC had significantly lower mean scores in all dimensions when matched with a reference population. Similarly, splitting patients into those with no symptoms, those with mild symptoms

and those with moderate or severe symptoms saw symptom severity being statistically significant across all dimensions among both patients with CD and patients with UC.

Bassi and Bodger,[128] available only in abstract, conducted face-to-face interviews with 120 IBD outpatients and nine IBD inpatients, directly measuring QoL using TTO, the VAS and the EQ-5D. The average utility scores for CD and UC were 0.84 and 0.89 using TTO, 0.62 and 0.70 using the VAS, and 0.71 and 0.77 using the EQ-5D. For the TTO utilities, disease severity showed a significant negative correlation: − 0.37 for CD when measured by the Harvey–Bradshaw Index, and − 0.42 for UC when measured by the Simple Colitis Activity Index.

Crohn's disease: quality-of-life studies

Arseneau et al.[129] within the context of an assessment of the cost-effectiveness of infliximab for CD perianal fistulas in the USA, undertook a TTO exercise among 32 patients with CD, 17 of whom were fistulising, or had a history of fistulising, and 15 who did not, and 20 health members of the general public (*Table 40*). For reasons that are unclear, the utilities for health states were also differentiated by whether a patient was receiving infliximab or was receiving 6-MP/metronidazole therapy. This resulted in the following values.

Note that in *Table 40*, the HRQoL values for pancreatitis health states are as per the footnote to table 2 of Arseneau et al.,[128] the values reported this table assuming that patients spent only one-quarter of their time with pancreatitis. These values appear to relate to the healthy respondents but this is not entirely clear from the text. Although the absolute values vary, the differences in the HRQoL values for those on treatment with fistula and improved fistula are reasonably consistent at between 0.12 and 0.14. The difference between pancreatitis with and without fistula of 0.10 was similar to the 0.09 difference between paraesthesias with and without fistula among patients with CD, but the corresponding 0.16 difference among healthy respondents was that bit larger.

Buxton et al.[130] (sponsored by Elan Pharmaceuticals) explored the possibility of mapping from the IBDQ and from the CDAI to utilities. The data set consisted of paired contemporaneous observations from patients with moderate to severe CD who participated in either of two natalizumab trials, with over 3000 observations.

TABLE 40 Arseneau et al.:[129] CD TTO utilities

Scenario	Patients with CD	Healthy individuals
Infliximab		
Fistula	0.73	0.77
Improved fistula	0.85	0.91
Perianal abscess	0.62	0.72
6-MP		
Fistula	0.69	0.75
Improved fistula	0.81	0.88
Pancreatitis + fistula	0.47	
Pancreatitis	0.57	
Paraesthesias + fistula	0.66	0.68
Paraesthesias	0.75	0.84

Demographic data were not presented. Both the SF-36 and the EQ-5D were considered, with these being transformed on to utilities using the SF-6D and the UK social tariff, respectively. The mean SF-6D utility was 0.68, with the paper deriving a mapping function for the SF-6D that is non-linear in the IBDQ. The mean EQ-5D utility was 0.70, with the paper deriving a mapping function for the EQ-5D that is linear in the IBDQ: $0.03043 + 0.0043$IBDQ, with an R^2 of 0.45. The authors state that, of the two mapping functions, the EQ-5D mapping function is the preferred mapping function.

Benedini et al.[131] (sponsored by Merck, Sharp and Dohme) applied the EQ-5D to 162 Italian patients with active CD and a CDAI score of more than 150 at baseline, with an additional three 6-monthly follow-up visits. The mean baseline EQ-5D score was 0.558, with this showing a gradual improvement over the follow-up visits to 0.682, 0.728 and 0.739. The valuation method for the EQ-5D is unclear, with the paper referencing the UK social tariff but stating that the values fall on the interval 0–1.

Casellas et al.[132] measured the QoL to 49 Spanish patients receiving infliximab and in remission. The number in remission fell to 42 at 12 months, 32 at 24 months, 13 at 36 months and 13 at 48 months. Casellas et al. (2007)[132] report the 25th percentile, the median and the 75th percentile of the EQ-5D among the patients in remission, using the preference set of the Spanish EQ-5D. At baseline these were 0.8, 1.0 and 1.0, with the median among those in remission remaining at 1.0 over the period of the study, and the 25th percentile never dropping to < 0.8.

Gregor et al.[133] recruited 180 inpatients and outpatients with CD from a single Canadian tertiary centre in order to evaluate QoL using TTO, SG and the VAS.[132] Follow-up data from a second visit 8 weeks later was obtained from 164 of these patients. Patients were ineligible if they required imminent surgical treatment, had a significant comorbidity, had undergone surgery in the last 4 weeks or were not 'judged by the investigators to comprehend the choices being offered by the HRQoL questionnaires'.

Patients were divided into four groups:

1. *Chronically active therapy resistant* Treatment with prednisone at a dose of ≥ 10 mg daily, continuous methotrexate of purine antimetabolites for a minimum of 6 months and a CDAI score of ≥ 150 – 52 patients of whom 62% were women, and a mean age of 35 years.
2. *Chronically active therapy responsive* Treatment with prednisone at a dose of ≥ 10 mg daily, continuous methotrexate of purine antimetabolites for a minimum of 6 months and a CDAI score of < 150 – 34 patients of whom 53% were women, and a mean age of 31 years.
3. *Acute disease exacerbation* A recent flare in activity with a CDAI score of ≥ 150, no steroid or immunosuppressive drug therapy in the 12 weeks preceding the flare, and the initiation of prednisone or 5-aminosalicylic acid treatment – 45 patients of whom 49% were women, and a mean age of 34 years.
4. *Remission* A CDAI score of < 150 for a minimum of 6 months and no systemic glucocorticoid or immune-suppressive drug therapy – 49 patients, of whom 59% were women, and a mean age of 37 years.

At baseline the TTO, the SG and the VAS mean values for these groups were 0.88, 0.74 and 0.61, respectively; for the chronically active therapy resistant, 0.98, 0.86 and 0.82; for the chronically active therapy responsive, 0.89, 0.77 and 0.60; for the acute disease exacerbation, 0.96, 0.88 and 0.84; and for remission and 0.92, 0.81 and 0.71 across all patients.

Three hypothetical disease states were outlined:

1. *Mild CD* Four or fewer bowel movements per day associated with occasional abdominal pain, only occasionally absent from school of work because of illness, and rarely tired or having disturbed sleep.

2. *Moderate CD* More than four but fewer than eight bowel movements per day associated with tolerable abdominal pain and occasional blood, tiredness most days, frequent frustration and concern about the side effects of medication, and frequent absences from school or work because of illness.
3. *Severe CD* More than eight bowel movements per day, frequent abdominal pain and bloody stools, always tired with difficulty sleeping, depressed and frustrated and worries about the need for surgery and the side effects of medication.

The mean results were broadly consistent between the first assessment and the follow-up assessment, with the mean TTOs being 0.95 and 0.96 for mild disease, 0.88 and 0.88 for moderate disease, and 0.73 and 0.71 for severe disease. The mean SGs were 0.81 and 0.82 for mild disease, 0.72 and 0.73 for moderate disease, and 0.50 and 0.54 for severe disease. The mean VAS scores were 0.80 and 0.82 for mild disease, 0.57 and 0.61 for moderate disease, and 0.27 and 0.31 for severe disease. Within these results, although the absolute values for the TTO lie above those of the SG, the net HRQoL changes from moving from mild to moderate disease, 0.07 to 0.09, and from moderate to severe disease, 0.15 to 0.22, are reasonably aligned between the TTO and the SG. The net changes estimated using the VAS are somewhat different: from moving from mild to moderate disease, 0.21 to 0.23, and from moderate to severe disease, 0.30.

Gibson *et al.*[134] (sponsored by Schering-Plough) surveyed 143 patients with the Assessment of Quality of Life (AQoL) questionnaire, recruited from five Australian outpatient clinics. Patients had had their diagnosis of CD confirmed by a specialist physician on standard clinical, radiological, endoscopic and histopathological criteria. Patients with significant comorbidities were excluded. A total of 110 patients were without fistulas, whereas 23 had fistulas. The overall mean CDAI score of 171 was slightly lower at 169 in those without fistulas, and higher at 177 in those with fistulas, but this difference was not significant. Those without fistulas were roughly equally balanced between ileal, ileocolonic and colonic, whereas 64% of those with fistulas were colonic. The AQoL website outlines that the AQoL utilities scoring system is based upon TTO, although the External Assessment Group (EAG) has not explored this in any depth. Among those without fistulas the average HRQoL was 0.646 compared with 0.606 for those with fistulas. For those without fistulas the average HRQoL was 0.766 for those with a CDAI score of < 150; 0.680 for those with a CDAI score of between 150 and 219; and, 0.450 for those with a CDAI score of more than 220. Relating the HRQoL to the CDAI score there was a broadly negative relationship, although this showed quite a wide dispersion of points around the regression line of HRQoL = $0.8198 - 0.00107 \times$ CDAI score and the R^2 was only 0.27.

Although of relatively limited usefulness for cost-effectiveness modelling purposes, Hill *et al.*[135] report QoL values to 41 Australian Crohn's patients who were paediatric at diagnosis. QoL was measured using the IMPACT III questionnaire, composed of 35 questions, each of which was scored on a 1–5 Likert scale, giving a range of possible values from 35 to 205, with a higher score being taken to indicate a better QoL. These were further related to the PCDAI, with patients being grouped into remission with a PCDAI score of ≤ 10, mild disease with a PCDAI score of 11–29, or moderate to severe disease with a PCDAI score of ≥ 30. A multivariate analysis found QoL as measured by the IMPACT III questionnaire to be significantly affected by the PCDAI. Age, gender, disease duration and whether diagnosis was within 6 months were not found to be significant. Whether patients were receiving treatment, either drug or enteral nutrition, was of borderline significance ($p = 0.07$), as was whether the patient was growth impaired as measure by the height z-score ($p = 0.06$).

Ulcerative colitis: quality-of-life studies

Connolly *et al.*[136] (sponsored by Ferring Pharmaceuticals) analysed EQ-5D data from a study of Western European patients with mild to moderately active UC scoring of between three and eight points on the Ulcerative Colitis Disease Activity Index (UCDAI). This compared oral mesalazine plus a daily mesalazine enema, $n = 71$, with oral mesalazine plus a daily placebo enema, $n = 56$, over a 4-week period, with an additional 4-week follow-up period with no enemas.[136] The proportion of women in the mesalazine enema was 38% compared with 43% in the placebo enema arm, and the median ages were 42 years and

47 years, respectively. The EQ-5D was administered at baseline, week 2, week 4 and week 8. The paper does not appear to report what value set was used to convert the EQ-5D to utility scores. At baseline the mean EQ-5D index values were 0.778 in the mesalazine enema arm and 0.762 in the placebo enema arm. These showed continuous improvement over the study period, including between week 4 and week 8 when the enemas had been discontinued, reaching 0.914 and 0.862 at week 8.

Bryan et al.,[137] in the Evidence Review Group (ERG) report for the STA of infliximab at a dose of 5 mg/kg for the treatment of acute exacerbations of UC, summarised the HRQoL data within the manufacturer submission. This mainly relied upon the Health Outcomes Data Repository (HODaR) study, which measured the EQ-5D among 171 Welsh UC patients. Additional data for the HRQoL for surgery with complications health state of the submitted model was drawn from Arseneau et al.,[138] which applied the TTO among 48 US patients with UC. The ERG report tabulated these as shown in Table 41.

Note that Feagan et al.[139] (sponsored by Centocor Inc.) used the IBDQ in a study of infliximab treatment for patients with moderate to severely active UC disease at doses of 5 and 10 mg/kg with a placebo control arm. The average IBDQ scores at baseline were 125, 130 and 124, respectively. Applying the IBDQ to HRQoL mapping function derived by Buxton et al.[130] for CD to these mean scores results in HRQoL values of 0.568, 0.589 and 0.564. At week 8 the mean improvements in the IBDQ were 40, 36 and 21, respectively, which would translate into HRQoL gains of 0.202, 0.185 and 0.121. These were broadly maintained to week 30. Feagan et al.[139] also noted mean improvements in the IBDQ for those with mucosal healing of 48 at week 8 and 58 at week 30, which translate into HRQoL gains of 0.237 and 0.280. For those without mucosal healing the corresponding IBDQ improvements were only 16 and 7, which translate into HRQoL gains of 0.099 and 0.061.

Waljee et al.[140] used TTO to measure the QoL and perceived QoL with and without colectomy among US ulcerative patients recruited from primary care without a colectomy and ulcerative patients post colectomy. Unfortunately, throughout their paper Waljee et al.[140] report only the median values, although this is mitigated by the 25th percentiles and the 75th percentiles also being reported. For current purposes, the more interesting results are the QoL values recorded among patients without a colectomy living with chronic mild ($n = 55$), moderate ($n = 47$) and severe ($n = 48$) UC. The medians (IQRs) for these were 0.96 (0.91 to 1.00), 0.94 (0.86 to 0.98) and 0.96 (0.88 to 0.99) respectively, whereas across the group as a whole they were 0.96 (0.89 to 0.99).

Poole et al.[141] (sponsored by Ferring Pharmaceuticals) used trial data from the PINCE clinical trials to map between the UCDAI score of patients and the individual dimensions of the EQ-5D with these subsequently being mapped to utilities, presumably using the EQ-5D UK social tariff, although this does not appear to be stated. The observed EQ-5D utilities were compared with those estimated for both the PINCE trial and the separate PODIUM trial. For those in remission, the mean utilities for PINCE observed, PINCE estimated and PODIUM estimated were 0.944, 0.939 and 0.940, respectively; for those with mild/moderate disease the mean utilities were 0.811, 0.801 and 0.775, respectively; and for those in severe relapse the mean utilities were 0.700, 0.630 and 0.660, respectively. The reasons for the estimated utilities falling below the observed utilities for mild/moderate disease and for severe relapse is not clear.

TABLE 41 Utility estimates associated with health states (from Bryan et al.[137])

Health state	Arseneau TTO		HODaR EQ-5D	
	Mean	SD	Mean	SD
Remission	0.79	0.24	0.88	0.14
Active UC	0.32	0.31	0.42	0.32
Surgical remission	0.63	0.30	0.60	0.38
Surgical complications	0.49	0.32		

Colonoscopy patient impacts and adverse events

Baudet et al.[142] followed up 1126 randomly selected Spanish colonoscopy patients, 78% of whom received sedation for the colonoscopy. Sedation was on request, not randomly allocated. There were two episodes of bleeding, both of which followed the removal of very large polyps. There were no perforations. Early adverse events of bradycardia and hypoxia rates were 7.2% and 4.6% in the sedated group compared with 3.2% and 1.2% in the non-sedated group, whereas tachycardia was less frequent in the sedated group at 2.5% compared with 9.2%. Nausea and vomiting occurred at an average 5.6% across the groups, whereas abdominal pain during with the procedure was less in the sedated group, at 5.1% compared with 47.8% among the non-sedated group.

Patients were followed up by telephone interview 30 days after their colonoscopy. Abdominal pain occurred on average across 7.2% of those responding, although was lower at 1.9% among the sedated than the 29.7% among the non-sedated. Abdominal distension and bloating was also relatively common at 4.9%. Rectal bleeding occurred among 2.4% of patients. Baudet et al.[142] conclude that minor complications of colonoscopy are reasonably common. But it is unclear whether the reported events at 30 days were necessarily due to the colonoscopy or could also be linked to the condition under investigation.

Also of note is that sedation reduced the frequency of some adverse events during colonoscopy, including pain and discomfort, and allowed more extensive investigations, such as intubation of the caecum.

De Jonge et al.[143] followed up 1144 Dutch colonoscopy patients by telephone interview. Major events were defined as those requiring hospital intervention. For the major events that were definitely procedure related in the 30 days' follow-up period, 0.36% required hospitalisation owing to rectal bleeding, whereas 0.18% required hospitalisation owing to abdominal discomfort, whereas dizziness, perforation and angina pectoris each occurred in an additional 0.09% of patients. Only 3% of patients had major events. However, 41% had minor adverse events. Those that were definitely procedure related in the 30 days' follow-up period were abdominal discomfort in 17% of patients, rectal blood loss in 5.6%, and a change in bowel habit in 5.4%.

Dominitz et al.[144] undertook a TTO study to investigate the amount of survival people would be willing to sacrifice to avoid 5-yearly screening with sigmoidoscopy or colonoscopy. Four patient groups were involved: those with no experience of screening, those undergoing screening with sigmoidoscopy, those undergoing screening with colonoscopy, and those with colorectal cancer. Those with no experience of screening were willing to trade off reasonable median amounts of time to avoid sigmoidoscopy and more time to avoid colonoscopy. Those screened with sigmoidoscopy were not willing to trade off any time to avoid sigmoidoscopy when measured at the median, but were willing to trade off some time to avoid colonoscopy. Those screened with colonoscopy and those with colorectal cancer were not willing to trade off any time to avoid sigmoidoscopy or to avoid colonoscopy when measured at the median. Although there might be a degree of patient choice among those being screened by sigmoidoscopy and among those being screened by colonoscopy, the results would seem to suggest that the anticipation of the procedures may be worse than the reality.

Niv et al.[145] assessed QoL using the SF-36 both pre, immediately post and 30 days after colonoscopy among 100 Israeli patients. There were no significant changes before and immediately after the colonoscopy in any of the SF-36 parameters, with all of the scores having similar scores pre and post procedure. Similarly, scores were also similar at the 1-month point, although there was a decrease noted in the physical functioning score. This applied among the non-IBD patients and not among the IBD patients, which might be suggestive of it being condition related rather than being procedure related.

Spiegel et al.[146] retrospectively evaluated 458 US patients with IBS using the SF-36, to examine whether having had a previous colonoscopy affected QoL. Controlling for potential confounding variables, Spiegel et al.[146] found no relationship between having had a colonoscopy and QoL. They conclude that there was no evidence that the reassurance provided by a negative colonoscopy improved the QoL of IBS patients.

Warren *et al.*[147] undertook a retrospective analysis of a random sample of 5% of Medicare beneficiaries aged 66–95 years who had undergone a colonoscopy (*n* = 53,220), matching these with controls in order to estimate whether colonoscopy raised event rates within 30 days of the colonoscopy. Patients were matched by date of birth, race, sex, state and a comorbidity score. Adjusting for covariates, they found that diagnostic colonoscopies were associated with a 0.42% risk of a serious GI event compared with 0.18% for those with no colonoscopy, an 8.9% risk of other GI events compared with 5.7% for those with no colonoscopy, but the same risks of cardiovascular events. But it remains unclear to what extent the patient matching would have controlled for the patient group having colonoscopies being inherently more likely to have GI conditions that would themselves lead to other GI events, without these being necessarily related to the colonoscopy.

Levin *et al.*[148] undertook a retrospective analysis of the medical records of 16,318 patients who had undergone a colonoscopy between 1 January 1994 and 16 July 2002 within the Kaiser Permanente health-care system of Northern California to determine rates of serious complications within 30 days of the procedure. Patients were eligible for inclusion if they were older than 40 years of age. Among the 5235 procedures carried out without a biopsy, none resulted in a serious bleed but three resulted in a perforation. Among the 11,083 procedures carried out without a biopsy 53 resulted in a serious bleed and 12 resulted in a perforation.

The economic modelling for the NICE CG118 (colorectal cancer – screening with colonoscopy) assumes that people on surveillance have no complications caused by colonoscopy, such as perforations or bleeding.[149] This is probably due to the rarity of these events.

In contrast, the School of Health and Related Research (ScHARR) Report to the English Bowel Cancer Screening Working Group estimates rates of bleeds and of perforation for colonoscopy and flexible sigmoidoscopy, and the mortality rates associated with the perforations.[150]

The UK flexible sigmoidoscopy trial reported 12 patients being admitted for bleeding following screening among the 40,764 people screened using it: a rate of 0.0295%.[151] A total of 9 of the 2051 patients undergoing colonoscopy with polypectomy were re-admitted to hospital with bleeding: a rate of 0.4390%.

The UK flexible sigmoidoscopy trial, as reported in Atkin *et al.*[152] apparently suggested only one perforation among the 40,764 people screened using flexible sigmoidoscopy. For colonoscopy with polypectomy, Atkin *et al.*[152] reported four perforations out of 2377 colonoscopies performed: a rate of 0.168%. The ScHARR report halved this rate for colonoscopies without polypectomy.

The probability of dying following a perforation was drawn from the study by Gatto *et al.*,[153] which randomly sampled 5% of Medicare beneficiaries within certain regions of the USA (*Table 42*). These figures need to be treated with caution, as all patients were over 65 years of age, but from a total of 108 perforations recorded six patients who died: 5.56%.

TABLE 42 Probability of dying following a perforation (from Gatto *et al.*[153])

Event	Probability (%)	
Colonoscopy bleed	0.4390	
Colonoscopy perforation	0.0800 no polypectomy	0.1680 with polypectomy
Colonoscopy mortality given perforation	5.2	
Sigmoidoscopy bleed	0.0295	
Sigmoidoscopy perforation	0.0025	
Sigmoidoscopy mortality given perforation	6.4	

Bleeds were assumed to require one night as an inpatient, whereas treating a perforation was assumed to require major surgery. Based upon 2011–12 NHS reference costs,[154] bleeds could be costed at the non-elective inpatient stay FZ38F: £561 (IQR £339 to £783), whereas the cost most in line with the ScHARR study for perforations appears to be the non-elective inpatient stay FZ77A: £5360 [IQR £3368 to £6390].

Cost–utility modelling

Summary of modelling approach

The modelling required for a full cost–utility modelling exercise is complicated by there being at least three main conditions under consideration – IBS, CD and UC – and a range of other considerations when PIBD is compared with non-IBD because IBS, although still the commonest non-IBD diagnosis, is less common than in adults. Modelling induction and maintenance subsequent to diagnosis for these three conditions is quite involved. As a consequence, an initial consideration of the immediate QoL impacts for time spent as false negatives is presented, which can be considered alongside the costs of the initial test sequences and likely periods of time spent as false negatives. This QoL impact is restricted to the direct detrimental QoL impacts from not being correctly treated and not entering remission, this being limited by the time spent being incorrectly treated prior to representing for testing. EAG expert opinion suggests that 12 weeks is a reasonable base-case assumption for the duration of false negatives being incorrectly treated prior to the possibility of IBD being reconsidered.

But these immediate QoL impacts among false negatives are formally incorrect, as not all IBD patients when diagnosed with either CD or UC will immediately enter remission after treatment and remain in remission thereafter. Moreover, achieving remission and maintaining it is not costless. As a consequence, it appears that there is a requirement for a full cost–utility modelling exercise that takes into account the costs and benefits of induction therapy and maintenance therapy in both CD and UC, bearing in mind the potential problems of false negatives (IBD missed). This is the approach adopted by the EAG.

The modelling for the full cost–utility approach is eased by the modelling of induction and maintenance of remission for CG152 Crohn's Disease: Management in Adults, Children and Young People[108] and the modelling for the draft clinical guideline for UC being available.[109] Both sets of models adopt a similar framework. Induction therapy with the aim of remission but with subsequent induction therapies for those not achieving remission. Those achieving remission enter a maintenance of remission model, most patients being on treatment but a relatively small minority maintaining remission without active therapy. Remission can be lost, however, which leads to a further sequence of induction therapies. Note that the sequences of induction therapies in the initial induction therapy modelling and the sequences of induction therapies among those having lost remission in the maintenance of remission modelling therapy modelling differ, and even where the same therapy is involved it may have different clinical effectiveness estimates. All induction therapy sequences have as their final option surgery. This, in common with the modelling for the clinical guidelines, is assumed to achieve a permanent remission without the requirement for any further therapy, although this may be optimistic given the 10-year time horizon of the modelling.

Within this modelling, in common with the clinical guidelines' modelling, there is no explicit consideration of possible disease progression, such as the development of fistula during the period of time spent being incorrectly treated as false negative or during periods of loss of remission. Were this to apply, it is likely that the relative importance of sensitivity over specificity would increase compared with the current modelling approach. We note that in the study by Shaoul et al.[92] two cases had fistulae at diagnosis.

The above has made little reference to the modelling of relief of symptoms in IBS patients. The approach adopted is broadly in line with that of the YHEC model,[50] informed by the modelling for CG61: Diagnosis and Management of Irritable Bowel Syndrome in Primary Care.[35] This is simpler than the modelling of induction and maintenance in CD and in UC. But, given the assumed 100% specificity of colonoscopy

meaning that there are no false positives at the end of the test sequence, the modelling of IBS and its treatment subsequent to diagnosis is of lesser importance within the overall cost–utility modelling. It is, in effect, a common residual to all comparators. Its main impact is to determine the costs incurred among false negatives being incorrectly treated. Given this, the full cost–utility model can be viewed as both an IBD versus IBS model and a reasonable IBD versus non-IBD model, provided that for the latter the costs among the false negatives are appropriately adjusted to take into account any additional testing and treatments that may occur among the non-IBD patients, noting the lower proportion with IBS in children.

Two scenarios are modelled:

- adult patients in primary care, with test accuracies for IBD compared with IBS
- paediatric patients in secondary care, with test accuracies for IBD compared with non-IBD.

Perspective, time horizon and discounting

The modelling adopts the NICE reference case perspective of patient benefits and NHS and PSS costs, over a 10-year time horizon for the base case, with discounting of costs and benefits at an annual 3.5%.

Model structure

For reasons of space, the cost–utility model is most simply presented as a set of interlinked models:

- the test model
- the induction and maintenance model among true positive patients with CD
- the induction and maintenance model among true positive patients with UC
- the induction and maintenance model among true negative patients with IBS and false-negative patients with IBD.
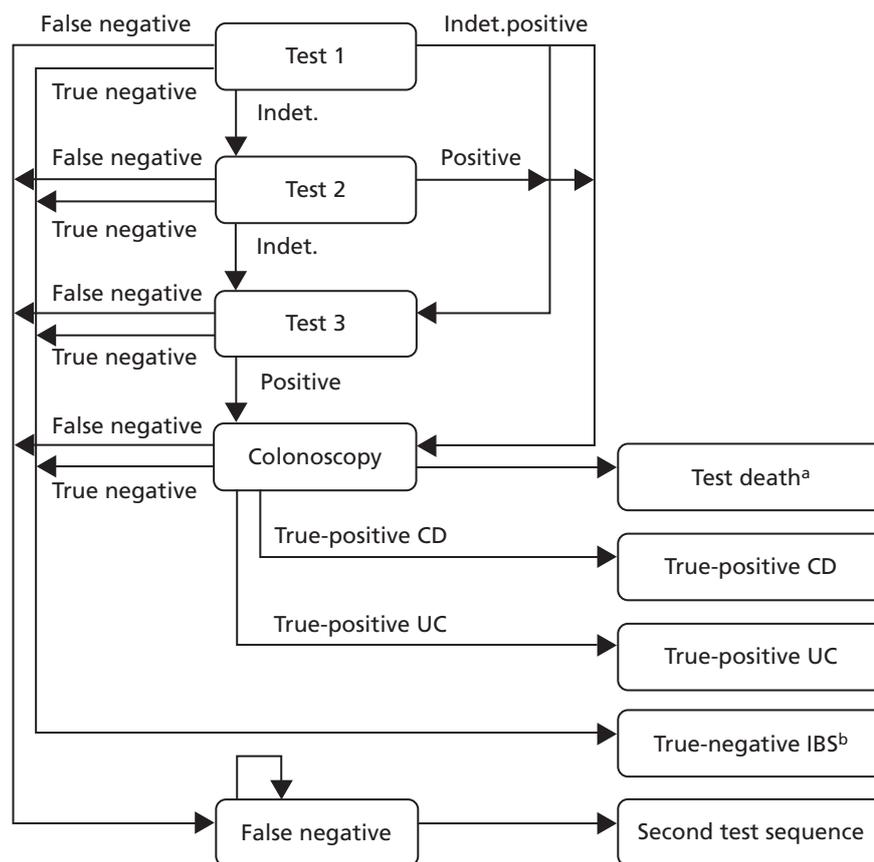
The model structure for the initial testing sequence is shown in *Figure 21*.

Owing to the timing of testing and the possible delays between tests, all of the models use a weekly cycle. The delay between referral and colonoscopy is assumed to be 4 weeks and the time to retesting among those testing negative but not responding to IBS therapy is assumed to be 12 weeks, both estimates being based upon expert opinion. This may be optimistic, as noted in *Chapter 1*, because a sequence of unsuccessful treatments may be pursued for IBS, and so is explored in sensitivity analyses.

The above permits a range of test sequences to be compared. For instance, an initial POCT with a poor specificity could be followed by an ELISA test. FC testing does not have to result in an immediate referral for colonoscopy for all positive results. In a similar vein, the model structure also permits the exploration of rates of indeterminate test results having a follow-up test prior to any referral to colonoscopy. Owing to a lack of data, this latter option has not been formally explored in the analyses that follow, although a consideration of the possible impact of retesting indeterminate results is presented.

The key assumption in all of the above is that all of those who test positive, possible after a sequence of tests, receive a colonoscopy. The current modelling (*Figure 22*) assumes that referrals to secondary care result in colonoscopy. The model structure allows for referral to secondary care to result in assessment by a gastroenterologist, with only a proportion of those referred going on to colonoscopy. But this requires that the sensitivity and specificity of any gastroenterology assessment be estimated. A lack of data means that this option has not been considered. The sensitivity and specificity of an ELISA test could be seen as the closest available proxy for this.

For the initial induction of remission in CD, the most cost-effective strategy within the modelling of CG152[108] was an 8-week course of prednisolone, followed by an 8-week course that adds azathioprine to prednisolone, followed by a 6-week course of anti-TNF. Within this, the more cost-effective anti-TNF was adalimumab, and this is applied within the base-case modelling. This appears to be broadly in line with

FIGURE 21 Model structure of initial test sequences. a, Death after colonoscopy; and b, true negative = IBS.

TA187,[14] although this envisages treatment with an anti-TNF for unresponsive disease for up to 12 months.

For the modelling of the maintenance of remission for CD, this is again based upon the most cost-effective strategy identified within the modelling of CG152.[108] This assumes azathioprine as the maintenance therapy, followed by the same induction sequence as in the initial induction of remission modelling.

In *Figure 23*, for diagrammatical simplicity, remission from inpatient therapy receives azathioprine maintenance therapy. All other patients receive LASA maintenance therapy if on active treatment.

For the modelling of induction of remission for UC, this is based upon the most cost-effective strategies identified within the modelling for the draft UC guideline: strategy 10 of table 34 of appendix L: High-dose ASA (HASA) followed by addition of a topical ASA (aminosalicylic acid), followed by high-dose ASA with beclometasone (HASB), followed by prednisolone.[109] This appears to be broadly in line with the recommendations of the draft clinical guideline. Whether induction of remission would initiate with the HASA of strategy 10 or perhaps the low-dose ASA (LASA) of strategy 6 is a moot point. The net monetary benefits of the two strategies are similar: £8513 and £8323 at a willingness to pay of £20,000 per QALY, but the cost-effectiveness of strategy 10 versus strategy 6 is estimated to be £2818 per QALY with a likelihood of cost-effectiveness of 47% compared with 18% for strategy 6. The current modelling adopts the sequence of strategy 10.

For the modelling of the maintenance of remission for UC, this is again based upon the most cost-effective strategies identified within the modelling for the draft UC guideline: LASA maintenance, followed by LASA maintenance for any patients losing but then regaining remission, as in table 59 of appendix L.
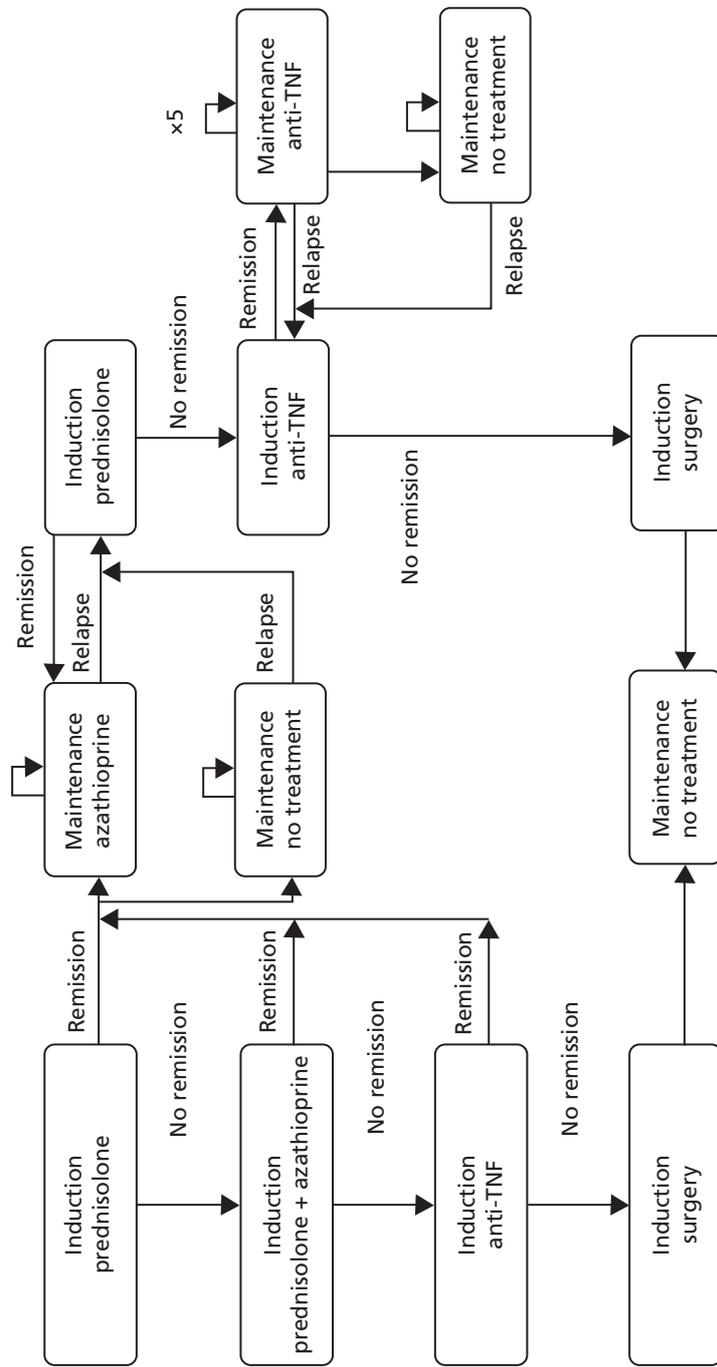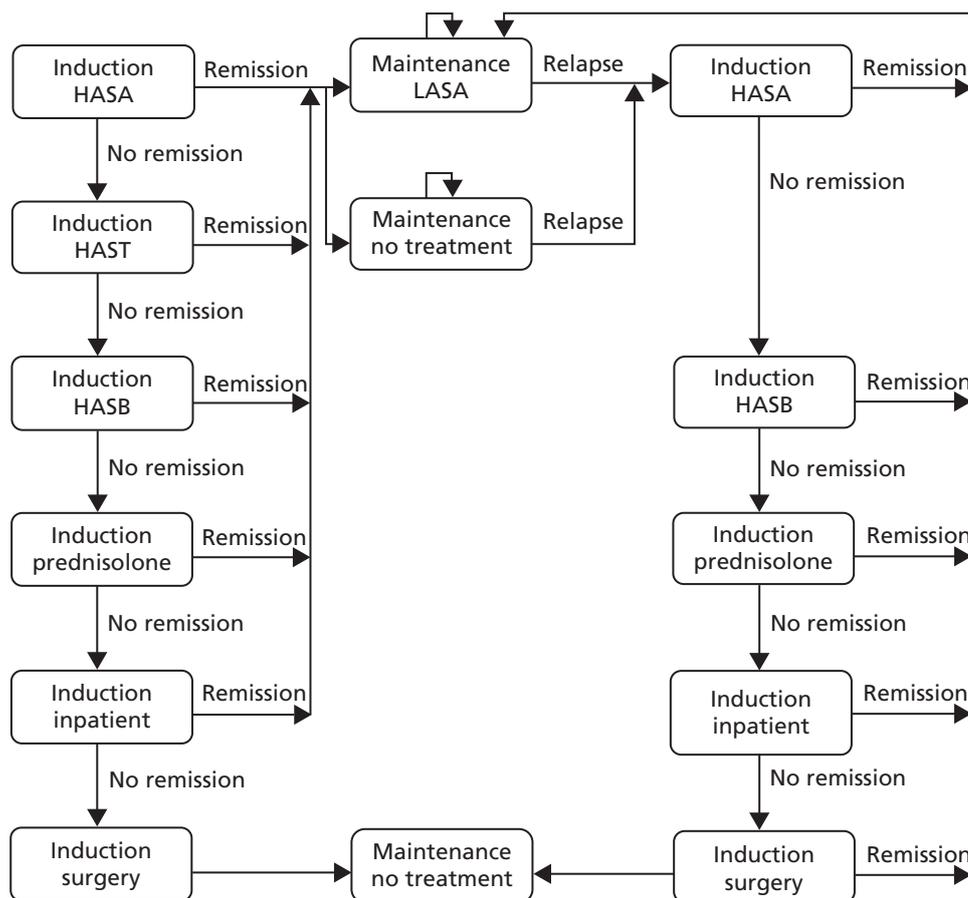
FIGURE 22 Model structure of CD true positives.

FIGURE 23 Model structure of UC true positives.
HASA, high-dose aminosalicylic acid; high-dose aminosalicylic acid with beclometasone; HASB, high-dose aminosalicylic acid with beclometasone; HAST, high-dose ASA with a topical aminosalicylic acid; LASA, low-dose aminosalicylic acid.

The rate of response to the various treatments for IBS affects the total costs, in that those not responding to initial treatment with dietary advice followed by medication are assumed to be referred for a colonoscopy if they have not already had one (*Figure 24*). The economic review and modelling of the IBS clinical guideline suggests a 45% response rate for placebo as drawn from Mearin *et al.*,[155] which for the base case will be taken to be the response rate to dietary advice. This is broadly in line with the 50% assumed in the YHEC report,[50] which was based on expert opinion.

The IBS clinical guideline also outlines a range of medical therapies for IBS, with RRs of response (compared with placebo), ranging from 1.32 for antispasmodic agents to 2.00 for anti-motility agents. As these can frequently be sequenced, this increases the overall response rates for medical therapies, for example sequencing in the antispasmodics modelling alone increases the estimated overall response rate to 78% in the modelling for CG61.[35] In the light of this, the 5% colonoscopy referral rate for those not responding to initial treatment with dietary advice followed by medication of the YHEC report,[50] based upon expert opinion, appears reasonable.

Note that for IBS patients who have already had a colonoscopy it is assumed that the colonoscopy is not repeated. There is some inconsistency of approach in this, in that the small proportion of false-negative patients who have previously had a colonoscopy are assumed to receive a colonoscopy. This can be justified upon grounds of presentation subsequent to incorrect treatment for IBS, but is not without objection.
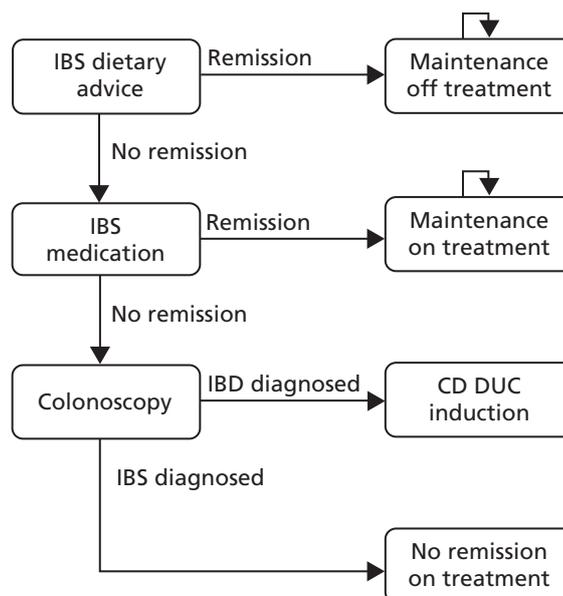
FIGURE 24 Model structure of IBS true negatives and IBD false negatives.

The above has been applied in the modelling of IBD versus non-IBD diagnosis, the implicit assumption being that similar delay to representation and costs are incurred among false negatives. This may not be the case, and given the other conditions than IBS within non-IBD it may be that the average weekly cost for false negatives in the IBD versus non-IBD modelling should be higher than those for the IBD versus IBS modelling.

For the costs of IBS, Akehurst *et al.*[7] estimated a net additional annual cost of £188 [£123] compared with a control group of patients. This suggests a weekly average of £9.29. For patients receiving medication for IBS a relatively minor additional £1.41 has been included, based upon a simple average of the cost of generic mebevirine and the cost of the branded mebevirine, Colofac (Abbott Healthcare).

Primary care modelling: base-case test characteristics

The base case considers the cost-effectiveness of GP testing without FC being available; CalDetect at the 15 µg/g cut-off as drawn from Otten *et al.*;[73] and ELISA at the 50 µg/g cut-off as drawn from Figure 6 of the clinical review (*Table 43*). The base case using the lower 15 µg/g cut-off of Otten *et al.*[73] may initially seem surprising, but the data for the 60 µg/g cut-off of Otten *et al.* suggest only a slight gain in terms of a better specificity, 97.8% compared with 94.5%, but significant loss in terms of a worse sensitivity, 60.9% compared with 100.0%.

TABLE 43 Base-case test characteristics

Test	GP no testing	CalDetect	ELISA RoC	Colonoscopy
Cut-off, µg/g	–	15	50	–
Sensitivity, %	100.0	100.0	93.0	95.0
Specificity, %	78.8	94.5	94.0	100.0
Cost, £	Nil additional	24.03	22.79	741.68[a]

a Weighted average of NHS reference cost outpatient and day case FZ51Z without biopsy, or FZ52Z with biopsy. Base case assumes 100% colonoscopy with biopsy. Sigmoidoscopy, where included, is costed as the weighted average of outpatient and day case FZ54Z without biopsy, or FZ55Z with biopsy. Also includes the cost of a gastroenterology outpatient appointment.

The costs of Caldetect are based upon the list price for KST11005 PreventID CalDetect on the AlphaLabs website, coupled with 15 minutes of GP practice nurse time. ELISA testing is based upon an assumption of 40 patient samples per 96-well plate, costed at the list price on the AlphaLabs website, coupled with an average of 11–12 minutes of staff time at grade 6/7 of the NHS terms and conditions of service handbook: Annex C table 12. These staff costs have been proportionately increased for oncosts in line with the oncosts estimated for a hospital pharmacist within the PSSRU 2011 *Unit Costs of Health and Social Care*.[107] (Note that for the later scenario analyses, a £28.27 cost of Quantum Blue is based upon the list price for Quantum Blue on the AlphaLabs website, the list price for an extraction kit on the Biohit website, coupled with an average of 12–13 minutes of staff time at grade 6/7. Staff times for sample preparation for Quantum Blue and ELISA have been equalised.)

The resulting staff costs are very similar between all FC tests, and the main cost differences are due to the publicly available prices for the consumables. During this assessment it has been consistently noted by both suppliers and those using the tests within the NHS that most if not all tests are sold at a discount to the publicly available prices. These discounts will differ between tests, and possibly even between suppliers of a given test. There will be further geographic variation. Probably the best that the EAG can do is to consider a range of hypothetical discounts common across all tests, but, given the base-case results that follow, there is little requirement for this.

The accuracy of colonoscopy is drawn from expert opinion. We have assumed that colonoscopy is only 95% sensitive, for reasons including failure to intubate the terminal ileum, CD being restricted to parts of the small bowel not accessible by the colonoscope, technical problems, failure of the patient to tolerate the procedure, and inadequate bowel preparation.

Note that in *Table 43* the colonoscopy cost includes the cost of an outpatient gastroenterology visit at a cost of £164 (NHS reference cost: 301 Gastroenterology Consultant led First Face to face Non Admitted) and the costs of adverse events. Given the rarity of bleeds and perforations, despite the large cost associated with perforation these add very little to the overall costs of investigation: only around £12 to the cost of each colonoscopy.

Secondary care modelling: base-case test characteristics

The base case considers the cost-effectiveness ELISA at the 50 µg/g cut-off as drawn from Figure 13 of the clinical review; the cost-effectiveness ELISA at the 100 µg/g cut-off as drawn from Figure 16 of the clinical review; and all patients being referred directly to colonoscopy. This results in the test characteristics for the secondary care modelling shown in *Table 44*.

Characterising uncertainty around the receiver operating characteristic curves

For the probabilistic sensitivity analysis (PSA), a possible and correct approach to characterising uncertainty around the central estimates drawn from the RoC curves is to draw a sufficient number of estimates from iterations of the WinBUGs code (MRC Biostatistics Unit, Cambridge, UK) to characterise the distributions. This has the advantage of correlating the sensitivity and specificity that underlies the RoC curve estimates. But for current purposes this required a large number of iterations for their means to converge to the

TABLE 44 Base-case characteristics: secondary care

	Test		
	ELISA, cut-off (µg/g)		
	50	100	Colonoscopy
Sensitivity, %	99.0	94.0	95.0
Specificity, %	74.0	82.0	100.0

central estimates of the RoC curves. Given the size and complexity of the other modelling this would have led to each PSA requiring a very long time to run.

In light of this, a simpler approach has been adopted. The sensitivity and specificity, or rather their deviation from 100% accuracy, is simulated using the gamma distribution. Over 1000 iterations, this results in the same central estimates as Figures 6, 13 and 16, but slight differences in the upper and lower confidence limits as outlined in *Table 45*. Any discrepancies appear minor, with the possible exception of the lower CI limit for the specificity of figure 6, possibly owing to these data being very heavily skewed compared with the other estimates. But it should be borne in mind that these simulations assume independence between the sensitivity and specificity of each RoC.

Baseline patient characteristics: primary care

For the primary care adult population, the model adopts a baseline age of 25 years for those presenting, as drawn from the CD CG152 modelling,[108] although this may be quite low for IBS patients. In line with the CD CG152 modelling,[108] the female proportion is taken to be 50% for both CD and UC. IBS appears to have a higher proportion of women presenting, the Brazier *et al.* (2004)[117] sample being 86% female, although this estimate may be towards the upper end. The base case adopts a 75% female proportion for IBS. Note that these estimates affect only the all-population mortality risks. As these are low during mid-adulthood for both women and men, the average age and proportion of women inputs have minimal impact upon results.

The base case 6.3% (7/111) prevalence of IBD is drawn from the Durham data, whereas the 60% (539/904) prevalence of UC among IBD patients is drawn from Shivananda *et al.*[29]

Baseline patient characteristics: secondary care

For the secondary care paediatric population, female proportions of 38% (35/91) for IBD patients and for 44% (44/99) non-IBD patients are drawn from Henderson *et al.*[30] An average age of 16 years is assumed, although as for the adult modelling this has minimal impact upon results.

The base case 48% (91/190) prevalence of IBD and the 75% (62/83) prevalence of CD among IBD patients are drawn from Henderson *et al.* (2012).[30]

Health-related quality of life

Although Konig *et al.*[122] provide QoL estimates for both CD and UC using the EQ-5D, their relevance is limited by the German mapping function being used. Similarly, although Casellas *et al.*[126] also provide QoL estimates for both CD and UC using the EQ-5D, their relevance is limited by the Spanish mapping function

TABLE 45 Probabilistic sensitivity analysis-simulated sensitivities and specificities for the ROC curves

ROC curve		Central, %	Lower CI, %		Upper CI, %	
			Actual	Simulated	Actual	Simulated
Figure 6	Sensitivity	93	83	85	97	98
	Specificity	94	73	76	99	100
Figure 13	Sensitivity	99	95	95	100	100
	Specificity	74	59	59	86	85
Figure 16	Sensitivity	94	86	87	98	99
	Specificity	82	67	68	91	92

being used. Stark *et al.*[124] provide QoL estimates for both CD and UC using the EQ-5D and the UK social tariff. Their sample sizes are also reasonably large: 270 with CD and 253 with UC, although there may be some concerns around sample selection given that only a little over one-third of the 1447 originally contacted agreed to participate. Despite this, Stark *et al.*[124] appear to provide the most coherent set of utility values in line with the NICE reference case, the values of interest being 0.890 for remission and 0.610 for active disease for patients with CD, and 0.910 for remission and 0.710 for active disease for patients with UC. These imply utility decrements for active disease compared with remission of 0.280 for patients with CD and 0.200 for UC patients.

Of the CD-specific QoL papers, Gregor *et al.*[133] is of the most interest. Their TTO results suggest QoL values of 0.955 for mild disease, 0.880 for moderate disease and 0.720 for severe disease: decrements from mild disease of 0.075 for moderate disease and 0.235 for severe disease. While the 0.235 decrement for mild to severe disease is in line with the 0.280 estimate of Stark *et al.*, for current purposes the 0.075 decrement for mild to moderate disease might be the relevant estimate, or at least be applicable to a larger proportion of patients. This could suggest a somewhat lesser impact from false negatives being incorrectly treated than occurs with the Stark *et al.*[124] estimates, although note that within the Stark *et al.*[124] patient population only 0.4% of patients with CD had severe disease. CG152 uses the 0.280 decrement of Stark *et al.*[124]

Of the UC-specific QoL papers, the reporting in Bryan *et al.*[137] of the HODaR EQ-5D values of 0.880 for remission and for 0.420 active disease suggest quite a large decrement of 0.460. This appears to be out of line with the other estimates that are available. The draft clinical guideline for UC uses the values of Poole *et al.*[139] of 0.940 for remission and 0.775 for mild to moderate disease, suggesting a decrement of 0.165. This is slightly less than the decrement of 0.200 of Stark *et al.*[124] and would also suggest a lesser impact from false negatives being incorrectly treated than occurs with the Stark *et al.*[124] estimates, although note again that within the Stark *et al.*[124] patient population only 2.2% of UC patients had severe disease.

In light of this, the base case will apply the QoL decrements from remission to active disease of 0.280 for CD and 0.200 for UC of Stark *et al.*[124] But sensitivity analyses applying the QoL decrements from mild to moderate disease of 0.075 for CD as drawn from Gregor *et al.*[133] and of 0.165 as drawn from Poole *et al.*[141] will also be explored.

It can be argued that the QoL data for those with missed IBD who are being incorrectly treated for IBS may differ from that of patients with IBD who are not in remission but are being correctly treated for IBD. But in the absence of QoL estimates for those with missed IBD who are being incorrectly treated for IBS, the best proxies are the QoL estimates for those patients with IBD who are not in remission but are being correctly treated for IBD.

In this context it is important to bear in mind that there is also uncertainty around the average duration that false negatives will be incorrectly treated for IBS before re-presenting and being further investigated due to IBS treatment not inducing remission. EAG expert opinion suggests that an average of 3 months is reasonable, with the main QoL impacts being broadly proportionate to this duration. Given this, the total QALY decrements among false negatives during their period of incorrect treatment for IBS can be presented for the QoL decrements outlined above, coupled with a range of possible durations of incorrect treatment (*Table 46*). For the base-case 3-month duration, the QALY decrement is simply one-quarter of the QoL decrement.

For CD, retaining the estimates of Stark *et al.*[124] and moving from a 2 months' average duration to a 4 months' average duration doubles the overall QALY decrement as would be anticipated. While it can be argued that the sensitivity analysis using the decrement of Gregor *et al.*[133] is more speculative as it is experimental data, applying the decrement of Gregor *et al.* within the context of a 2 months' average duration results in an overall decrement of only 0.013 QALYs compared with 0.093 QALYs when applying

TABLE 46 Quality-adjusted life-year decrements for different utility estimates and durations of false negatives

Condition	Source	Decrement	Total QALY decrement from being a false negative for:					
			1 month	2 months	3 months	4 months	5 months	6 months
CD	Stark et al.[124]	0.280	0.023	0.047	0.070	0.093	0.117	0.140
	Gregor et al.[133]	0.075	0.006	0.013	0.019	0.025	0.031	0.038
UC	Stark et al.[124]	0.200	0.017	0.033	0.050	0.067	0.083	0.100
	Poole et al.[141]	0.165	0.014	0.028	0.041	0.055	0.069	0.083

the decrement of Stark et al.[124] within the context of a 4 months' average duration: over a sevenfold difference.

For UC, applying the decrement of Poole et al.[141] within the context of a 2 months' average duration results in an overall decrement of 0.028 QALYs compared with 0.067 QALYs when applying the decrement of Stark et al.[124] within the context of a 4 months' average duration: between two and three times the amount.

These QALY decrements will be qualified by the prevalence of IBD in the presenting patient population, and the proportion of these who are modelled as being diagnosed as false negatives. For instance, an IBD prevalence of 5% coupled with a sensitivity of 90% results in only 0.5% of the total patient population being diagnosed as false negatives.

The above underlines that however complicated the full cost–utility modelling is, the QALY decrements among false negatives will be dependent upon:

- the source of the QoL values
- the assumed duration of patients remaining as false negatives
- the prevalence of IBD in the presenting population
- the sensitivity of the tests under consideration.

Owing to the low prevalence of IBD in the primary care population and the quite high sensitivities of the various tests, the total QALY decrements among false negatives are likely to be quite small. Results may be mainly driven by the direct upfront test costs, including the costs of colonoscopies. This may also cause the adverse events and associated mortality from colonoscopy to come more to the fore, despite the assumed mortality rate also being very low.

The utility decrements for IBS are less important for current modelling purposes, given the 100% specificity assumed for colonoscopy meaning that there are no false positives by the end of the initial test sequence. For the base case, the 0.071 increment for response to treatment estimated within CG61[35] will be applied. The 0.662 baseline HRQoL that this increment is applied to is taken from Brazier et al.[117] A sensitivity analysis using the EQ-5D values of Spiegel et al.[5] can also be considered; 0.780 for response to treatment and 0.730 for no response to treatment, but recall that the mapping function employed by Spiegel et al.[5] clear. Note that the baseline HRQoL value for IBS will also have an impact owing to the small mortality rate associated with colonoscopy, with this impact enduring for the 10-year time horizon of the model.

Other model inputs
Given the extent of the downstream modelling, the full set of model inputs is presented in *Appendix 7*, coupled with their treatment within the PSA.

Primary care modelling: sensitivity and scenario analyses

The prevalence of IBD within the presenting patient population determines the relative importance of sensitivity and specificity. This is explored through sensitivity analyses that vary this from 5% to 25% in 5% increments.

A scenario analysis that speculates that FC testing might be used in a wider patient group than would be referred in the absence of FC testing is then presented.

Further sensitivity analyses are then presented, which:

- vary the time to representation among false negatives from the base case 12 weeks to 8 weeks, 16 weeks and 24 weeks
- change the source of utility estimates from Stark *et al.*[124] and CG61[35] to Gregor *et al.*,[133] Poole *et al.*[141] and Spiegel *et al.*[5]
- remove the cost of the gastroenterology outpatient appointment from the cost of colonoscopy (this was the approach used in the YHEC 2010 report[50] and is included here to allow comparison with their figures; it also provides a sensitivity analysis around the cost of referral and colonoscopy; usual UK practice would be to refer for a gastroenterology opinion that would lead to colonoscopy, but, in some countries, direct referral to colonoscopy appears to apply)
- vary the assumed non-response to medication among IBS patients from the base case 5% to 0% and 10% (this applies to those in whom dietary advice has failed)
- remove the mortality associated with colonoscopy.

The clinical effectiveness section also presents a range of sources that include estimates of the accuracy of FC testing at various cut-off levels. These are considered within individual scenario analyses, with the different cut-offs being directly compared, although these estimates are not integrated into the primary care base case.

Primary care modelling: other estimates of test characteristics

Additional effectiveness estimates for further analyses are drawn from Otten *et al.*[73] (*Table 47*) for CalDetect at the 60 μg/g cut-off; from Hessells *et al.*[98] for CalDetect at the 15 and 60 μg/g cut-offs, and for Quantum Blue at the 30, 40, 50 and 60 μg/g cut-offs (*Table 48*); and, from Basumani *et al.*[83] for ELISA at the 50, 100 and 150 μg/g cut-offs (*Table 49*). The central estimates from Otten *et al.*[73] are shown in *Table 47*.

In order to characterise the sensitivities and specificities of Hessells *et al.*[98] for the probabilistic modelling, the numbers of true positives, false negatives, true negatives and false positives has to be calculated. Unfortunately, Hessells *et al.*[98] only report the overall sample size, the numbers of correct diagnoses and the sensitivities and specificities. Given this, on the basis of a sample size of 85 the EAG has calculated the number with IBD to be 23, which implies the following numbers of true positives, false negatives, true negatives and false positives. These in turn imply sensitivities and specificities. In the main these correspond with those of Hessells *et al.*,[98] although there is some very minor disagreement of the order of 1% for a few of the percentages.

TABLE 47 Otten *et al.*:[73] CalDetect test accuracy

	CalDetect, cut-off (μg/g)	
	15	60
Sensitivity, %	100.0	60.9
Specificity, %	94.5	97.8

TABLE 48 Hessells *et al.*:[98] POCT accuracy

	Test					
	Quantum Blue, cut-off (µg/g)				CalDetect, cut-off (µg/g)	
	30	40	50	60	15	60
n with	23	23	23	23	23	23
True positive	22	21	20	18	22	20
False negative	1	2	3	5	1	3
Implied sensitivity, %	96	91	87	78	96	87
Hessells *et al.* (2012)[97] table 1, %	96	92	88	79	96	88
n without	62	62	62	62	62	62
True positive	43	52	52	54	33	46
False positive	19	10	10	8	29	16
Implied specificity, %	69	84	84	87	53	74
Hessells *et al.* (2012)[97] table 1, %	69	84	84	87	53	74

TABLE 49 Basumani *et al.*:[83] ELISA test accuracy

	ELISA, cut-off (µg/g)		
	50	100	150
Sensitivity, %	100.0	91.7	83.3
Specificity, %	60.2	81.6	85.7

Secondary care modelling: sensitivity and scenario analyses

Sensitivity analyses are presented which:

- vary the prevalence of IBD within the presenting patient population from the base case 48% to 40% and 60%
- vary the time to re-presentation among false negatives from the base case 12 weeks to 8 weeks and 16 weeks
- change the source of utility estimates from Stark *et al.*[124] and CG61[35] to Gregor *et al.*,[133] Poole *et al.*[141] and Spiegel *et al.*[5]
- doubling the annualised net cost amongst false negatives from £188 to £376
- remove the mortality associated with colonoscopy.

Base-case results: primary care

For the primary care base case, the patient numbers receiving the initial test and being referred for colonoscopy are as shown in *Table 50*.

Note that the above relates to the initial test sequence of, for example, CalDetect 15 µg/g followed by colonoscopy. Within this test sequence, among those with IBD the initial Caldetect test identifies all 6.3% of patients with IBD as true positives. The colonoscopy subsequent to this identifies 6.0% of the 6.3% referred by Caldetect as true positives, owing to its 95% sensitivity. Among those with IBS, the initial Caldetect test identifies 5.1% of the 93.7% of patients with IBD as false positives. These are referred on, with the subsequent colonoscopy identifying all these as true negatives owing to its 100% specificity.

TABLE 50 Primary care: base-case results of initial test sequence

	GP			CalDetect 15 µg/g			ELISA		
	First test	Colonoscopy	Final	First test	Colonoscopy	Final	First test	Colonoscopy	Final
IBD tested, %	6.3	6.3		6.3	6.3		6.3	5.9	
True positive, %	6.3	6.0	6.0	6.3	6.0	6.0	5.9	5.6	5.6
False negative, %	0.0	0.3	0.3	0.0	0.3	0.3	0.4	0.3	0.7
IBS tested, %	93.7	19.8		93.7	5.1		93.7	5.6	
True negative, %	73.9	19.8	93.7	88.5	5.1	93.7	88.1	5.6	93.7
False positive, %	19.8	0.0	0.0	5.1	0.0	0.0	5.6	0.0	0.0

As a consequence, although the initial test referred on a proportion of false positives, these are all eliminated by the colonoscopy and at the end of the test sequence there are no false positives.

Immediately apparent from the above is that GP opinion (without calprotectin) results in a somewhat larger number of false positives being referred for unnecessary colonoscopies: 19.8% of the total patient population or 21.2% of those with IBS, as would be anticipated from the 78.8% specificity. CalDetect 15 µg/g is somewhat better: 5.1% of the total patient population or 5.4% of those with IBS, as would be anticipated from the 94.5% specificity. The ELISA test, although perhaps marginally cheaper than the CalDetect test, is estimated to have an inferior sensitivity and a very slightly inferior specificity. The proportion correctly referred to colonoscopy is lower than for CalDetect, although the proportion incorrectly referred to colonoscopy is slightly higher.

Given the above, when coupled with representations for testing among false negatives and IBS patients not responding to IBS therapy who have not previously been scoped, this results in the following test costs, with the other costs from downstream modelling of treatment for induction and maintenance of remission yielding the total estimated costs for the cost–utility modelling.

Within *Table 51*, in part due to the quite low base-case prevalence assumed for IBD within the presenting population, the average QALYs and downstream costs of treatment are broadly in line between the three comparators. There are very slight differences between the comparators' QALYs, with very slight gains from CalDetect over ELISA, and larger, although still slight, QALY gains over the GP with no FC testing. But the main differences are in the upfront average test costs, with CalDetect and ELISA having similar test costs, both of which are somewhat less than those of the GP in the absence of FC testing due to their superior specificity.

The central estimates and cost-effectiveness acceptability frontiers (CEAFs) from the probabilistic modelling run over 1000 iterations are shown in *Table 52* and *Figure 25*, respectively. Within this, it should be borne in mind that the prevalence of IBD is also treated as being probabilistic within the PSA.

In the above, the probability of being cost-effective for the comparator of the GP without FC testing never rises above the horizontal axis, i.e. it is estimated that there is no probability of GP without calprotectin testing being cost-effective compared with calprotectin testing.

As for the deterministic modelling, the probabilistic model central estimates suggest small QALY gains from FC testing coupled with cost savings compared with GP referrals in the absence of FC testing. There are very minor differences between CalDetect and ELISA in terms of the central estimates for costs and

TABLE 51 Primary care: base-case results

Comparators	QALYs	Costs (£)		
		Tests	Other	Total
GP no FC				
CD	0.1832	22	493	515
UC	0.2771	32	144	176
IBS	5.7682	202	2404	2606
Total	6.2285	257	3041	3297
POCT: CalDetect 15 µg/g				
CD	0.1832	23	493	516
UC	0.2771	33	144	177
IBS	5.7691	114	2408	2522
Total	6.2293	170	3044	3214
ELISA				
CD	0.1831	23	492	515
UC	0.2770	34	143	177
IBS	5.7690	116	2407	2524
Total	6.2291	173	3042	3215

TABLE 52 Primary care: probabilistic modelling central estimates

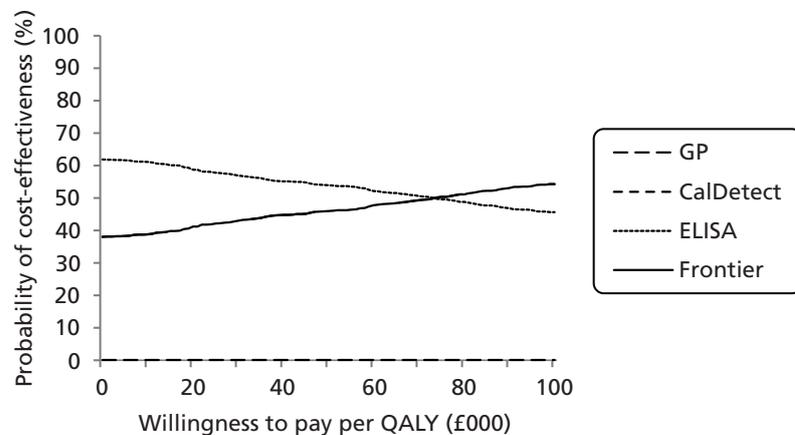| Test	Base case	
	QALYs	Costs (£)
GP	6.2637	3312
POCT	6.2646	3230
ELISA	6.2643	3230

FIGURE 25 Cost-effectiveness acceptability frontiers: primary care – base case.

QALYs. Interestingly, the CEAFs suggest that ELISA is the most likely to be cost-effective, although, again, the differences between the two tests are not marked, but that CalDetect has the highest expected net monetary benefit at all but very low willingness-to-pay values. Further sensitivity analyses that increase the prevalence of IBD suggest that CalDetect will remain at the frontier for all but very low willingness-to-pay values, but, that as the prevalence of IBD rises, the likelihood of CalDetect being the most likely to be cost-effective increases; i.e. the crossover point moves towards lower willingness-to-pay values. But perhaps not too much should be read into this given the similarity of the central estimates, and that a few outlier iterations could have a more marked effect than usual.

Sensitivity analyses: primary care: prevalence of inflammatory bowel disease

Varying the prevalence of IBD in the presenting patient population between 5% and 25% provides the results shown in *Table 53*.

As in the base-case modelling, the GP with no FC testing is estimated to have the smallest overall QALYs and also to cost more than the other comparators for IBD prevalences up to 25%. At an IBD prevalence of 25%, the GP with no FC testing is still estimated to result in QALY losses compared with CalDatect, but in very slight patient gains compared with ELISA owing to the less-than-perfect 93.0% sensitivity of ELISA. The average cost savings from ELISA also fall from £83 at a 5% prevalence to £63 at a 25% prevalence, as there are lower cost offsets from avoiding fewer incorrect referrals for colonoscopy. But, owing to the very limited differences in QALY estimates and the additional average £63 cost, at a prevalence of 25% the cost-effectiveness of the GP without FC testing compared with ELISA is estimated to be £378,000 per QALY.

For CalDetect an increase in the prevalence of IBD increases the net QALY gain over ELISA, this mainly being due to fewer false negatives incorrectly receiving treatment for IBS. This also slightly increases the costs of CalDetect, to the extent that if the prevalence of IBD is $\geq 20\%$ it is no longer cost-saving compared with ELISA. At a prevalence of 25% the incremental cost-effectiveness ratio (ICER) for CalDetect compared with ELISA is estimated to be £1697 per QALY but the differences in terms of the net costs and net QALYs between CalDetect and ELISA remain slight.

Sensitivity analyses: primary care: population tested and general practitioner sensitivity and specificity

As reviewed in the clinical effectiveness section, if FC is made available in primary care the patient group being tested may widen beyond that which GPs previously considered for referral. If patient numbers being tested with FC are double those who would previously have been seriously considered for referral in the absence of FC testing being available, this will affect specificity of the scenario of GP referral in the absence of FC testing. Assuming a doubling of the patient group for FC testing compared with that previously seriously considered for referral, coupled with a small percentage of additional IBD patients within the additional patient group, could suggest a prevalence of only 3.4% among those being tested

TABLE 53 Sensitivity analyses: primary care – prevalence of IBD

IBD prevalence (%)	QALYs			Costs (£)		
	GP	**CalDetect**	**ELISA**	**GP**	**CalDetect**	**ELISA**
5	6.2135	6.2144	6.2142	3190	3106	3107
10	6.2706	6.2715	6.2711	3599	3520	3521
15	6.3277	6.3285	6.3280	4008	3935	3935
20	6.3848	6.3856	6.3849	4416	4349	4348
25	6.4419	6.4426	6.4418	4825	4763	4762

with FC, and a sensitivity and specificity for GP referral in the absence of FC testing of 94.1% and 89.7% respectively in this wider patient group. This results as shown in *Tables 54* and *55*.

With the following results from the full cost–utility modelling.

Faecal calprotectin testing within the wider patient population increases the absolute number of false positives, whereas by construction it has been assumed to remain constant for those assessed by the GP in the absence of FC testing. This tends to reduce the difference between the costs of the test sequences between calprotectin testing and GP assessment in the absence of FC testing. But despite the assumed doubling in the size of the patient group from those who would seriously be considered for *referral* by the GP in the absence of FC testing to those who would be *tested* were FC made available to GPs, the introduction of FC testing is still estimated to be cheaper or at worst broadly cost neutral

TABLE 54 Primary care: alternative presenting population sensitivity analyses test results

	GP			CalDetect, 15 µg/g			ELISA		
	First test (%)	Colonoscopy (%)	Final (%)	First test (%)	Colonoscopy (%)	Final (%)	First test (%)	Colonoscopy (%)	Final (%)
IBD tested	3.4	3.2		3.4	3.4		3.4	3.1	
True positive	3.2	3.0	3.0	3.4	3.2	3.2	3.1	3.0	3.0
False negative	0.2	0.2	0.4	0.0	0.2	0.2	0.2	0.2	0.4
IBS tested	96.6	9.9		96.6	5.3		96.6	5.8	
True negative	86.7	9.9	96.6	91.3	5.3	96.6	90.8	5.8	96.6
False positive	9.9	0.0	0.0	5.3	0.0	0.0	5.8	0.0	0.0

TABLE 55 Primary care: alternative presenting population sensitivity analyses test results – cost–utility

		Costs (£)		
Comparators	QALYs	Tests	Other	Total
GP no FC				
CD	0.0973	12	261	273
UC	0.1472	17	76	94
IBS	5.9507	129	2482	2612
Total	6.1952	158	2820	2978
POCT: CalDetect 15 µg/g				
CD	0.0973	12	262	274
UC	0.1473	18	76	94
IBS	5.9510	118	2484	2601
Total	6.1956	147	2822	2969
ELISA				
CD	0.0973	12	261	273
UC	0.1472	18	76	94
IBS	5.9510	120	2483	2604
Total	6.1955	150	2821	2971

compared with the previous situation of FC testing not being available. Small QALY gains still accrue from FC testing as well.

Sensitivity analyses: primary care – test cut-offs

For the cut-offs for CalDetect reported within Otten et al.[73] the results shown in *Table 56* are estimated.

The slightly better specificity of the 60 μg/g cut-off results in slight cost savings of £27 compared with the 15 μg/g cut-off, but gains of 0.0012 QALYs are anticipated from the 15 μg/g cut-off. This suggests a cost-effectiveness of £23,635 per QALY for the 15 μg/g cut-off which could be seen as borderline cost-effectiveness. Whether the 60 μg/g cut-off with an estimated sensitivity of only 61% would be acceptable in practice is a doubtful. Note also the ICER is almost exactly inversely proportionate to the prevalence if IBD in the presenting population, i.e. if it doubles then the ICER halves.

For the cut-offs for CalDetect reported within Hessells et al.[98] the results shown in *Table 57* are estimated.

The results from the estimates of Hessells et al.[98] for CalDetect are almost the opposite of those estimated using those of Otten et al.[73] CalDetect with a cut-off of 60 μg/g is estimated to dominate CalDetect with a cut off of 15 μg/g. The specificity is notably better for 60 μg/g at 74.2% compared with 53.2% for 15 μg/g, and the sensitivity, although worse at 87.0%, is still somewhat closer to the 95.7% for 15 μg/g than the corresponding figures within Otten et al.[73] But it should be borne in mind that Hessells et al.[98] did not fit the eligibility criteria of the clinical effectiveness review.

For the cut-offs for Quantum Blue reported within Hessells et al.[98] the results shown in *Table 58* are estimated.

TABLE 56 Sensitivity analysis: primary care – Otten et al.,[73] CalDetect cut-offs

Cut-off (μg/g)	QALYs	Costs (£)
15	6.2293	3214
60	6.2281	3187

TABLE 57 Sensitivity analysis: primary care – Hessells et al.,[98] CalDetect cut-offs

Cut-off (μg/g)	QALYs	Costs (£)
15	6.2269	3496
60	6.2278	3352

TABLE 58 Sensitivity analysis: primary care – Hessells et al.,[98] Quantum Blue cut-offs

Cut-off (μg/g)	QALYs	Costs (£)
30	6.2278	3390
40	6.2285	3290
50	6.2283	3290
60	6.2282	3267

As would be anticipated, the 50 µg/g cut-off is dominated, as is the 30 µg/g cut-off, by the 40 µg/g cut-off. The 40 µg/g cut-off is estimated to cost an additional £24 on average compared with the 60 µg/g cut-off, whereas small QALY gains of 0.0003 QALYs suggest an ICER of around £87,000. But these QALY differences are extremely minor and the ICER will swing possibly quite wildly if the underlying inputs and assumptions are changed. Note also that the better specificity of the 60 µg/g cut-off could have resulted in some further QALY gain from avoidance of the minor adverse effects of colonoscopy had these been included in the modelling. But the better sensitivity of the 40 µg/g cut-off would see increases in the prevalence of IBD increase its net QALYs further, the converse being true for a lower IBD prevalence.

For the cut-offs for ELISA reported within Basumani et al.[83] the results shown in *Table 59* are estimated.

The 100 µg/g cut-off is only slightly inferior to the 150 µg/g cut-off in terms of specificity, but better in terms of sensitivity. This leads to a very slight QALY gain, but an additional £30 cost. The poor specificity of the 50 µg/g cut-off leads to it being dominated by the 100 µg/g cut-off.

Other sensitivity analyses: primary care
The univariate sensitivity analyses results in the estimates shown in *Table 60*.

TABLE 59 Sensitivity analysis: primary care – Basumani et al.,[83] ELISA cut-offs

Cut-off (µg/g)	QALYs	Costs (£)
50	6.2274	3448
100	6.2284	3300
150	6.2283	3271

TABLE 60 Primary care: univariate sensitivity analyses

	QALYs					Costs (£)				
	GP	POCT	ELISA	Net		GP	POCT	ELISA	Net	
Analysis	S1	S2	S3	S2 – S1	S2 – S3	S1	S2	S3	S2 – S1	S2 – S3
Base case	6.2285	6.2293	6.2291	0.0009	0.0002	3297	3214	3215	−83.17	−1.48
8-week re-presentation	6.2312	6.2320	6.2319	0.0009	0.0002	3304	3218	3220	−86.21	−2.08
16-week re-presentation	6.2258	6.2266	6.2263	0.0009	0.0003	3274	3191	3192	−83.27	−1.38
24-week re-presentation	6.2204	6.2212	6.2208	0.0009	0.0005	3229	3146	3147	−83.39	−0.77
Utilities, non-Stark	6.6371	6.6377	6.6376	0.0006	0.0001	3297	3214	3215	−83.17	−1.48
No outpatient	6.2285	6.2293	6.2291	0.0009	0.0002	3251	3191	3192	−59.88	−0.73
IBS non-responder rate 0%	6.2437	6.2445	6.2443	0.0009	0.0002	3281	3195	3196	−86.37	−1.58
IBS non-responder rate 10%	6.2133	6.2141	6.2139	0.0008	0.0002	3313	3233	3235	−79.96	−1.37
No colonoscopy mortality	6.2286	6.2294	6.2292	0.0008	0.0002	3297	3214	3216	−83.20	−1.48

The changes appear to broadly affect the three comparators in a like manner, such that although the estimates of costs and QALYs change there is only a limited impact upon net costs and net QALYs. FC testing remains cost saving compared with no-FC testing, and confers some small additional patient benefits. The costs of the POCT CalDetect FC testing and ELISA FC testing remain very similar throughout, with very slight patient gains from POCT CalDetect FC testing being estimated.

Scenario analyses: outpatient appointment sensitivity and specificity

As previously noted, the base case assumes that all those referred receive both an outpatient appointment and a colonoscopy. Owing to a lack of data, the possibility that the outpatient appointment results in some of those referred not receiving a colonoscopy cannot be given one formal estimate but it can be explored through additional scenario analyses.

For these scenario analyses, given the limited importance of test sensitivity to the net costs for the base case, in order to avoid too many permutations a 100% sensitivity is assumed for the consultant's clinical assessment at the outpatient clinic (outpatient assessment). The specificity of this outpatient assessment can then be explored using values of 25%, 50%, 75% and 95%. As the specificity of the outpatient assessment rises so the importance of the specificity of the initial test falls. If the outpatient assessment has a very high specificity such as 95%, a good specificity for the initial test still avoids the costs of unnecessary referrals and outpatient assessments but it is no longer instrumental in avoiding the costs of unnecessary colonoscopies.

These analyses have been undertaken for two scenarios: the base-case referral scenario with an IBD prevalence of 6.3% and a resulting 100% sensitivity and 79% specificity and 24% PPV for GP nous without calprotectin testing, and the testing scenario analysis of a larger tested population (50% of those with symptoms) with an IBD prevalence of 3.4% and a resulting 94% sensitivity and 90% specificity for GP nous without calprotectin testing as shown in *Table 61*.

The key to this is that it assumes that there is a sequence of independent 'tests', for example the GP clinical assessment ('GP nous') followed by the outpatient assessment, or CalDetect followed by outpatient assessment. The sensitivity and specificity of the first test in the sequence are as per the table above. The outpatient assessment specificity is applied subsequent to this. For instance, GP nous refers 100% of those with IBD and 21.2% (100% – 78.8%) of those with IBS. By assumption, the outpatient assessment passes all the 100% of the referred IBD patient on to colonoscopy. But for the scenario of a 95% specificity for the outpatient assessment, only 5% (100%–95%) of the 21.2% of IBS patients who were referred for outpatient assessment are passed on to colonoscopy; i.e. a total of 1.06% = (100% – 95%) × (100% – 78.8%) of those with IBS among the whole population assessed by the

TABLE 61 Test accuracies and outpatient assessment

	Source							
	Durham Dales		**Otten**[73]	**Figure 6**	**Basumani**[83]			**Assumption**
			CalDetect, cut-off (µg/g)	**ELISA cut-off (µg/g)**	**ELISA cut-off (µg/g)**			
	GP nous		**15**	**50**	**50**	**100**	**150**	**Outpatient assessment**
Sensitivity, %	100.0	94.1	100.0	93.0	100.0	91.7	83.3	100
Specificity, %	78.8	89.7	94.5	94.0	60.2	81.6	85.7	variable

Outpatient assessment relates to the consultant's clinical assessment at the outpatient clinic.

GP. The parallel calculation for CalDetect would be 0.275% = (100% – 95%) × (100% – 94.5%) of those with IBS among the whole population assessed by the GP. (Note that in terms of the model structure the additional 'test' of the outpatient assessment also slightly alters the timing of tests within the sequence. In itself, this has minimal impact upon results.)

Note that for the options involving calprotectin testing, the above assumes that the specificity of the outpatient assessment does not vary by calprotectin level. It may be reasonable to assume that someone with IBS and calprotectin level 125 µg/g is more likely to be further investigated that someone with calprotectin of 55. That is, the outpatient assessment specificity may fall as calprotectin rises.

For the base case, this results as shown in *Table 62*.

The total costs reported in the above include not only the immediate test, referral and colonoscopy costs, but also the downstream costs of induction of remission and maintenance of remission therapy. However, the net costs are mainly composed of the net costs of the immediate test, referral and colonoscopy. The net QALY differences between the comparators are relatively minor.

Calprotectin testing is estimated to remain cost saving, although the cost savings dwindle as the specificity of the outpatient assessment improves. This is for the reasons outlined above: a high test specificity avoids referrals for outpatient assessment, but with a high specificity for outpatient assessment the high test specificity is no longer instrumental in avoiding unnecessary colonoscopies.

As the specificity of the outpatient assessment improves, ELISA becomes marginally cheaper than CalDetect owing to its slightly lower staff and consumables cost. This aspect should be treated with some

TABLE 62 Base case and outpatient assessment specificity

Outpatient specificity	Cost	vs. 1	vs. 2	QALY	vs. 1	vs. 2
25%						
CalDetect 15 µg/g	£3207			6.2294		
ELISA figure 6	£3208	£1		6.2291	− 0.0002	
GP nous	£3275	£68	£67	6.2287	− 0.0007	− 0.0005
50%						
CalDetect 15 µg/g	£3199			6.2294		
ELISA figure 6	£3200	£0		6.2292	− 0.0002	
GP nous	£3246	£47	£47	6.2289	− 0.0005	− 0.0003
75%						
ELISA figure 6	£3191			6.2293		
CalDetect 15 µg/g	£3192	£1		6.2295	0.0002	
GP nous	£3218	£26	£26	6.2291	− 0.0001	− 0.0004
95%						
ELISA figure 6	£3185			6.2293		
CalDetect 15 µg/g	£3186	£1		6.2295	0.0002	
GP nous	£3195	£10	£9	6.2293	0.0000	− 0.0002

caution in the light of the staffing time estimates being drawn from different expert opinion and publicly available list prices being used, when all tests are apparently sold at some discount to these.

In the alternative presenting population scenario, it is assumed that calprotectin testing will be carried out in double the number of patients previously considered for referral in the absence of calprotectin testing. This results as follows in *Table 63*.

The relatively small cost savings from calprotectin testing of the original analysis are reversed, as the specificity of the outpatient assessment rises above 50%. The QALY differences are relatively insignificant.

But *Table 63* illustrates that results are sensitive to any assumed increase in the proportion of patients GPs may choose to test with calprotectin if it becomes available.

The Basumani data[83] permit a comparison of the different ELISA cut-offs. But note that the EAG cost-effectiveness estimates using the Basumani data[83] suggested that the 50 µg cut-off was dominated by the 100 µg cut-off owing to the poor specificity of the 50 µg cut-off. The specificities for the three Basumani data[83] cut-offs – 50 µg, 100 µg and 150 µg – are also somewhat worse than that drawn from figure 6 for the pooled 50 µg cut-off. Applying these within the base-case results as follows in *Table 64*.

For the comparison between ELISA with the 150 µg cut-off and ELISA with the 100 µg cut-off, the slight net additional costs from ELISA with the 100 µg cut-off fall slightly as the outpatient assessment specificity improves, this lessening the importance of the better specificity of ELISA with the 150 µg cut-off. Very small net patient gains are estimated for ELISA with the 100 µg cut-off compared with ELISA with the 150 µg cut-off, although these are not sufficient to render it cost-effective at conventional thresholds.

For the scenario of the alternative presenting population, associating the outpatient assessment with a specificity results as shown in *Table 65*.

TABLE 63 Alternative presenting population and outpatient assessment specificity

Outpatient specificity	Costs	vs. 1	vs. 2	QALY	vs. 1	vs. 2
25%						
CalDetect 15 µg	£2961			6.1957		
ELISA figure 6	£2962	£1		6.1955	− 0.0001	
GP nous	£2965	£4	£3	6.1954	− 0.0003	− 0.0002
50%						
GP nous	£2951			6.1955		
CalDetect 15 µg	£2953	£2		6.1957	0.0003	
ELISA figure 6	£2954	£3	£1	6.1956	0.0001	− 0.0001
75%						
GP nous	£2936			6.1956		
ELISA figure 6	£2945	£9		6.1957	0.0001	
CalDetect 15 µg	£2945	£9	£0	6.1958	0.0002	0.0001
95%						
GP nous	£2925			6.1957		
ELISA figure 6	£2938	£14		6.1957	0.0000	
CalDetect 15 µg	£2939	£14	£1	6.1958	0.0002	0.0001

TABLE 64 Base-case, Basumani data[83] and outpatient assessment specificity

Outpatient specificity 25%	Costs	vs. 1	vs. 2	QALY	vs. 1	vs. 2
ELISA 150 μg	£3255			6.2285		
ELISA 100 μg	£3280	£25		6.2286	0.0001	
ELISA 50 μg	£3407	£152	£152	6.2279	− 0.0006	− 0.0006
Outpatient specificity 50%	**Costs**	**vs. 1**	**vs. 2**	**QALY**	**vs. 1**	**vs. 2**
ELISA 150 μg	£3236			6.2286		
ELISA 100 μg	£3255	£20		6.2288	0.0001	
ELISA 50 μg	£3353	£118	£118	6.2283	− 0.0003	− 0.0003
Outpatient specificity 75%	**Costs**	**vs. 1**	**vs. 2**	**QALY**	**vs. 1**	**vs. 2**
ELISA 150 μg	£3216			6.2288		
ELISA 100 μg	£3231	£14		6.2289	0.0002	
ELISA 50 μg	£3300	£83	£83	6.2287	− 0.0001	− 0.0001
Outpatient specificity 95%	**Costs**	**vs. 1**	**vs. 2**	**QALY**	**vs. 1**	**vs. 2**
ELISA 150 μg	£3201			6.2289		
ELISA 100 μg	£3211	£10		6.2291	0.0002	
ELISA 50 μg	£3256	£56	£56	6.2290	0.0002	0.0002

TABLE 65 Alternative presenting population, Basumani data and outpatient assessment specificity

Outpatient specificity	Costs	vs. 1	vs. 2	QALY	vs. 1	vs. 2
25%						
ELISA 150 μg	£3011			6.1950		
ELISA 100 μg	£3037	£25		6.1949	0.0000	
ELISA 50 μg	£3167	£156	£156	6.1941	− 0.0009	− 0.0009
50%						
ELISA 150 μg	£2991			6.1951		
ELISA 100 μg	£3011	£20		6.1951	0.0000	
ELISA 50 μg	£3112	£120	£120	6.1945	− 0.0006	− 0.0006
75%						
ELISA 150 μg	£2971			6.1953		
ELISA 100 μg	£2986	£14		6.1953	0.0001	
ELISA 50 μg	£3056	£85	£85	6.1950	− 0.0003	− 0.0003
95%						
ELISA 150 μg	£2956			6.1954		
ELISA 100 μg	£2965	£9		6.1955	0.0001	
ELISA 50 μg	£3012	£56	£56	6.1953	− 0.0001	− 0.0001

The lower prevalence of IBD increases the importance of specificity. Given the tests' sensitivities, as outlined in *Table 65*, as the specificity of the outpatient assessment increases the better specificity of ELISA with the 150 µg cut-off relative to the specificity of ELISA with the 100 µg cut-off falls in importance. The increase in net costs associated with ELISA with the 100 µg cut-off is reduced, and very small patient gains result.

Indeterminate results

In addition to the model structure permitting a sequence of tests to be explored as outlined above, it also permits indeterminate test results to be treated differently than determinate results. For instance, those whose first calprotectin test was within a range deemed to be indeterminate, such as 50–125 µg, might receive a second calprotectin test after a period of time, and be only referred if the result of this second calprotectin test was > 50 µg. Expert opinion indicates that perhaps 10–15% of first calprotectin test results might be indeterminate.

Retesting those who receive an indeterminate result from their first test will increase the cost of the calprotectin testing element. But whether this increases or decreases overall costs will mainly depend upon what tends to happen during the period between the tests to the calprotectin level among the IBS patients who received an indeterminate first test result. If for the vast majority of these patients it remains within the indeterminate range then referral rates would remain largely unchanged and total costs would increase. But if for a reasonable proportion of these patients it falls back to < 50 µg by the time of the second test, referrals to secondary care will be avoided. Given relative costs, only a minority of these patients would have to fall back to < 50 µg for the additional costs of the second test to be more than offset by the reduction in the costs of unnecessary referrals and further investigations. The secondary care costs avoided will be a function of the outpatient assessment specificity, as outlined above. Given the uncertainty around the impact of the outpatient assessment specificity, building further speculation around the impact of indeterminate results upon this seems to be of limited worth.

Base-case results: secondary care – inflammatory bowel disease versus non-inflammatory bowel disease

For the primary care base case, the patient numbers receiving the initial test and being referred for colonoscopy are as shown in *Table 66*.

TABLE 66 Secondary care: base-case results (%) of initial test sequence

| | Colonoscopy | | | ELISA | | | | | |
| | | | | 50 µg/g | | | 100 µg/g | | |
	First test	Colonoscopy	Final	First test	Colonoscopy	Final	First test	Colonoscopy	Final
IBD tested	–	47.9		47.9	47.4		47.9	45.0	
True positive	–	45.5	45.5	47.4	45.0	45.0	45.0	42.8	42.8
False negative	–	2.4	2.4	0.5	2.4	2.8	2.9	2.3	5.1
Non-IBD tested	–	52.1		52.1	13.5		52.1	9.4	
True negative	–	52.1	52.1	38.6	13.5	52.1	42.7	9.4	52.1
False positive	–	0.0	0.0	13.5	0.0	0.0	9.4	0.0	0.0

The ELISA with 50 µg/g cut-off results in 13.5% false positives being referred onward for colonoscopy, or 26.0% of the non-IBD patient population, as would be anticipated given the 74.0% specificity. This is at the minor cost of 0.5% false negatives not being referred on for colonoscopy at first presentation, or 1.0% of the IBD patient population as would be anticipated given the 99.0% sensitivity. The final results after colonoscopy are 45.0% true positives, 2.8% false negatives and 52.1% true negatives.

The ELISA with 100 µg/g cut-off results in only 9.4% false positives being referred onward for colonoscopy, or 18.0% of the non-IBD patient population, as would be anticipated given the 82.0% specificity. This is at the slightly larger cost of 2.9% false negatives not being referred on for colonoscopy at first presentation, or 6.0% of the IBD patient population as would be anticipated given the 96.0% sensitivity. The final results after colonoscopy are 42.8% true positives, 5.1% false negatives and 52.1% true negatives.

The differences between the two ELISA cut-offs are more marked in terms of false positives with the ELISA cut-off of 50 µg/g, resulting in 13.5% false positives being referred to colonoscopy compared with 9.4% for the 100 µg/g cut-off. But there is also a difference in the end results in terms of true positives being diagnosed at first presentation: 45.0% for the 50 µg/g cut-off and 42.8% for the 100 µg/g cut-off, a net difference of 2.3% of the presenting population or 4.4% of the presenting IBD population. Given the above this results in the following estimates (*Table 67*).

Faecal calprotectin testing with ELISA is estimated to be both cost saving and more effective than all patients receiving a colonoscopy. There are limited differences between the two ELISA cut-offs, with the 50 µg/g cut-off being slightly more expensive on average by £35, owing to an additional average £22 for tests among non-IBD patients and an additional average £13 among patients with CD owing to earlier

TABLE 67 Secondary care: base-case results

Comparators	QALYs	Tests	Other	Total
Colonoscopy				
CD	2.5773	£244	£6938	£7183
UC	0.8942	£83	£463	£546
Non-IBD	3.2094	£338	£629	£967
Total	**6.6809**	**£665**	**£8031**	**£8696**
ELISA				
50 µg/g				
CD	2.5767	£254	£6934	£7188
UC	0.8941	£86	£463	£549
Non-IBD	3.2117	£120	£634	£754
Total	**6.6824**	**£460**	**£8031**	**£8491**
100 µg/g				
CD	2.5757	£256	£6921	£7177
UC	0.8938	£87	£462	£549
Non-IBD	3.2119	£95	£634	£729
Total	**6.6814**	**£438**	**£8018**	**£8456**

diagnosis. It is also marginally more effective by 0.001 QALYs, which suggests a cost-effectiveness estimate of £33,982 per QALY, but it should be stressed that the estimated net QALYs are extremely small and that any change in the underlying inputs would have a large swing effect upon the ICER.

The central estimates and CEAFs from the probabilistic modelling run over 1000 iterations are as shown in *Table 68* and *Figure 26*, respectively.

The central estimate of net cost of ELISA with the 50 µg/g cut-off compared with ELISA with the 100 µg/g cut-off remains in line with the deterministic modelling at £35 as are the net QALYs at 0.001. The probabilistic central estimate of the cost-effectiveness of ELISA with the 50 µg/g cut-off compared with ELISA with the 100 µg/g cut-off is £33,088 per QALY.

Up to a willingness to pay of around £30,000 per QALY it is estimated that the ELISA with the 100 µg/g cut-off is most likely to be cost-effective and has the highest monetised health benefits net of costs. Thereafter, as the willingness to pay rises further it is estimated that the ELISA with the 50 µg/g cut-off is most likely to be cost-effective and has the highest monetised health benefits net of costs.
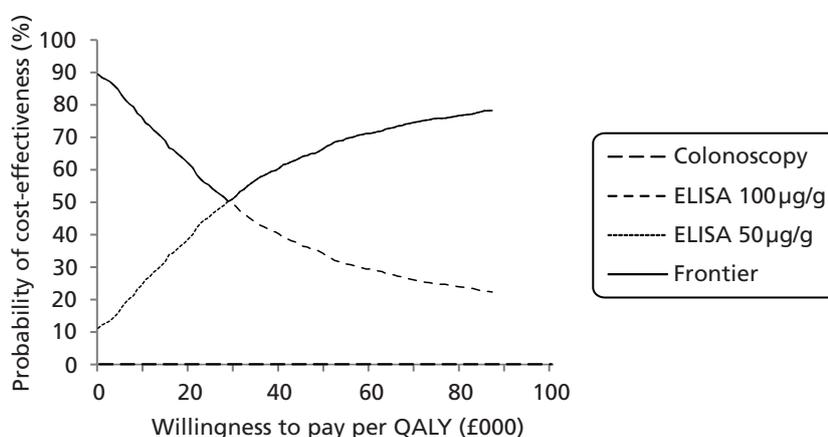
Sensitivity analyses: secondary care
The univariate sensitivity analyses for secondary care result as shown in *Table 69*.

As for primary care, most of the changes appear to broadly affect the three comparators in a like manner. The main difference arises from varying the prevalence of IBD, which tends to reduce the cost savings from FC testing as the prevalence rises, as would be anticipated. The source of utilities also has an impact upon the anticipated net gain from ELISA with the 50 µg/g cut-off compared with ELISA with the 100 µg/g

TABLE 68 Secondary care: probabilistic modelling central estimates

| | | Base case | |
		QALYs	Costs (£)
Colonoscopy		6.6960	8553
ELISA	50 µg/g	6.6975	8348
	100 µg/g	6.6965	8313

FIGURE 26 Cost-effectiveness acceptability frontier: secondary care – base case.

TABLE 69 Secondary care: univariate sensitivity analyses

Analysis	QALYs					Costs (£)					
	Colonoscopy	50 µg/g	100 µg/g	Net		Colonoscopy	50 µg/g	100 µg/g	Net		
	S1	S2	S3	S2 – S1	S2 – S3	S1	S2	S3	S2 – S1	S2 – S3	
Base case	6.6809	6.6824	6.6814	0.0015	0.0010	8696	8491	8456	–205	35.21
40% IBD prevalence	6.5950	6.5970	6.5962	0.0020	0.0008	7569	7330	7292	–239	37.11
60% IBD prevalence	6.8127	6.8135	6.8121	0.0008	0.0014	10,425	10,271	10,239	–154	32.28
8-week re-presentation	6.6828	6.6845	6.6839	0.0017	0.0006	8707	8493	8463	–214	30.34
16-week re-presentation	6.6789	6.6805	6.6789	0.0015	0.0016	8689	8479	8442	–209	37.18
Utilities non-Stark	7.2055	7.2069	7.2066	0.0014	0.0003	8696	8491	8456	–205	35.21
No colonoscopy mortality	6.6815	6.6829	6.6818	0.0013	0.0011	8697	8492	8456	–205	35.21

cut-off, the ICER for which worsens to £117,000 per QALY. But this may be to overstate the effect given the prevalence of CD within the presenting population and the perhaps rather small QoL decrement sourced from Gibson *et al.*[134]

Summary and discussion

Previous economic analyses have typically concluded that FC testing is cost saving compared with the situation without it. Given test specificities and the assumed prevalences of IBD in the presenting population, the additional cost of the FC testing is more than offset by the reduction in the cost of unnecessary colonoscopies. The YHEC 2010 report[50] for the CEP concluded that FC testing not only saved money through the diagnostic pathway, but that it also resulted in more true positives and true negatives owing to its superior sensitivity and specificity and so dominated the situation of the GP referral in the absence of FC testing.

Sensitivities

A distinction of the EAG modelling from that of the literature is that for primary care the GP in the absence of FC testing is estimated to have at least as good a sensitivity as the FC tests, and for some comparisons a better sensitivity. In this circumstance, the GP results in as many or more true positives than FC testing. As a consequence, despite quite large cost savings and fewer false positives still being estimated for FC testing compared with the GP referring in the absence of FC testing, dominance for FC testing cannot be definitively concluded on the basis the diagnostic pathway alone. There is an argument that the cost and QALY impacts among the false negatives also need to be considered.

The costs and QALY impacts among the false negatives is in the first instance dependent upon the prevalence of IBD in the presenting population and the sensitivities of the tests under consideration. A low IBD prevalence and high test sensitivities mean that there will be few false negatives, although higher IBD prevalences and lower test sensitivities will increase the number of false negatives and so the importance of considering the costs and QALY impacts among them. These latter are dependent upon the average time spent as false negative prior to re-consideration of IBD within a diagnostic pathway. The QALY impacts are also dependent upon the source of the QoL estimates for false negative being incorrectly treated and for those correctly diagnosed, the latter requiring QoL estimates for remission and no remission. In the absence of other data it has been assumed that the QoL among those remaining as a false negative and being incorrectly treated is the same as among those correctly diagnosed but not in remission. The longer the period of time spent as a false negative and the larger the QoL gain from achieving remission, the larger the impact of false negatives and so the greater the importance of tests' sensitivities.

Modelling induction of remission and maintenance of remission is eased for the current assessment by the relevant models for both CD and UC being available as appendices to the respective clinical guidelines, although the UC guidelines is still under consultation. A key assumption within these is that there is no disease progression, such as the development of fistula, when patients are not in remission. Were this to apply, it would also increase the importance of tests' sensitivities.

Adverse events

The modelling also needs to consider the adverse impacts of unnecessary colonoscopies. Owing to data constraints, the cost impacts have been limited to modelling the cost impacts of the relatively rare serious adverse events of bleeds and perforations. The QoL impacts are limited to the mortality associated with perforations. Although perforations are rare, so resulting in a very low mortality rate, the QALY impact of this persists for the duration of the model.

There is evidence from the literature that colonoscopies result in minor adverse events among a reasonable proportion of patients; for example, de Jonge et al.[143] suggest that perhaps around 40% of those investigated with colonoscopy have some effects persisting 30 days subsequent to the colonoscopy. In common with the CG118 guideline on screening for colorectal cancer with colonoscopy,[149] these minor adverse events have not been taken into account in the modelling principally because of a lack of QoL data. The effects of minor and transient colonoscopy side effects seem unlikely to affect the conclusions for the comparisons of no FC testing with FC testing, but they may take on a greater significance in the context of comparing different FC tests or different cut-offs. Depending upon the prevalence of IBD in the presenting population, inclusion of these minor adverse events would increase the importance of tests' sensitivities.

Primary care modelling

For the primary care base case for diagnosis of IBD versus IBS in an adult population, GP referral in the absence of FC testing is compared with CalDetect at the 15 μg/g cut-off and with ELISA at the 50 μg/g cut-off. The choice of CalDetect at the 15 μg/g cut-off may initially seem surprising, but the data from Otten et al.[73] suggests a very much worse sensitivity for the 60 μg/g cut-off of only 61%, which renders it of questionable clinical relevance. A 6.3% prevalence of IBD is drawn from the Durham primary care data.

Within the total patient population both the GP without FC and the initial Caldetect test identifies all 6.3% of patients with IBD as true positives. The colonoscopy subsequent to this identifies 6.0% of the 6.3% referred as true positives owing to its 95% sensitivity. The ELISA test is slightly worse, identifying only 5.9% as true positives, with 0.4% being wrongly classified as false negatives. Of the 5.9% referred to colonoscopy, 5.6% are identified as true positive, resulting in a total of 0.7% false positives.

Within the total patient population, the GP without FC testing incorrectly identified 19.8% as false positives requiring referral to colonoscopy. The rates of false positives incorrectly referred to colonoscopy for CalDetect and ELISA are much lower: 5.1% and 5.6%, respectively.

Despite its additional initial test costs, FC testing is estimated to result in cost savings compared with the GP without FC testing: £83 for CalDetect and £82 for ELISA. This is on average per patient. This is due mainly to the lower number of colonoscopies. Small QALY gains of around 0.001 QALYs also accrue, although these are limited, as the low prevalence of IBD and the similar high sensitivities of the tests result in relatively few false negatives. Some of the QALY differences accrue from the very slightly lower mortality associated with the lower number of colonoscopies. CalDetect and ELISA are estimated to be broadly equivalent with only minor differences between them. Probabilistic modelling results in similar estimates.

Sensitivity analyses around the base case suggest that FC testing results in patient gains and remains cost saving compared with the GP without FC testing up to an IBD prevalence of 25%. At this point, owing to ELISA having a less-than-perfect sensitivity ELISA starts to result in very slight QALY losses compared with the GP without FC testing, although retains cost savings of around £63 per patient on average. The resulting estimate for the cost-effectiveness of the GP without FC testing compared with ELISA is £378,000 per QALY. Owing to its perfect sensitivity, CalDetect remains both more effective and cheaper than the GP without FC testing.

The primary care patient group in whom FC is used may be wider than the data set used for the estimation of the sensitivity and specificity of the GP without FC testing. Doubling the size of this patient group and allowing for some additional IBD patients within the wider patient group results in a lower IBD prevalence of only 3.3%, and also sensitivity and specificity estimates for the GP without FC testing in this wider patient group of 94.1% and 89.7%, respectively. Despite this improvement in specificity, the GP without FC testing is still estimated to result in higher costs and lower QALYs than both CalDetect and

ELISA FC testing, although the margin between the with FC testing and without FC testing is narrows quite significantly.

Other univariate sensitivity analyses suggest that the primary care base-case results are reasonably robust. The main sensitivity of the results of CalDetect compared with ELISA arise from changing the source of utilities and shortening the time spent as false negatives. These both tend to reduce the importance of false negatives and so reduce the importance of tests' relative sensitivities, and so reduce the estimated net QALY gain from CalDetect over ELISA. But in all of this, it should be stressed that the QALY differences between the FC tests are very small.

Secondary care

For the secondary care paediatric population for the diagnosis of IBD versus non-IBD, direct referral to colonoscopy is compared with ELISA with the 50 μg/g cut-off and ELISA with the 100 μg/g cut-off. The base-case prevalence of IBD of 47.9% increases the importance of test sensitivities compared with the primary care setting, and so the effect of false negatives upon the modelling outputs. Within the total patient population ELISA with the 50 μg/g cut-off refers 47.4% as true positives for colonoscopy, whereas ELISA with the 100 μg/g cut-off refers 45.0% as true positives for colonoscopy. Colonoscopy is assumed to have a sensitivity of 95%, so the end diagnosis if all are referred immediately to colonoscopy is 45.5% being diagnosed with IBD. For those referred to colonoscopy by ELISA with the 50 μg/g cut-off 45.0% are diagnosed as having IBD, while for those referred to colonoscopy by ELISA with the 100 μg/g cut-off 42.8% are diagnosed as having IBD; a net difference between the cut-offs of 2.2%.

Despite the higher IBD prevalence in the secondary care population, the main test differences still lie in the number of unnecessary colonoscopies. Without FC testing all 52.1% of non-IBD patients receive a colonoscopy, compared with 13.5% for the ELISA with the 50 μg/g cut-off and only 9.4% for ELISA with the 100 μg/g cut-off.

The additional ELISA test costs are more than offset by the savings from reduced colonoscopies. Compared with referring all directly to colonoscopy, ELISA with the 50 μg/g cut-off is estimated to save £205 per patient on average, whereas ELISA with the 100 μg/g cut-off is estimated to save £240. Small QALY gains of around 0.001 QALYs are modelling for ELISA compared with direct referral to colonoscopy, these being slightly larger for ELISA with the 50 μg/g cut-off owing to its better sensitivity. But given the additional average £35 cost, the cost-effectiveness estimate for ELISA with the 50 μg/g cut-off compared with ELISA with the 100 μg/g cut-off is £35,000 per QALY. As before for the primary care modelling, it should be stressed that the QALY differences between the FC tests are very small and perhaps not too much should be read into these differences. The central estimates from the probabilistic modelling are in line with those of the deterministic modelling.

Sensitivity analyses suggest that the base-case results are reasonably robust, although the anticipated QALY gain from ELISA with the 50 μg/g cut-off compared with ELISA with the 50 μg/g cut-off shows some sensitivity the prevalence of IBD, the source of the utilities and the assumed average period of time spent as false negatives as would be expected.

For the modelling in secondary care, compared with the primary care modelling there is additional uncertainty in terms of the model structure. The model is principally a model of IBD versus IBS in an adult population. It may not be as suited to the secondary care paediatric population, for which the distinction is between IBD and non-IBD. The non-IBD paediatric patients also have a higher proportion of conditions other than IBS compared with the adult patient population. But the main differences in terms of costs arise from the upfront test costs, and these will apply within any model construct. A distinction also needs to be

drawn between the additional costs of incorrect treatment among false negatives, which for a given set of inputs will be correctly estimated by the model structure, and the structural uncertainty around the appropriate model inputs for CD and for UC in a paediatric population.

Quality-of-life summary
For details see *Tables 70–73*.

TABLE 70 Inflammatory bowel syndrome HRQoL studies reporting utilities

Paper	Brazier[117]		Puhan[121]			Spiegel[5]								Wang[4]	
Year	2004		2007			2009								2012	
Country	UK		Canada			USA								Singapore	
Setting	Primary		Gastrenterology			General								Symposium	
n	161		96			257	41	82	134	–	–	–	–	198	251
Mean age, years	47		40			43	NA	NA	NA	–	–	–	–	52	57
Female, %	86		84			79	NA	NA	NA	–	–	–	–	47	58
Mean duration, years	NA		2.7			11.2	NA	NA	NA	–	–	–	–	NA	–
Condition	IBS		IBS			All	IBS-C	IBS-D	IBS-M	–	–	–	–	IBS	No IBS
Severity	NA		NA			Mild–severe	–	–	–	Non-severe	Severe	3-montd relief	3-montd no relief	NA	–
Measurement	EQ-5D	SF-6D	TTO	SG	SF-6D	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D[a]	EQ-5D	EQ-5D
HRQoL	0.662	0.666	0.760	0.840	0.850	0.744	0.760	0.760	0.730	0.800	0.700	0.780	0.730	0.739	0.849

NA, not available.
a Mapping function unclear.

TABLE 71 Inflammatory bowel disease HRQoL studies reporting utilities

Paper	Konig[122]				Stark[124]					
Year	2002				2010					
Country	German				German					
Setting	Outpatient	Inpatient	Outpatient and inpatient		General					
n	121	31			270			232		
Mean age, years			41	46	41			44		
Female, %			53	28	63			55		
Mean duration, years			15	10	14			13		
Condition			CD	UC	CD			UC		
Severity, %	79 remission	7 remission	60 remission	70 remission	All	Remission	Active	All	Remission	Active
Measurement	EQ-5D[b]	EQ-5D[b]	EQ-5D[b]	EQ-5D[b]	EQ-5D	EQ-5D	EQ-5D	EQ-5D	EQ-5D	EQ-5D
HRQoL	0.803	0.619	0.875	0.627	0.770	0.890	0.610	0.840	0.910	0.710

Paper	[a]Casellas[126]	
Year	2005	
Country	Spain	
Setting	Outpatient and inpatient	
n	628	528
Mean age, years	34	38
Female, %	56	50
Mean duration, years	5.0	4.0
Condition	CD	UC
Severity, %	Remission	Mild	Moderate–severe	Remission	Mild	Moderate–severe
Measurement	EQ-5D[c]	EQ-5D[c]	EQ-5D[c]	EQ-5D[c]	EQ-5D[c]	EQ-5D[c]
HRQoL	0.800	0.720	0.600	1.000	0.720	0.500

a Median values reported.
b German mapping function.
c Spanish mapping function.

TABLE 72 Crohn's disease HRQoL studies reporting utilities

Paper	Buxton[130]		Casellas[132]	Gregor[133]							
Year	2007		2007	1997							
Country	UK		Spain	Canada							
Setting	Trial		Outpatient	Outpatient and inpatient							
n	NA		49	180	52	34	45	49			
Mean age, years	NA		40	35	35	31	34	37			
Female, %	NA		47	56	62	53	49	59			
Mean duration, years	NA		8.0	8.9	10.3	8.0	7.7	9.3			
Condition	CD		CD	CD	CD	CD	CD	CD	CD	CD	CD
Severity	Moderate-severe		Remission	All	Chronic refractive	Chronic responsive	Active	Remission	Mild	Moderate	Severe
Measurement	ED-5D	SF-6D	EQ-5D[b]	TTO / SG	TTO / SG	TTO / SG	TTO / SG	TTO / SG	TTO / SG	TTO / SG	TTO / SG
HRQoL	0.700	0.680	1.00	0.920 / 0.810	0.880 / 0.740	0.980 / 0.860	0.890 / 0.770	0.960 / 0.880	0.955 / 0.815	0.880 / 0.725	0.720 / 0.519

NA, not available.
a Median values reported.
b Spanish mapping function.

TABLE 73 Ulcerative colitis HRQoL studies reporting utilities

Paper	Connolly[136]	Bryan[137]			Arseneau[138]					Poole[141]		
Year	2009	2008			2006					2009		
Country	Western Europe	UK								UK		
Setting	Trial	HODaR[a]								–		
n	127	171								126		
Mean age, years	42–47	NA								43		
Female, %	40	NA								41		
Mean duration, years	NA	NA								–		
Condition	UC	UC								UC		
Severity	Mild to moderate	Remission	Active	Surgical remission	Remission	Active	Surgical remission	Complications of surgery		Remission	Mild/moderate	Severe
Measurement	EQ-5D	EQ-5D	EQ-5D	EQ-5D	TTO	TTO	TTO	TTO		EQ-5D	EQ-5D	EQ-5D
HRQoL	0.771[b]	0.880	0.420	0.600	0.790	0.320	0.630	0.490		0.944	0.811	0.700

NA, not available.
a Possibly a manufacturer-commissioned QoL study.
b Baseline pooled between arms.

Chapter 4 Discussion

Principal findings

The key findings of this review are:

- In adults, FC is a good indicator of inflammation in the bowel and can be used to distinguish between IBS and IBD in cases for which the differential diagnosis is in doubt.
- Calprotectin could be very useful for GPs as a way of confirming a clinical diagnosis of IBS, although it will not be required in all people with IBS, because in some, other features such as a long history, comorbidities, relationship to stress and an absence of weight loss, may tilt the balance of probability to IBS.
- It is not a perfect test because some patients with IBS have raised calprotectin levels but false-negative IBD is unusual if we use the cut-off of 50 µg/g (for ELISA tests) and 15 µg/g (for PreventID POCT) recommended by the manufacturers.
- In children, it is useful for distinguishing between IBD and non-inflammatory conditions.
- From the clinical perspective, the balance of risk between sensitivity (not missing any cases of IBD) and specificity (avoiding false positives – people with IBS thought to have IBD) may best be towards sensitivity because missed IBD can lead to much more serious consequences than an unnecessary colonoscopy, but given the low prevalence of IBD in the primary care population, it is specificity that drives relative costs in this setting.
- There are a few patients who have slightly raised levels (50–150 µg, or perhaps to 200 µg in children) who may only need monitoring. In many cases, calprotectin level will fall and no further investigation will be necessary. In those who have low-grade IBD, calprotectin will usually rise.
- There are few head-to-head comparisons of different tests, but such data as there are do not suggest significant differences in clinical reliability.
- There are no published studies in patients drawn only from primary care.
- If calprotectin testing is made available in primary care, GPs could be much more selective in whom they refer to specialist care. Referrals will fall considerably.
- In secondary care, both paediatric and adult, the availability of calprotectin testing could lead to a reduction in the number of colonoscopies performed.
- It is likely that delays in diagnosing IBD could be reduced, as a raised calprotectin will alert clinicians. This may be particularly useful in children where the onset can be insidious, as it can also be in some adults.
- Calprotectin testing would lead to cost savings, mainly in secondary care from a reduction in colonoscopies.
- Measurement of ESR and CRP in patients with ?IBS, ?IBD, should cease.

Uncertainties

Evidence from primary care

As noted by several commentators, nearly all of the evidence on calprotectin comes from studies from large GI clinics and referral centres.[33,52] The value of the test in primary care is for ruling out IBD, and confirming a presumptive diagnosis of IBS. High sensitivity is therefore required. In theory, sensitivity and specificity are not influenced by prevalence but they may be if the spectrum of disease alters. However, NPV will be affected, and for GPs is more useful than sensitivity and specificity – a very high NPV will be used to rule out IBD.

Borderline faecal calprotectin results

This usually refers to patients with FC levels in the 50 to 150 µg/g or 200 µg/g range. Most may come to little harm, may have little visible pathology on endoscopy or video capsule imaging, but some may have very mild CD. They could be monitored for abdominal pain, diarrhoeas and weight loss. The calprotectin test should be repeated at intervals, perhaps around 6–8 weeks, as deemed necessary. The trend over time may be a useful guide to management.

Moroni et al.[156] (abstract only) from Glasgow followed up 158 patients newly referred to a GI clinic, aged 16 to 50 years, with FC values of 50–100 µg/g, using the Bühlmann ELISA test. They excluded anyone with known IBD, and anyone who had had a previous FC of > 100 µg/g. Colonoscopy was carried out in 82, and no IBD was found. There were six patients with abnormalities: three with single small (< 10 mm) adenomas, one helminth infection, one diverticulosis and one acute inflammation. They also studied a group with FC < 50 µg/g, and found no IBD. They conclude that a FC level of < 100 µg/g excludes significant pathology.

Zayyat et al.[157] (abstract only), from the Kings College group, linked data from a very large database of FC results, with electronic patient records, to find out what happened to people who had had a borderline FC ('50–100 mg/f', which, presumably, means 50–100 µg/g), and who had had further investigation with lower GI endoscopy or MRI or CT. Of 433 patients, only 10 had IBD confirmed, and in almost all of these, a repeat FC had shown an increase. The remaining 423 patients were followed up for an average of 3.6 years, but none developed IBD. This suggests that the threshold for action might be raised to 100.

Mohammed and Smale[158] from Leeds (abstract only) report a group of patients who had an abnormal calprotectin (over 50 µg/g), but in whom follow-up endoscopic or radiological investigations were normal. After 3 years, none of those whose baseline calprotectin had been < 225 µg/g developed organic disease. However, details in the abstract are limited, with no duration of symptoms prior to first calprotectin testing. Some may have had post-infectious, self-limiting inflammation.

Lee et al.[159] from County Durham trialled calprotectin using the manufacturer's recommended cut-off of > 60 µg/g in a series of 122 patients. The NPV was high (94%), with only 4 of 71 patients with negative FC being found to have any pathology (although not all underwent endoscopy), and none of them had IBD. However, the results amongst those with positive FC tests were more mixed: 19 of the 51 had organic disease, including 9 (18%) with IBD. But 32 (63%) were diagnosed as having functional bowel disorders. All of those with IBD had FC levels well above 100 µg/g.

Henderson et al.[30] from Edinburgh report that in their paediatric group, they tend to regard a cut-off of 200 µg/g as most useful.

Demir et al.[160] from King's College Hospital (abstract only) followed up 66 patients with borderline calprotectin (50–150 µg/g) for 2 years. None developed IBD and calprotectin tended to fall.

Koulaouzidis et al.[161] from Edinburgh reported results in a highly selected group of 70 patients suspected of having IBD but in whom both colonoscopy and gastroscopy had found no lesions, and in whom localised small bowel CD was suspected. No patient with calprotectin under 100 µg/g had evidence of CD on small bowel capsule endoscopy, whereas 43% (15/35) of those with calprotectin over 100 µg/g were found to have CD (mean 326 µg/g; range 116–1430 µg/g).

One issue concerning cut-offs is what the sensitivity cut-off should be based on. Should it be based purely on presence of disease – IBD or no IBD? Or should it be based on likely need for treatment? If the latter, perhaps we need two cut-offs and three groups: negative, presumed not to have IBD; positive, IBD requiring treatment; and intermediate, IBD that requires only monitoring meantime.

Manufacturers' current recommended cut-offs appear to be based on the first of the two cut-offs above, and on ensuring that nothing is missed. This seems reasonable in the present state of knowledge.

Delayed diagnosis is a well-known problem in IBD, especially CD. A study by Vavricka et al.[162] from the Swiss IBD Cohort Study group reported that among 1591 patients with IBD, median delay was 9 months in CD and 4 months in UC. Vavricka et al.[162] noted that the delay could be because in patients with mild IBD, symptoms were similar to those of IBS. In one-quarter of patients with CD, the diagnosis took longer than 24 months, and in 10% it took 8 years or longer.

There were two phases to the delay – time from symptoms to consulting a physician; and time from consulting a physician to getting the diagnosis. The second phase was longer, especially in younger patients.

Peyrin-Biroulet et al.[163] reviewed population-based studies of the natural history of CD and reported that up to one-third of patients had intestinal complications, such as strictures at diagnosis.

Diagnosis is also often delayed in paediatric UC and delays are associated with poorer outlook. Gower-Rousseau et al.[21] reported that a delay of more than 6 months in diagnosis was associated with an increased risk of extensive disease.

One crucial issue for the economics is whether someone with CD and mild, or no, symptoms and lowish FC could develop a serious complication, such as a stricture or fistula, or develop an ileal mass requiring surgery. Low-grade inflammation can continue with little in the way of symptoms. 70–80% of patients with CD will require surgery within 5 years of diagnosis. Note that the average time from onset of symptoms to diagnosis of IBD is 12–18 months, perhaps longer if symptoms are mild.

One problem for this group if FC was not available would be that GPs might monitor with CRP, ESR and Hb, which would all be likely to be normal.

Comparative data on the relative performance of the point-of-care tests in the intermediate 50 to 200 µg/g zone would be useful.

Clinical activity scores

The value of clinical activity scores may be overestimated. This applies more to monitoring of disease activity after diagnosis. Schoepfer et al.[26] surveyed Swiss gastroenterologists and found that most considered clinical activity to be adequate for monitoring disease activity, rather than using markers such as calprotectin. Only 28% of gastroenterologists used calprotectin in more than 70% of their IBD patients. Schoepfer et al.[26] concluded that clinical practice in Switzerland was not keeping up with clinical science.

It has been suggested that IBS and IBD may co-exist. This may have arisen because some patients who appear to be in remission according to CDAI and Ulcerative Colitis Activity Index (UCAI), have symptoms suggestive of IBS and meet the Rome II criteria. Farrokhyar et al.[164] reported that the majority of their patients with 'inactive' IBD had symptoms matching the Rome II criteria. The inactivity was based on them not having had a change in therapy for 12 months.

However, the advent of calprotectin has shown that what many of these patients have is ongoing inflammation. Keohane et al.[165] from Cork studied a group of patients with CD and UC who were apparently in remission, as judged by physicians assessment, CRP of < 10 mg/l, no treatment in last 6 months, and CDAI score of ≤ 150, or UCAI score of ≤ 3. Sixty per cent of the CD group and 39% of the UC group had symptoms that met the Rome II criteria. Calprotectin testing revealed raised levels, indicating that the IBS-like symptoms were due to active IBD. A control group of people with true IBS had normal FC levels. (Note that figure 3 of the paper is wrongly labelled, with IBD that should be IBS.)

In many patients with IBD, symptoms persist. A Finnish study[166] reported that 77% of people with IBD who responded to a survey questionnaire (response rate 40%) had symptoms that impaired QoL. The mean HRQoL was 169 using the IBDQ, which has a range of 32 to 224, with high being better.

Implications of wider use of calprotectin

There are implications of calprotectin testing for the NICE guidance on drug treatment of CD. TA187[14] states: 'Infliximab and adalimumab . . . are recommended as treatment options for adults with severe Crohn's disease whose disease has not responded to conventional therapy (including immunosuppression and/or corticosteroid treatments).' Severe is defined by reference to CDAI score of 300 or more or Harvey–Bradshaw score of 8–9 or more. Paragraph 2.9 notes that: 'The CDAI is frequently used to assess disease severity.'

The trials used to underpin the NICE guidance used CDAI as main outcome.

There is no mention of calprotectin in TA187,[14] because at the time it was written, there was insufficient evidence to support its use. There is an important implication of the use of calprotectin for the NICE guidance. We have noted that clinical scores such as CDAI do not correlate well with mucosal inflammation. We have also noted that some people with CD in apparent remission with 'IBS symptoms' have been shown by calprotectin testing to have ongoing inflammation.[167]

Calprotectin testing could reveal a group with few or no symptoms but ongoing inflammation, in whom the anti-TNFs are not recommended. So the present NICE guidance may leave many people with inadequately controlled CD. Treatment of this group may be cost-effective.[28] TA187[14] may need to be reviewed in the light of calprotectin data. There is evidence from Cardiff of an association between increasing use of immunosuppressants and decreasing the need for surgery.[168] Calprotectin can also be used to predict relapse.[169]

Calprotectin is also useful in children, as a non-invasive guide to mucosal inflammation and disease activity in previously diagnosed IBD,[170,171] in an admittedly small series of teenagers with IBD, examined clinical activity indices (PUCAI and PCDAI) CRP and FC for predicting relapse.[172] Calprotectin was more useful than clinical scores but CRP was not helpful.

The use of calprotectin for monitoring disease activity is outwith the scope of this review and appraisal, but we recommend that the NICE Technology Appraisal Programme should consider when best to assess the impact of calprotectin testing on current guidance on treatment.

Possible other implications

Without calprotectin testing, there could theoretically also be QALYs gained by the incidental finding and removal of polyps that might over years turn malignant. However, this might be balanced by a reduction of colonoscopies in people with IBS releasing endoscopy resources for more timely investigation of those in the colorectal cancer age range, for example after screening.

Earlier diagnosis and earlier treatment

Would earlier treatment based on a sensitive test to identify inflammation enable treatment to be started earlier? Could this, by inducing 'deep remission', reduce the risk of later complications? (For review, see Panaccione et al.[173]) D'Haens et al.[174] have reported a close correlation between calprotectin and lesions seen on endoscopy.

As noted above, the current NICE guidance does not recommend biological agents such as infliximab (which might be regarded as 'disease modifying') until after treatment aimed at relieving symptoms has failed.

Another issue is that people with IBS may have a succession of different treatments, possibly involving several therapeutic trials, as outlined in *Chapter 1*, based on the NICE IBS guidelines. So it might take 6 months or longer before someone with IBD, treated as IBS, was recognised as not IBS.

Irritable bowel syndrome can be difficult to manage, and some people with IBS will still be referred for specialist advice. So the advent of calprotectin testing would not mean that none of those with negative calprotectin would be referred. One issue raised by an anonymous referee was the relative reassurance gained by anxious patients after endoscopy (which might only be by flexible sigmoidoscopy) versus the reassurance provided by a negative calprotectin. However, Spiegel *et al.*[146] found no difference in SF-36 scores in people with IBS who had undergone colonoscopy from those who had not. The proportions reassured were similar.

Inflammatory irritable bowel syndrome?

Some patients in some studies that were diagnosed as having IBS, after investigation, had raised FC levels, and it does appear that some people with IBS have an inflammatory component. It has been suggested that this may be due to disturbances in the intestinal bacteria, followed by a mucosal response.[163]

In the County Durham study the range of calprotectin results in those who tested positive (> 60 µg/g) but whose final diagnosis was functional bowel disease was 61–547 µg/g (mean 153 µg/g).[175]

D'Haens *et al.*[174] reported an overlap of calprotectin results between patients with IBD and a group with IBS: range for IBS 16–139 µg/g.

Therefore, some people with IBS are positive on calprotectin testing. Are they 'false positives' or does IBS represent a mix of conditions, some of which have inflammation?

One possibility could be if people with IBS use NSAIDs (including over-the-counter ibuprofen) that raise calprotectin. However, this cannot be the sole explanation, because raised calprotectin levels have been reported in studies that exclude people using NSAIDs.[60,82]

As noted in a previous assessment report for NICE, some people with IBS-D may have bile acid malabsorption.[176] This is a condition in which bile acids are not absorbed as they usually mostly (90%) are in the ileum. The SeHCAT report has no mention of calprotectin. It does not appear to be raised in bile acid malabsorption (except in those whose bile acid malabsorption is due to CD in the bile acid absorption site in the ileum) and so that will not be a source of false positives.

Others with IBS may have it subsequent to infectious gastroenteritis, where calprotectin would be raised during the infectious episode, but would then be expected to return to normal. Or does inflammation sometimes continue?

The answer in this group may be repeat testing. In those with raised calprotectin after bowel infection, the level will fall after a few months. However, there does appear to be a small number of patients (< 1%) with no evidence of IBD after thorough investigation (A Dhar, Darlington and Bishop Auckland General Hospitals, 2013, personal communication).

There is also a small group with IBS and mild inflammation that responds to NSAIDs (N Read, 2013, personal communication).

The high calprotectin levels seen in some people with IBS are perhaps not surprising, as raised levels have been reported in a random sample of the general population by Poullis *et al.*[177] They examined the association between lifestyle factors associated with colorectal cancer, and FC levels, in a group of people aged 50–70 years, and found that one-quarter had calprotectin levels above the upper limit of normal of

65 µg/g. The range was 2 to 440 µg/g. Some recruits may have had colonic adenomas, which are common in this age range.

Inflammatory bowel disease with normal calprotectin

Several studies showed that a few patients with confirmed IBD did not have raised calprotectin. One possibility might be that some have fibrotic post-inflammatory CD, although that seems unlikely in newly presenting patients.

Raised calprotectin in larger adenomas

The raised calprotectin seen in some patients with larger adenomas, reported best by Kok et al.,[39] might reflect bleeding, or possibly that the progression from small adenomas to large ones or cancer may have an inflammatory component. It raises the question whether calprotectin might have a role in colorectal cancer screening. Consideration of that is outwith the scope of this review but we note that Hoff et al.[178] from Norway compared calprotectin (PhiCal, Eurospital) with FOBT (using an immunochemical method) and concluded that FOBT was better.

Use of calprotectin in routine care

Trials and other studies may be prone to patient selection bias, and may be an imperfect guide to the use of calprotectin testing in routine care. As previously mentioned, we have data from the Durham Dales pilot of implementation that show that considerable savings can be made. The Cardiff data show that a considerable proportion of referrals by GPs are to confirm IBS (by exclusion of IBD).[38] So the main value of calprotectin testing may be to confirm presumptive diagnoses of IBS, and that can be done in general practice. Other studies (Rotherham unpublished) report that > 60% of colonoscopies in the 'pre-calprotectin era' showed no pathology.

Alrubaiy et al.[179] from Llanelli report results in 74 patients referred to a District General Hospital with intestinal symptoms. Depending on local practice, some had colonoscopy with, or before, calprotectin testing. Two were confirmed to have IBD and both had raised calprotectin levels (mean 271 µg/g). Another 14 had raised levels but further investigations were normal. In the group of 18 who had colonoscopy before calprotectin testing, all colonoscopies were normal, but calprotectin was tested later because of continuing symptoms. In six of the 18, calprotectin was raised – but not by much: mean was 114 µg/g. All were finally diagnosed as having no IBD. In the group of 23 that had calprotectin measured before colonoscopy, it was raised in eight (mean 171 µg/g) who did not have confirmed IBD.

Taylor et al.,[180] from the Isle of Wight, present some data from a small audit. Twenty-three patients had calprotectin levels of > 50 µg/g, of whom 18 had endoscopy. Only six of these had IBD, and all had calprotectin levels of > 200 µg/g, with all but one having levels over 300 µg/g. However, one patient with an initially negative calprotectin later developed IBD.

A key implication of calprotectin testing is the likely reduction in the number of colonoscopies required. Without it, many patients with IBS will undergo endoscopy. Data from various centres show that over 60% of colonoscopies in patients under the age of 45 years show no abnormal findings. Data from Newark and Sherwood show 64% of colonoscopies were normal (Newark and Sherwood CCG, 2012, personal communication).

Results of studies reported earlier, and the consensus of the NICE expert panel, suggest that a positive point-of-care test is sufficient as a guide to referral, without quantitative ELISA testing. The latter may be done as a baseline for assessing need for, or response to, treatment.

A recent study from Brighton by Pavlidis et al.[181] reported on the use of calprotectin in routine care.

The starting point was submission of a sample for calprotectin testing. The study focused on patients aged 18–45 years with suspected IBS and without alarm ('red flag') symptoms (although it turned out that some

did have such symptoms). The test used was the Bühlmann ELISA EK-CAL. A positive calprotectin result was taken to mean FC of > 50 μg/g. It was left to individual GPs to decide how to manage patients with positive results, including whether to refer them to gastroenterology. Similarly, GPs could refer patients with calprotectin levels of < 50 μg/g if they so wished.

Inevitably in a 'real-life' study, there are limitations, such as the absence of a 'gold standard' reference test such as endoscopy for all. That would have been impractical and unethical outwith the setting of a RCT. There was no routine follow-up testing of all those with calprotectin levels of < 50 μg/g, or of those with higher levels, whom GPs did not refer. The final diagnosis was based on clinical follow-up for 12 months. In patients referred to specialist care, the decision on further investigation, including endoscopy, was left to the individual specialist. Fourteen per cent (134) of patients had colonoscopy and 11% (104) had a flexible sigmoidoscopy. The study would therefore not be eligible for inclusion in our adult meta-analyses because we regarded colonoscopy as the reference test. Another issue may have been spectrum bias, in that GPs could refer patients in whom they suspected IBD, by another pathway. So those with IBD in this study probably had milder forms.

However, some useful findings emerged:

- Few people (2.5%, 17 out of 686) who had calprotectin levels of < 50 μg/g were diagnosed at the 12-month follow-up as having organic disease.
- Raising the cut-off to 100 μg/g would have little effect (4%) on sensitivity but much more (14%) on specificity. The NPV hardly changes (98% vs. 97%) but the PPV improves from 28% to 49%.
- Raising the cut-off to 150 μg/g gives NPV 97% and PPV 71%.
- Considerable savings could result, although the authors note that a considerable number of those with calprotectin levels of < 50 μg/g were still referred (reasons not given) and underwent endoscopy (reasons not given), and so they suggest that repeat calprotectin testing of people with levels of < 150 μg/g should be considered.

Other tests

S100A12

The S100A12 protein is part of the S100 superfamily also known as calgranulin C. Unlike calprotectin (part of the S100A8/9 family), it is derived exclusively from granulocytes and monocytes.[182] This has resulted in the suggestion it is perhaps more specific than calprotectin in distinguishing inflammatory-related conditions compared with functional types, for example IBS.[82] However, both markers show correlation with endoscopic and histological inflammation.[183,184]

The reported ranges for sensitivity with S100A12 are 86–97%) and specificity 92–100%.[82,183,184] The cut-off most commonly used was 10 μg/g. However, it should be noted that, unlike calprotectin, most of the studies performed on S100A12 were in children and on a much smaller data set.

Calprotectin levels are consistently raised in children and fall to reach adult levels by the age of 5 years but there is a suggestion rather that S100A12 does not correlate with age, which may suggest an advantage in certain age groups.[82] These calgranulin peptides are also raised in pseudoinflammatory conditions, for example colorectal cancer, colorectal polyps and during use of non-steroidal anti inflammatory drugs. Such findings have been reported with calprotectin,[185,186] but little is known of its effects on expression of S100A12.

Thus if S100A12 were to be considered an alternative marker to distinguish between IBD and IBS, further larger studies would need to be performed (especially in adults) and to determine its expression in other gut-related conditions. One particular area would be to confirm reports that, unlike calprotectin, S100A12 level is elevated in bacterial gastroenteritis and not viral gastroenteritis, which may prove useful in clinical practice.[182]

Faecal haemoglobin

Mooiweer et al.[187] suggested that, at least in surveillance of IBD, faecal Hb might be an alternative to calprotectin. Their study was based on findings in 119 patients having surveillance colonoscopy. Faecal Hb was as accurate an indicator of inflammation as calprotectin. However, the mean calprotectin level in patients with active inflammation was 451 µg/g, so the spectrum of disease is rather different from the patients in the NICE decision problem group.

Lactoferrin

Lactoferrin was not included by NICE in the scope of this appraisal. There is less evidence on lactoferrin than on calprotectin. Testing for lactoferrin uses mainly ELISA methods but a point-of-care test is also available.[73] Lactoferrin is an iron-binding protein present in neutrophils.[188] Faecal lactoferrin (FL) is stable at room temperature for 4 days.[189] There are suggestions that FL is as good as FC for differentiating IBD from non-IBD.[189]

For differentiating IBD from IBS, the sensitivity of lactoferrin ranged between 78% and 82%, whereas the specificity ranged between 85% and 100% in studies.[188] Schoepfer et al.[75] compared the accuracy of faecal markers, blood leucocytes, CRP and IBD antibodies in discriminating IBD from IBS and found that the overall accuracies of FC ELISA (PhiCal test – 89%) and lactoferrin (IBD-SCAN – 90%) were similar. In another study[190] using colonoscopy as the reference standard, the sensitivity, specificity, PPV and diagnostic efficacy were slightly higher with lactoferrin (80%, 85%, 87% and 81%, respectively) than with calprotectin (78%, 83%, 86% and 80%).

Otten et al.[73] compared the diagnostic performance of the two new rapid tests for FL and FC against ELISA, and also assessed their potential to differentiate IBD from IBS. The sensitivity and NPV for the FC rapid test were higher than the FL rapid test (100% vs. 78%, 100% vs. 95%), whereas specificity and PPV were higher for the FL rapid test than the FC rapid test (99%, 95% vs. 95%, 82%). The diagnostic accuracy for both rapid tests was similar to ELISA tests (Cohen's kappa = 0.69 for FC, 0.68 for lactoferrin). Schroder et al.[77] found that the sensitivity with FC was better than FL (93% vs. 82%), whereas specificity was 100% for both. One study (Joishy et al. 2009[191]) compared the use of FC and FL as non-invasive markers in children and young adults with IBD, and found that using both FC and FL as diagnostic tests was better than using them in isolation. In contrast, another study (Schroder et al.[77]) found that using two tests together did not provide additional benefit.

The above findings suggest that FL can be used to differentiate IBD from IBS. The new FL rapid test seems to be as good as the ELISA in differentiating IBD from IBS thus there may be a place for this test in primary care. However, further research in primary care populations should compare the rapid FL test against rapid FC tests.

Neopterin

Nancey et al.[192] reported a comparison of neopterin and calprotectin, with endoscopic scores. Both distinguished active from inactive IBD, although with better accuracy in UC than CD. Neopterin was as accurate as calprotectin.

Ongoing research and research needs

Uncertainty remains as to how to deal with patients who have calprotectin levels of between 50 and 200 µg/g. One option is repeat testing after 6–8 weeks. If this group includes some people with IBD, is it a mild form of IBD?

Some of the parameters used in the economic analysis had to be based on assumptions or sensitivity analyses. Uncertainties include:

- What proportion of people with symptoms will have calprotectin level tested by GPs?
- Will the proportion of people with IBS that currently manage their symptoms without consulting their GPs, change, if there is publicity about a new test?
- What harm might accrue from false negatives? Will physicians whether in primary or secondary care, be too reassured by a negative calprotectin test?
- What proportion of patients with negative calprotectin levels will still be referred?

Arasaradnam et al.[193] from Coventry and Rotherham have pioneered an 'electric nose' test for IBD detection using urine testing to distinguish between those in remission and those not, and between those with IBD and healthy controls. This could potentially be used for diagnosis but requires a comparison of people with newly presenting IBD and those with IBS. The rationale behind this test is that abnormal gut permeability allows fermentation breakdown products into the bloodstream and hence into urine.

A Canadian group is carrying the FOCUS study (The Future of FC Utility Study: NCT0167324) to find out how often calprotectin results would change management of patients.[194]

Some patients with IBS (diagnosed after negative endoscopies) have raised calprotectin levels. The reasons for this are not known, and research into this group may be indicated. It may be due to an inflammatory component in some patients with IBS, perhaps especially those whose IBS follows GI infection.

Comparative data on the relative performance of the point-of-care tests, including in the intermediate 50 to 200 µg/g zone, are required.

Conclusions

Faecal calprotectin testing appears to be a useful method of distinguishing between inflammatory and non-inflammatory chronic bowel disease.

The current evidence base does not suggest any preference for any test over the others on diagnostic grounds. Relative cost will be more important in choice of test. The test kits are not expensive, and relative cost may depend more on labour costs and local discounts.

The NICE scope raised questions, abbreviated in italic text below.

Is calprotectin testing a reliable way of differentiating inflammatory disease of the bowel from non-inflammatory ones?

Yes. FC testing identifies patients with inflammation of the bowel, who need referral to specialist care. The majority of younger adult patients seen with lower abdominal symptoms in general practice have IBS, and the absence of inflammation as indicated by a negative calprotectin test means that IBD is very unlikely. They can then be managed in primary care and spared further investigations.

What are the optimal cut-offs for use in primary and secondary care

The same cut-off should be used in primary and secondary care: 50 µg/g. This is based on ensuring high sensitivity, and not missing people with IBD. Some people assessed as positive by this cut-off will have borderline levels of 50–150 µg/g, and may initially be monitored with repeat calprotectin testing, but some of this group will progress to definite IBD.

How do the rapid point-of-care tests compare to the laboratory tests?

There are few studies directly comparing tests, and on clinical effectiveness grounds, there is insufficient evidence to recommend one test over the others. The point-of-care tests can provide faster results. Costs vary among tests. None of the test kits is expensive but labour costs vary. The evidence base varies amongst tests. There are currently no grounds on either diagnostic reliability or cost-effectiveness considerations for preferring one test over another.

How will calprotectin testing perform in primary care?

Sensitivity and specificity will be as good in primary care, but the lower prevalence will increase the NPV. The main benefit in primary care could be to confirm the clinical diagnosis by GPs of IBS. Making calprotectin testing available to GPs could greatly reduce the number of younger adults referred to specialist care, and the need for invasive investigations, such as colonoscopy.

Impact in secondary care

In secondary care, the main benefit should be a marked reduction in colonoscopies that find no abnormalities. Calprotectin testing could considerably reduce the number of colonoscopies required. In various studies, over 60% of colonoscopies in this group of adult patients have found no abnormalities.

Calprotectin testing could lead to considerable savings to the NHS, as well as the avoidance of an unpleasant invasive procedure in people whose symptoms are due to IBS.

Calprotectin testing could also reduce the need for colonoscopy in children who do not have IBD, and could reduce diagnostic delays in those who do. It could also reduce loss of work time for parents and loss of school time for children.

Acknowledgements

We thank the following people for useful advice or information:

- Dr Hazel Borthwick.
- Dr Anjan Dhar.*
- Dr Paul Henderson.
- Dr Gurleen Juhti, NICE.
- Dr Robert Logan.*
- Barbara Mascialino and Steve Ferguson, Thermo Fisher Scientific.
- Dr John O'Malley.*
- Dr Nick Read.*
- Dr Jeremy Tibble.
- Dr Simon Whitehead.*
- Professor David Wilson.
- The Rotherham group: Dr A Kumar, Dr A Basumani, Professor K Bardhan.

*Indicates members of the NICE expert advisory group for this appraisal.

And we thank the following for commenting on sections of a near-final draft of this report:

- Dr Robert Logan.
- Dr Nick Read.
- Dr Marie Westhouse and Nigel Armstrong, Kleijnen Systematic Reviews, York.
- Professor David Wilson, University of Edinburgh.

Contributions of authors

References

1. Thompson WEG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Müller-Lissner SA. Functional bowel disorders and functional abdominal pain. *Gut* 1999;**45**(Suppl. II):43–7. http://dx.doi.org/10.1136/gut.45.2008.ii43

2. Longstreth GF, Thompson WG, Chey WD, Heaton KW, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology* 2006;**130**:1480–91. http://dx.doi.org/10.1053/j.gastro.2005.11.061

3. Ford AC, Talley NJ. Irritable bowel syndrome. *BMJ* 2012;**345**:e5836. http://dx.doi.org/10.1136/bmj.e5836

4. Wang YT, Lim HY, Tai D, Krishnamoorthy TL, Tan T, Barbier S, *et al.* The impact of irritable bowel syndrome on health-related quality of life: a Singapore perspective. *BMC Gastroenterol* 2012;**12**:104. http://dx.doi.org/10.1186/1471-230X-12-104

5. Spiegel B, Harris L, Lucak S, Mayer E, Naliboff B, Bolus R, *et al.* Developing valid and reliable health utilities in irritable bowel syndrome: results from the IBS PROOF Cohort. *Am J Gastroenterol* 2009;**104**:1984–91. http://dx.doi.org/10.1038/ajg.2009.232

6. Maxion-Bergemann S, Thielecke F, Abel F, Bergemann R. Costs of irritable bowel syndrome in the UK and US. *Pharmacoeconomics* 2006;**24**:21–37. http://dx.doi.org/10.2165/00019053-200624010-00002

7. Akehurst RL, Brazier JE, Mathers N, O'Keefe C, Kaltenthaler E, Morgan A, *et al.* Health-related quality of life and cost impact of irritable bowel syndrome in a UK primary care setting. *Pharmacoeconomics* 2002;**20**:455–62. http://dx.doi.org/10.2165/00019053-200220070-00003

8. Badia X, Mearin F, Balboa A, Baro E, Caldwell E, Cucala M, *et al.* Burden of illness in irritable bowel syndrome comparing Rome I and Rome II criteria. *Pharmacoeconomics* 2002;**20**:749–58. http://dx.doi.org/10.2165/00019053-200220110-00004

9. British Society of Gastroenterology. *Chronic management: IBS/Functional symptoms*. 2009. URL: www.bsg.org.uk/clinical/commissioning-report/ibs/functional-symptoms.html (accessed 19 July 2013).

10. British Society of Gastroenterology. *Clinical Commissioning Report: Diagnosis Chonic Abdominal discomfort, Diarrhoea and Constipation*. 2013. URL: www.bsg.org.uk/clinical/commissioning-report/chonic-abdominal-discomfort-diarrhoea-and-constipation.html (accessed 19 July 2013).

11. Degen A, Buning C, Siegmund B, Prager M, Maul J, Preiss J, *et al. Reasons for a delayed diagnosis in inflammatory bowel disease*. ECCO 2013. Congress Abstracts. 2013. https://www.ecco-ibd.eu/publications/congress-abstract-s.html (accessed 19 July 2013).

12. Ordas I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ. Ulcerative colitis. *Lancet* 2012;**380**:1606–19. http://dx.doi.org/10.1016/S0140-6736(12)60150-0

13. Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet* 2012;**380**:1590–605.

14. National Institute for Health and Care Excellence (NICE). *Crohn's disease – infliximab (review) and adalimumab (review of TA40)*. Technical assessment 187. 2010. http://guidance.nice.org.uk/TA187 (accessed 19 July 2013).

15. Henderson P, Hansen R, Cameron FL, Gerasimidis K, Rogers P, Bisset WM, *et al.* Rising incidence of pediatric inflammatory bowel disease in Scotland. *Inflamm Bowel Dis* 2012;**18**:999–1005. http://dx.doi.org/10.1002/ibd.21797

16. Aldhous MC, Armitage E, Drummond H, Satsangi J. Evidence of north–south gradient in incidence of juvenile: onset of inflammatory bowel disease in Scotland. *Gastroenterology* 2003;**124**:A213. http://dx.doi.org/10.1016/S0016-5085(03)81071-5

17. Henderson P, Wilson DC. The rising incidence of paediatric-onset inflammatory bowel disease. *Arch Dis Child* 2012;**97**:585–6. http://dx.doi.org/10.1136/archdischild-2012-302018

18. Shamir R, Phillip M, Levine A. Growth retardation in pediatric Crohn's disease: pathogenesis and interventions. *Inflammatory Bowel Dis* 2007;**13**:620–8.

19. Ford AC, Moayyedi P, Hanauer SB. Ulcerative colitis. *BMJ* 2013;**346**:f432. http://dx.doi.org/10.1136/bmj.f432

20. Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* 2011;**140**:1785–94. http://dx.doi.org/10.1053/j.gastro.2011.01.055

21. Gower-Rousseau C, Dauchet L, Vernier-Massouille G, Tilloy E, Brazier F, Merle V, *et al.* The natural history of pediatric ulcerative colitis: a population-based cohort study. *Am J Gastroenterol* 2009;**104**:2080–8. http://dx.doi.org/10.1038/ajg.2009.177

22. Sprakes MB, Ford AC, Suares NC, Warren L, Greer D, Donnellan CF, *et al.* Costs of care for Crohn's disease following the introduction of infliximab: a single-centre UK experience. *Aliment Pharmacol Therapeut* 2010;**32**:1357–63. http://dx.doi.org/10.1111/j.1365-2036.2010.04482.x

23. Waters HC, Vanderpoel JE, Nejadnik B, McKenzie RS, Lunacsek OE, Lennert BJ, *et al.* Resource utilization before and during infliximab therapy in patients with inflammatory bowel disease. *J Med Econ* 2012;**15**:45–52. http://dx.doi.org/10.3111/13696998.2011.625746

24. Butcher RO, Mehta SJ, Ahmad OF, Boyd CA, Anand R, Stein J, *et al.* Incidental diagnosis of inflammatory bowel disease in a British bowel cancer screening cohort: a multi-centre study. *J Crohns Colitis* 2013;**7**:S97–S98. http://dx.doi.org/10.1016/S1873-9946(13)60242-1

25. Esch A, Bourrier A, Seksik P, Nion-Larmurier I, Sokol H, Beaugerie L, *et al.* What is the prognosis of silent Crohn's disease? *J Crohns Colitis* 2013;**7**:S55. http://dx.doi.org/10.1016/S1873-9946(13)60138-5

26. Schoepfer AM, Vavricka S, Zahnd-Straumann N, Straumann A, Beglinger C. Monitoring inflammatory bowel disease activity: clinical activity is judged to be more relevant than endoscopic severity or biomarkers. *J Crohns Colitis* 2012;**6**:412–18. http://dx.doi.org/10.1016/j.crohns.2011.09.008

27. NZ IBD Research Review. *Inflammatory Bowel Disease*. 2010. URL: www.researchreview.co.nz/nz/Clinical-Area/Internal-Medicine/Gastroenterology/Inflammatory-Bowel-Disease/NZ-Inflammatory-Bowel-Disease-Research-Review-(7).aspx (accessed 19 July 2013).

28. Ananthakrishnan AN, Korzenik JR, Hur C. Can mucosal healing be a cost-effective endpoint for biologic therapy in Crohn's disease? A decision analysis. *Inflamm Bowel Dis* 2013;**19**:37–44.

29. Shivananda S, Lennard-Jones J, Logan R, Fear N, Price A, Carpenter L, *et al.* Incidence of inflammatory bowel disease across Europe: is there a difference between north and south? Results of the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Gut* 1996;**39**:690–7. http://dx.doi.org/10.1136/gut.39.5.690

30. Henderson P, Casey A, Lawrence SJ, Kennedy NA, Kingstone K, Rogers P, *et al.* The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease. *Am J Gastroenterol* 2012;**107**:941–9. http://dx.doi.org/10.1038/ajg.2012.33

31. Nielsen OH, Vainer B, Schaffalitzky de Muckadell OB. Microscopic colitis: a missed diagnosis? *Lancet* 2004;**364**:2055–7. http://dx.doi.org/10.1016/S0140-6736(04)17518-1

32. Ford AC, Talley NJ, Veldhuyzen van Zanten SJ, Vakil NB, Simel DL, Moayyedi P. Will the history and physical examination help establish that irritable bowel syndrome is causing this patient's lower gastrointestinal tract symptoms? *JAMA* 2008;**300**:1793–805. http://dx.doi.org/10.1001/jama.300.15.1793

33. Jellema P, van der Windt DA, Schellevis FG, van der Horst HE. Systematic review: accuracy of symptom-based criteria for diagnosis of irritable bowel syndrome in primary care. *Aliment Pharmacol Ther* 2009;**30**:695–706. http://dx.doi.org/10.1111/j.1365-2036.2009.04087.x

34. National Institute for Health and Care Excellence (NICE). *Faecal calprotectin diagnostic tests for identifying inflammatory bowel disease: final scope*. 2012. URL: www.nice.org.uk/nicemedia/live/13789/61043/61043.pdf (accessed 19 July 2013).

35. National Institute for Health and Care Excellence (NICE). Irritable bowel syndrome. *Clinical guideline 61*. 2008. URL: www.nice.org.uk/CG061 (accessed 19 July 2013).

36. NHS Technology Adoption Centre (NTAC). *Faecal calprotectin testing in primary care report*. 2012. URL: www.ntac.nhs.uk/web/FILES/Publications/Faecal_Calprotectin_Report_Final.pdf (accessed November 2012).

37. Thompson WG, Heaton KW, Smyth GT, Smyth C. Irritable bowel syndrome: the view from general practice. *Eur J Gastroenterol Hepatol* 1997;**9**:689–92. http://dx.doi.org/10.1097/00042737-199707000-00008

38. Ramadas A, Datta K, Sunderraj L, Hawthorne AB. Faecal calprotectin as part of standardised assessment tool of suspected irritable bowel syndrome (IBS): a novel investigative algorithm. *Gut* 2011;**60**:A73–4. http://dx.doi.org/10.1136/gut.2011.239301.150

39. Kok L, Elias SG, Witteman BJ, Goedhard JG, Muris JW, Moons KG, *et al.* Diagnostic accuracy of point-of-care fecal calprotectin and immunochemical occult blood tests for diagnosis of organic bowel disease in primary care: the Cost-Effectiveness of a Decision Rule for Abdominal Complaints in Primary Care (CEDAR) study. *Clin Chem* 2012;**58**:989–98. http://dx.doi.org/10.1373/clinchem.2011.177980

40. Burgmann T, Clara I, Graff L, Walker J, Lix L, Rawsthorne P, *et al.* The Manitoba Inflammatory Bowel Disease Cohort Study: prolonged symptoms before diagnosis: how much is irritable bowel syndrome? *Clin Gastroenterol Hepatol* 2006;**4**:614–20. http://dx.doi.org/10.1016/j.cgh.2006.03.003

41. Porter CK, Cash BD, Pimentel M, Akinseye A, Riddle MS. Risk of inflammatory bowel disease following a diagnosis of irritable bowel syndrome. *BMC Gastroenterol* 2012;**12**:55. http://dx.doi.org/10.1186/1471-230X-12-55

42. Naismith GD, Smith LA, Barry SJ, Munro JI, Laird S, Rankin K, *et al.* A prospective single-centre evaluation of the intra-individual variability of faecal calprotectin in quiescent Crohn's disease. *Aliment Pharmacol Ther* 2013;**37**:613–21. http://dx.doi.org/10.1111/apt.12221

43. Moum B, Jahnsen J, Bernklev T. Fecal calprotectin variability in Crohn's disease. *Inflamm Bowel Dis* 2010;**16**:1091–2.

44. Haapamaki J, Tanskanen A, Roine RP, Blom M, Turunen U, Mantyla J, *et al.* Medication use among inflammatory bowel disease patients: excessive consumption of antidepressants and analgesics. *Scand J Gastroenterol* 2013;**48**:42–50. http://dx.doi.org/10.3109/00365521.2012.743584

45. Demir OM, Appleby R, Wang Y, Logan RPH. What factors might contribute to borderline faecal calprotectin levels? *J Crohns Colitis* 2012;**6**:S56. http://dx.doi.org/10.1016/S1873-9946(12)60134-2

46. Turvill J. High negative predictive value of a normal faecal calprotectin in patients with symptomatic intestinal disease. *Frontline Gastroenterol* 2012;**3**:21–8. http://dx.doi.org/10.1136/flgastro-2011-100011

47. Whitehead WE, Palsson OS, Feld AD, Levy RL, von Korff M, Turner MJ, *et al.* Utility of red flag symptom exclusions in the diagnosis of irritable bowel syndrome. *Aliment Pharmacol Ther* 2006;**24**:137–46. http://dx.doi.org/10.1111/j.1365-2036.2006.02956.x

48. Dolwani S, Metzner M, Wassell JJ, Yong A, Hawthorne AB. Diagnostic accuracy of faecal calprotectin estimation in prediction of abnormal small bowel radiology. *Aliment Pharmacol Therapeut* 2004;**20**:615–21. http://dx.doi.org/10.1111/j.1365-2036.2004.02128.x

49. Tibble JA, Sigthorsson G, Foster R, Forgacs I, Bjarnason I. Use of surrogate markers of inflammation and Rome criteria to distinguish organic from non-organic intestinal disease. *Gastroenterology* 2002;**123**:450–60.

50. York Health Economics Consortium (YHEC). *Economic report: Value of calprotectin in screening out irritable bowel syndrome: CEP09041*. 2010. URL: www.alphalabs.co.uk/asp/db/documents/cep09026_Calprotectin%20Economic_Jan10.pdf (accessed 19 July 2013).

51. Mascialino B, Hermansson L-L, Larsson A. Is the IBD pre-endoscopic screening F-Calprotectin test more cost-effective than the usage of serologic markers in selected European markets? *J Crohns Colitis* 2013;**7**:S91.

52. Burri E, Beglinger C. Faecal calprotectin: a useful tool in the management of inflammatory bowel disease. *Swiss Med Wkly* 2012;**142**:w13557. http://dx.doi.org/10.4414/smw.2012.13557

53. Van Rheenen PF, Van de Vijver Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ* 2010;**341**:c3369. http://dx.doi.org/10.1136/bmj.c3369

54. Van Roon AC, Karamountzos L, Purkayastha S, Reese GE, Darzi AW, Teare JP, *et al.* Diagnostic precision of fecal calprotectin for inflammatory bowel disease and colorectal malignancy. *Am J Gastroenterol* 2007;**102**:803–13. http://dx.doi.org/10.1111/j.1572-0241.2007.01126.x

55. Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J* 2009;**9**:211–29.

56. Macaskill P, Gatsonis C, Deeks J, Harbord RM, Takwoingi Y. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. Analysing and Presenting Results. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy,* version 1.0. The Cochrane Collaboration 2010.

57. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, 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.

58. Damms A, Bischoff SC. Validation and clinical significance of a new calprotectin rapid test for the diagnosis of gastrointestinal diseases. *Int J Colorectal Dis* 2008;**23**:985–92. http://dx.doi.org/10.1007/s00384-008-0506-0

59. Garcia SM V, Gonzalez R, Iglesias FE, Gomez CF, Casais JL, Cerezo RA, *et al.* [Diagnostic value of fecal calprotectin in predicting an abnormal colonoscopy.] *Med Clin* 2006;**127**:41–6.

60. Li XG, Lu YM, Gu F, Yang XL. [Fecal calprotectin in differential diagnosis of irritable bowel syndrome.] *Beijing da Xue Xue Bao Yi Xue Ban/Journal of Peking University Health Sciences* 2006;**38**:310–13.

61. Licata A, Randazzo C, Cappello M, Calvaruso V, Butera G, Florena AM, *et al.* Fecal calprotectin in clinical practice a noninvasive screening tool for patients with chronic diarrhea. *J Clin Gastroenterol* 2012;**46**:504–8. http://dx.doi.org/10.1097/MCG.0b013e318248f289

62. Shitrit AB, Braverman D, Stankiewics H, Shitrit D, Peled N, Paz K. Fecal calprotectin as a predictor of abnormal colonic histology. *Dis Colon Rectum* 2007;**50**:2188–93. http://dx.doi.org/10.1007/s10350-007-9038-x

63. Zippi M, Al AN, Siliquini F, Severi C, Kagarmanova A, Maffia C, *et al.* Correlation between faecal calprotectin and magnetic resonance imaging (MRI) in the evaluation of inflammatory pattern in Crohn's disease. *Clin Ter* 2010;**161**:e53–6.

64. Aomatsu T, Yoden A, Matsumoto K, Kimura E, Inoue K, Andoh A, *et al.* Fecal calprotectin is a useful marker for disease activity in pediatric patients with inflammatory bowel disease. *Dig Dis Sci* 2011;**56**:2372–7. http://dx.doi.org/10.1007/s10620-011-1633-y

65. Canani RB, de Horatio LT, Terrin G, Romano MT, Miele E, Staiano A, *et al.* Combined use of noninvasive tests is useful in the initial diagnostic approach to a child with suspected inflammatory bowel disease. *JPGN* 2006;**42**:9–15. http://dx.doi.org/10.1097/01.mpg.0000187818.76954.9a

66. Tomas AB, Vidal MV, Camps R. Fecal calprotectin as a biomarker to distinguish between organic and functional gastrointestinal disease. *Rev Esp Enferm Dig* 2007;**99**:689–93.

67. Harbord R. *Meta-analysis of diagnostic test accuracy studies: metandi&midas*. 2010. URL: www.STATA.com/meeting/uk10/UKSUG10.Harbord.pdf (accessed 19 July 2013).

68. Jellema P, van Tulder MW, van der Horst HE, Florie J, Mulder CJ, van der Windt DA. Inflammatory bowel disease: a systematic review on the value of diagnostic testing in primary care. *Colorectal Dis* 2011;**13**:239–54. http://dx.doi.org/10.1111/j.1463-1318.2009.02131.x

69. Kostakis ID, Cholidou KG, Vaiopoulos AG, Vlachos IS, Perrea D, Vaos G. Fecal calprotectin in pediatric inflammatory bowel disease: a systematic review. *Dig Dis Sci* 2013;**58**:309–19. doi:10.1007/s10620-012-2347-5.

70. Henderson P, Anderson NH, Wilson DC. The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease: a systematic review and meta-analysis [published online ahead of print May 14, 2013]. *Am J Gastroenterol* 2013. doi:10.1038/ajg.2013.131.

71. Bunn SK, Bisset WM, Main MJ, Gray ES, Olson S, Golden BE. Fecal calprotectin: validation as a noninvasive measure of bowel inflammation in childhood inflammatory bowel disease. *JPGN* 2001;**33**:14–22. http://dx.doi.org/10.1097/00005176-200107000-00003

72. Kolho K, Turner D, Veereman-Wauters G, Sladek M, de RL, Shaoul R, *et al.* Rapid test for fecal calprotectin levels in children with Crohn disease. *J Pediatr Gastroenterol Nutr* 2012;**55**:436–9. http://dx.doi.org/10.1097/MPG.0b013e318253cff1

73. Otten CM, Kok L, Witteman BJ, Baumgarten R, Kampman E, Moons KG, *et al.* Diagnostic performance of rapid tests for detection of fecal calprotectin and lactoferrin and their ability to discriminate inflammatory from irritable bowel syndrome. *Clin Chem Lab Med* 2008;**46**:1275–80. [Erratum appears in *Clin Chem Lab Med* 2008;**46**:1798.] http://dx.doi.org/10.1515/CCLM.2008.361

74. Schoepfer AM, Trummler M, Seeholzer P, Criblez DH, Seibold F. Accuracy of four fecal assays in the diagnosis of colitis. *Dis Colon Rectum* 2007;**50**:1697–706. http://dx.doi.org/10.1007/s10350-007-0303-9

75. Schoepfer AM, Trummler M, Seeholzer P, Seibold-Schmid B, Seibold F. Discriminating IBD from IBS: comparison of the test performance of fecal markers, blood leukocytes, CRP, and IBD antibodies. *Inflamm Bowel Dis* 2008;**14**:32–9. http://dx.doi.org/10.1002/ibd.20275

76. Tibble JA, Sigthorsson G, Bridger S, Fagerhol MK, Bjarnason I. Surrogate markers of intestinal inflammation are predictive of relapse in patients with inflammatory bowel disease. *Gastroenterology* 2000;**119**:15–22. http://dx.doi.org/10.1053/gast.2000.8523

77. Schroder O, Naumann M, Shastri Y, Povse N, Stein J. Prospective evaluation of faecal neutrophil-derived proteins in identifying intestinal inflammation: combination of parameters does not improve diagnostic accuracy of calprotectin. *Aliment Pharmacol Therapeut* 2007;**26**:1035–42. http://dx.doi.org/10.1111/j.1365-2036.2007.03457.x

78. Limburg PJ, Ahlquist DA, Sandborn WJ, Mahoney DW, Devens ME, Harrington JJ, *et al.* Fecal calprotectin levels predict colorectal inflammation among patients with chronic diarrhea referred for colonoscopy. *Am J Gastroenterol* 2000;**95**:2831–7. http://dx.doi.org/10.1016/S0002-9270(00)01987-0

79. Perminow G, Brackmann S, Lyckander LG, Franke A, Borthne A, Rydning A, *et al.* A characterization in childhood inflammatory bowel disease, a new population-based inception cohort from South-Eastern Norway, 2005–7, showing increased incidence in Crohn's disease. *Scand J Gastroenterol* 2009;**44**:446–56.

80. Van de Vijver E, Schreuder AB, Cnossen WR, Muller Kobold AC, van Rheenen PF. Safely ruling out inflammatory bowel disease in children and teenagers without referral for endoscopy [published online ahead of print 2012]. *Arch Dis Child* 2012;**12**:1014–18. http://dx.doi.org/10.1136/archdischild-2011-301206

81. El-Badry A, Sedrak H, Rashed L. Faecal calprotectin in differentiating between functional and organic bowel diseases. *Arab J Gastroenterol* 2010;**11**:70–3. http://dx.doi.org/10.1016/j.ajg.2010.04.009

82. Sidler MA, Leach ST, Day AS. Fecal S100A12 and fecal calprotectin as noninvasive markers for inflammatory bowel disease in children. *Inflammatory Bowel Dis* 2008;**14**:359–66. http://dx.doi.org/10.1002/ibd.20336

83. Basumani P, Bardhan K, Eyre R, Ellis R, The Rotherham Team. Faecal calprotectin: Rotherham experience (slide presentation). BSG Away day 28/06/2012 (accessed 19 July 2013). Unpublished

84. Manz M, Burri E, Rothen C, Tchanguizi N, Niederberger C, Rossi L, *et al.* Value of fecal calprotectin in the evaluation of patients with abdominal discomfort: an observational study. *BMC Gastroenterology* 2012;**12**:5. http://dx.doi.org/10.1186/1471-230X-12-5

85. Burri E, Manz M, Rothen C, Rossi L, Beglinger C, Lehmann FS. Monoclonal antibody testing for fecal calprotectin is superior to polyclonal testing of fecal calprotectin and lactoferrin to identify organic intestinal disease in patients with abdominal discomfort. *Clin Chim Acta* 2013;**416**:41–7. http://dx.doi.org/10.1016/j.cca.2012.11.008

86. Diamanti A, Panetta F, Basso MS, Forgione A, Colistro F, Bracci F, *et al.* Diagnostic work-up of inflammatory bowel disease in children: the role of calprotectin assay. *Inflamm Bowel Dis* 2010;**16**:1926–30. http://dx.doi.org/10.1002/ibd.21257

87. Carroccio A, Iacono G, Cottone M, Di PL, Cartabellotta F, Cavataio F, *et al.* Diagnostic accuracy of fecal calprotectin assay in distinguishing organic causes of chronic diarrhea from irritable bowel syndrome: a prospective study in adults and children. *Clin Chem* 2003;**49**:t–7. http://dx.doi.org/10.1373/49.6.861

88. Bharathi S, Moncur P, Holbrook IB, Kelly S. A normal facael calprotectin has a high negative predictive value in cases of suspected irritable bowel syndrome. *Gastroenterology* 2005;**128**:A459.

89. Ashorn S, Honkanen T, Kolho KL, Ashorn M, Valineva T, Wei B, *et al.* Fecal calprotectin levels and serological responses to microbial antigens among children and adolescents with inflammatory bowel disease. *Inflamm Bowel Dis* 2009;**15**:199–205. http://dx.doi.org/10.1002/ibd.20535

90. Lee H, Mainman H, Borthwick H, Dhar A. Faecal calprotectin testing in primary and secondary care – are the current manufacturer's cut-off levels clinically useful? *Unpublished data BSG 2013. Abstract submission* 2013.

91. Fagerberg UL, Loof L, Myrdal U, Hansson LO, Finkel Y. Colorectal inflammation is well predicted by fecal calprotectin in children with gastrointestinal symptoms. *JPGN* 2005;**40**:450–5. http://dx.doi.org/10.1097/01.MPG.0000154657.08994.94

92. Shaoul R, Sladek M, Turner D, Paeregaard A, Veres G, Wauters GV, *et al.* Limitations of fecal calprotectin at diagnosis in untreated pediatric Crohn's disease. *Inflamm Bowel Dis* 2012;**18**:1493–7. http://dx.doi.org/10.1016/S1873-9946(12)60116-0

93. Bremner A, Roked S, Robinson R, Phillips I, Beattie M. Faecal calprotectin in children with chronic gastrointestinal symptoms. *Acta Paediatr* 2005;**94**:1855–8. http://dx.doi.org/10.1080/08035250500254639

94. Komraus M, Wos H, Wiecek S, Kajor M, Grzybowska-Chlebowczyk U. Usefulness of faecal calprotectin measurement in children with various types of inflammatory bowel disease. *Mediators Inflamm* 2012. doi:10.1155/2012/608249.

95. Wassell J, Wallage M, Brewer E. Evaluation of the Quantum Blue rapid test for faecal calprotectin. *Ann Clin Biochem* 2012;**49**:55–8. http://dx.doi.org/10.1258/acb.2011.011106

96. Dolci A, Panteghini M. Comparative study of a new quantitative rapid test with an established ELISA method for faecal calprotectin. *Clin Chim Acta* 2012;**413**:350–1. http://dx.doi.org/10.1016/j.cca.2011.09.030

97. Coorevits L, Baert F, Vanpoucke H. Comparative study of the Quantum Blue rapid test for point of care quantification of faecal calprotectin with an established ELISA method. *J Crohns Colitis* 2012;**6**:S45. http://dx.doi.org/10.1016/S1873-9946(12)60107-X

98. Hessels J, Douw G, Yildirim DD, Meerman G, Van Herwaarden MA, Van Den Bergh FAJT. Evaluation of Prevent ID and Quantum Blue rapid tests for fecal calprotectin. *Clin Chem Lab Med* 2012;**50**:1079–82. http://dx.doi.org/10.1515/cclm-2011-0855

99. Sydora MJ, Sydora BC, Fedorak RN. Validation of a point-of-care desk top device to quantitate fecal calprotectin and distinguish inflammatory bowel disease from irritable bowel syndrome. *J Crohns Colitis* 2012;**6**:207–14. http://dx.doi.org/10.1016/j.crohns.2011.08.008

100. Vestergaard TA, Nielsen SL, Dahlerup JF, Hornung N. Fecal calprotectin: assessment of a rapid test. *Scand J Clin Lab Inv* 2008;**68**:343–7. http://dx.doi.org/10.1080/00365510701576198

101. Shastri Y, Schroder O, Stein J. Comparative evaluation of a new semi-quantitative, rapid, Office-based strip test with an ELISA-based assay for measuring fecal calprotectin to assess intestinal inflammation: prospective multicenter clinical study. *Gastroenterology* 2009;**136**:A34–5. http://dx.doi.org/10.1016/S0016-5085(09)60158-X

102. Labaere D, Smismans A, Van OG, Christiaens P, D'Haens G, Moons V, *et al.* Clinical comparison of eight different calprotectin immunoassays for the diagnosis and follow-up of inflammatory bowel disease. *J Crohns Colitis* 2013;**7**:S120. http://dx.doi.org/10.1016/S1873-9946(13)60299-8

103. Whitehead SJ, French J, Brookes MJ, Ford C, Gama R. Between-assay variability of faecal calprotectin enzyme-linked immunosorbent assay kits. *Ann Clin Biochem* 2013;**50**:53–61. http://dx.doi.org/10.1258/acb.2012.011272

104. Loitsch SM, Berger C, Hartmann F, Stein J. Comparison of two commercially available serologic kits for the detection of fecal calprotecin. *Gastroenterology* 2010;**138**:S528. http://dx.doi.org/10.1016/S0016-5085(10)62436-5

105. Tomkins C, Zeino Z, Nwokolo C, Smith SC, Arasaradnam R. Faecal calprotectin analysis: does the method matter? *Gut* 2012;**61**:A173–4. http://dx.doi.org/10.1136/gutjnl-2012-302514b.243

106. Spiller R, Aziz Q, Creed F, Emmanuel A, Houghton L, Hungin P, *et al.* Guidelines on the irritable bowel syndrome: mechanisms and practical management. *Gut* 2007;**56**:1770–98. http://dx.doi.org/10.1136/gut.2007.119446

107. Curtis L. Personal Social Services Research Unit (PSSRU): *Unit Costs of Health and Social Care.* 2011. URL: www.pssru.ac.uk/project-pages/unit-costs/2011/index.php (accessed 19 July 2013).

108. National Institute for Health and Care Excellence (NICE). Crohn's disease. Clinical guideline 152. 2012. URL: guidance.nice.org.uk/cg152 (accessed 19 July 2013).

109. National Institute for Health and Care Excellence (NICE). *Ulcerative colitis: management in adults, children and young people: NICE guideline*. Draft for consultation. 2013. URL: http://guidance.nice.org.uk/CG/Wave25/9/Consultation/Latest (accessed 19 July 2013).

110. Hornung T, Anwar GA. Faecal calprotectin: an audit to determine its clinical and cost effectiveness at North Tees and Hartlepool NHS Trust. *Gut* 2011;**60**:A203. http://dx.doi.org/10.1136/gut.2011.239301.427

111. Mindemark M, Larsson A. Ruling out IBD: estimation of the possible economic effects of pre-endoscopic screening with F-calprotectin. *Clin Biochem* 2012;**45**:552–5. http://dx.doi.org/10.1016/j.clinbiochem.2011.10.015

112. Goldfarb NI, Pizzi LT, Fuhr JP, Salvador C Jr, Sikirica V, Kornbluth A, *et al.* Diagnosing Crohn's disease: an economic analysis comparing wireless capsule endoscopy with traditional diagnostic procedures. *Dis Manag* 2004;**7**:292–304. http://dx.doi.org/10.1089/dis.2004.7.292

113. Mascialino B, Hermansson L, Meister S, Mummert E. Cost-effectiveness in diagnostic tests: comparison of the IBD pre-endoscopic screening F-Calprotectin test versus serologic markers in the United Kingdom. Poster presented at ECHE (European Confrence on Health Economics), July 2012, Zurich, Switzerland, (accessed 19 July 2013).

114. Dubinsky MC, Johanson JF, Seidman EG, Ofman JJ. Suspected inflammatory bowel disease: the clinical and economic impact of competing diagnostic strategies. *Am J Gastroenterol* 2002;**97**:2333–42. http://dx.doi.org/10.1016/S0002-9270(02)04345-9

115. Ricci JF, Walters S, Ward S, Akehurst R. Quality of life of IBS patients in the United Kingdom compared with a general population sample. *Gastroenterology* 2003;**124**:A396. http://dx.doi.org/10.1016/S0016-5085(03)82002-4

116. Bernklev T, Jahnsen J, Lygren I, Henriksen M, Vatn M, Moum B. Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: psychometric assessments and a comparison with general population norms. *Inflamm Bowel Dis* 2005;**11**:909–18. http://dx.doi.org/10.1097/01.mib.0000179467.01748.99

117. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;**13**:873–84. http://dx.doi.org/10.1002/hec.866

118. Bracco A, Jonsson B, Ricci JF, Drummond M, Nyhlin H. Economic evaluation of tegaserod vs. placebo in the treatment of patients with irritable bowel syndrome: an analysis of the TENOR study. *Value Health* 2007;**10**:238–46. http://dx.doi.org/10.1111/j.1524-4733.2007.00179.x

119. Dibonaventura MD, Prior M, Prieto P, Fortea J. Burden of constipation-predominant irritable bowel syndrome (IBS-C) in France, Italy, and the United Kingdom. *Clin Exp Gastroenterol* 2012;**5**:203–12. http://dx.doi.org/10.2147/CEG.S35568

120. Pare P, Gray J, Lam S, Balshaw R, Khorasheh S, Barbeau M, *et al.* Health-related quality of life, work productivity, and health care resource utilization of subjects with irritable bowel syndrome: baseline results from LOGIC (Longitudinal Outcomes Study of Gastrointestinal Symptoms in Canada), a naturalistic study. *Clin Therapeut* 2006;**28**:1726–35. http://dx.doi.org/10.1016/j.clinthera.2006.10.010

121. Puhan MA, Schunemann HJ, Wong E, Griffith L, Guyatt GH. The standard gamble showed better construct validity than the time trade-off. *J Clin Epidemiol* 2007;**60**:1029–33. http://dx.doi.org/10.1016/j.jclinepi.2007.03.001

122. Konig HH, Ulshofer A, Gregor M, von Tirpitz C, Reinshagen M, Adler G, *et al.* Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2002;**14**:1205–15. http://dx.doi.org/10.1097/00042737-200211000-00008

123. Leidl R, Reitmeir P, Konig HH, Stark R. The performance of a value set for the EQ-5D based on experienced health states in patients with inflammatory bowel disease. *Value Health* 2012;**15**:151–7. http://dx.doi.org/10.1016/j.jval.2011.08.004

124. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflamm Bowel Dis* 2010;**16**:42–51. http://dx.doi.org/10.1002/ibd.20989

125. Turunen P, Ashorn M, Auvinen A, Iltanen S, Huhtala H, Kolho KL. Long-term health outcomes in pediatric inflammatory bowel disease: a population-based study. *Inflamm Bowel Dis* 2009;**15**:56–62. http://dx.doi.org/10.1002/ibd.20558

126. Casellas F, Arenas JI, Baudet JS, Fabregas S, Garcia N, Gelabert J, *et al.* Impairment of health-related quality of life in patients with inflammatory bowel disease: a Spanish multicenter study. *Inflamm Bowel Dis* 2005;**11**:488–96. http://dx.doi.org/10.1097/01.MIB.0000159661.55028.56

127. Badia X, Roset M, Montserrat S, Herdman M, Segura A. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale.] *Med Clin* 1999;**112**(Suppl. 1):79–85.

128. Bassi A, Bodger K. Health state utilities and willingness-to-pay in inflammatory bowel disease (IBD): a study of feasibility and validity. *Gastroenterology* 2005;**128**:A555.

129. Arseneau KO, Cohn SM, Cominelli F, Connors AF Jr., Cost-utility of initial medical management for Crohn's disease perianal fistulae. *Gastroenterology* 2001;**120**:1640–56. http://dx.doi.org/10.1053/gast.2001.24884

130. Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. *Value Health* 2007;**10**:214–20. http://dx.doi.org/10.1111/j.1524-4733.2007.00171.x

131. Benedini V, Caporaso N, Corazza GR, Rossi Z, Fornaciari G, Cottone M, *et al.* Burden of Crohn's disease: economics and quality of life aspects in Italy. *Clinicoecon Outcomes Res* 2012;**4**:209–18.

132. Casellas F, Rodrigo L, Nino P, Pantiga C, Riestra S, Malagelada JR. Sustained improvement of health-related quality of life in Crohn's disease patients treated with infliximab and azathioprine for 4 years. *Inflamm Bowel Dis* 2007;**13**:1395–400.

133. Gregor JC, McDonald JW, Klar N, Wall R, Atkinson K, Lamba B, *et al.* An evaluation of utility measurement in Crohn's disease. *Inflamm Bowel Dis* 1997;**3**:265–76. http://dx.doi.org/10.1002/ibd.3780030405

134. Gibson PR, Weston AR, Shann A, Florin TH, Lawrance IC, Macrae FA, *et al.* Relationship between disease severity, quality of life and health-care resource use in a cross-section of Australian patients with Crohn's disease. *J Gastroenterol Hepatol* 2007;**22**:1306–12. http://dx.doi.org/10.1111/j.1440-1746.2007.04930.x

135. Hill R, Lewindon P, Muir R, Grange I, Connor F, Ee L, *et al.* Quality of life in children with Crohn disease. *JPGN* 2010;**51**:35–40.

136. Connolly MP, Poole CD, Currie CJ, Marteau P, Nielsen SK. Quality of life improvements attributed to combination therapy with oral and topical mesalazine in mild-to-moderately active ulcerative colitis. *Digestion* 2009;**80**:241–6. http://dx.doi.org/10.1159/000235916

137. Bryan S, Andronis L, Hyde C, Connock M, Fry-Smith A, Wang D. *Infliximab for acute exacerbations of ulcerative colitis*. 2008. URL: www.hta.ac.uk/erg/reports/1730.pdf (accessed 19 July 2013).

138. Arseneau KO, Sultan S, Provenzale DT, Onken J, Bickston SJ, Foley E, *et al.* Do patient preferences influence decisions on treatment for patients with steroid-refractory ulcerative colitis? *Clin Gastroenterol Hepatol* 2006;**4**:1135–42. http://dx.doi.org/10.1016/j.cgh.2006.05.003

139. Feagan BG, Reinisch W, Rutgeerts P, Sandborn WJ, Yan S, Eisenberg D, *et al.* The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. *Am J Gastroenterol* 2007;**102**:794–802. http://dx.doi.org/10.1111/j.1572-0241.2007.01094.x

140. Waljee AK, Higgins PD, Waljee JF, Tujios SR, Saxena A, Brown LK, *et al.* Perceived and actual quality of life with ulcerative colitis: a comparison of medically and surgically treated patients. *Am J Gastroenterol* 2011;**106**:794–9. http://dx.doi.org/10.1038/ajg.2011.39

141. Poole CD, Connolly MP, Nielsen SK, Currie CJ, Marteau P. A comparison of physician-rated disease severity and patient reported outcomes in mild to moderately active ulcerative colitis. *J Crohns Colitis* 2010;**4**:275–82. http://dx.doi.org/10.1016/j.crohns.2009.11.010

142. Baudet JS, Diaz-Bethencourt D, Aviles J, Aguirre-Jaime A. Minor adverse events of colonoscopy on ambulatory patients: the impact of moderate sedation. *Eur J Gastroenterol Hepatol* 2009;**21**:656–61. http://dx.doi.org/10.1097/MEG.0b013e328314b7e3

143. de Jonge V, Nicolaas JS, van BO, Brouwer JT, Stolk MF, Tang TJ, *et al.* The incidence of 30-day adverse events after colonoscopy among outpatients in the Netherlands. *Am J Gastroenterol* 2012;**107**:878–84. http://dx.doi.org/10.1038/ajg.2012.40

144. Dominitz JA, Provenzale D. Patient preferences and quality of life associated with colorectal cancer screening. *Am J Gastroenterol* 1997;**92**:2171–8.

145. Niv Y, Bogolavski I, Ilani S, Avni I, Gal E, Vilkin A, *et al.* Impact of colonoscopy on quality of life. *Eur J Gastroenterol Hepatol* 2012;**24**:781–6. http://dx.doi.org/10.1097/MEG.0b013e328352deff

146. Spiegel BM, Gralnek IM, Bolus R, Chang L, Dulai GS, Naliboff B, *et al.* Is a negative colonoscopy associated with reassurance or improved health-related quality of life in irritable bowel syndrome? *Gastrointest Endosc* 2005;**62**:892–9. http://dx.doi.org/10.1016/j.gie.2005.08.016

147. Warren JL, Klabunde CN, Mariotto AB, Meekins A, Topor M, Brown ML, *et al.* Adverse events after outpatient colonoscopy in the Medicare population. *Ann Int Med* 2009;**150**:849–57. http://dx.doi.org/10.7326/0003-4819-150-12-200906160-00008

148. Levin TR, Zhao W, Conell C, Seeff LC, Manninen DL, Shapiro JA, *et al.* Complications of colonoscopy in an integrated health care delivery system. *Ann Intern Med* 2006;**145**:880–6. http://dx.doi.org/10.7326/0003-4819-145-12-200612190-00004

149. National Institute for Health and Care Excellence (NICE). *Colonoscopic surveillance for prevention of colorectal cancer in people with ulcerative colitis, Crohn's disease or adenomas.* Clinical guideline 118. 2011. URL: http://guidance.nice.org.uk/CG118 (accessed 19 July 2013).

150. Tappenden P, Eggington S, Nixon R, Chilcott J, Sakai H, Karnon J. *Colorectal cancer screening options appraisal: report to the English Bowel Cancer Screening Working Group*. 2004. URL: www.cancerscreening.nhs.uk/bowel/scharr.pdf (accessed 19 July 2013).

151. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, *et al.* Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010;**375**:1624–33. http://dx.doi.org/10.1016/S0140-6736(10)60551-X

152. Atkin WS, Cook CF, Cuzick J, Edwards R, Northover JM, Wardle J. Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial. *Lancet* 2002;**359**:1291–300.

153. 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**:230–6. http://dx.doi.org/10.1093/jnci/95.3.230

154. Department of Health (DoH). *NHS Reference Costs 20011–12*. London: DoH; 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213060/2011-12-reference-costs-publication.pdf (accessed January 2013).

155. Mearin F, Roset M, Badia X, Balboa A, Baro E, Ponce J, *et al.* Splitting irritable bowel syndrome: from original Rome to Rome II criteria. *Am J Gastroenterol* 2004;**99**:122–30. http://dx.doi.org/10.1046/j.1572-0241.2003.04024.x

156. Moroni F, Winter JW, Morris AJ, Gaya DR. What is the clinical relevance of a mildly elevated faecal calprotectin detected in new referrals to the gastroenterology clinic? *Gut* 2012;**61**:A78–9. http://dx.doi.org/10.1136/gutjnl-2012-302514b.15

157. Zayyat R, Appleby RN, Logan RPH. Can we improve the negative predictive value of faecal calprotectin for the diagnosis of ibs in primary care? *Gut* 2011;**60**:A49–50. http://dx.doi.org/10.1136/gut.2011.239301.97

158. Mohammed N, Smale S. Positive calprotectin but negative investigations-what next? *Gut* 2012;**61**:A236. http://dx.doi.org/10.1136/gutjnl-2012-302514c.125

159. Lee SH, Mainman H, Borthwick H, Dhar A. PTH-094 faecal calprotectin testing in primary and secondary care: are the current manufacturer's cut-off levels clinically useful? *Gut* 2013;**62**:A249. http://dx.doi.org/10.1136/gutjnl-2013-304907.581

160. Demir OM, Ahmed Z, Logan RPH. Optimising the use of faecal calprotectin for early diagnosis of IBD in primary care. *J Crohns Colitis* 2013;**7**:S8–9. http://dx.doi.org/10.1016/S1873-9946(13)60018-5

161. Koulaouzidis A, Douglas S, Rogers MA, Arnott ID, Plevris JN. Fecal calprotectin: a selection tool for small bowel capsule endoscopy in suspected IBD with prior negative bi-directional endoscopy. *Scand J Gastroenterol* 2011;**46**:561–6. http://dx.doi.org/10.3109/00365521.2011.551835

162. Vavricka SR, Spigaglia SM, Rogler G, Pittet V, Michetti P, Felley C, *et al.* Systematic evaluation of risk factors for diagnostic delay in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:496–505. http://dx.doi.org/10.1002/ibd.21719

163. Peyrin-Biroulet L, Loftus EV, Jr, Colombel JF, Sandborn WJ. The natural history of adult Crohn's disease in population-based cohorts. *Am J Gastroenterol* 2010;**105**:289–97. http://dx.doi.org/10.1038/ajg.2009.579

164. Farrokhyar F, Marshall JK, Easterbrook B, Irvine EJ. Functional gastrointestinal disorders and mood disorders in patients with inactive inflammatory bowel disease: prevalence and impact on health. *Inflamm Bowel Dis* 2006;**12**:38–46. http://dx.doi.org/10.1097/01.MIB.0000195391.49762.89

165. Keohane J, O'Mahony C, O'Mahony L, O'Mahony S, Quigley EM, Shanahan F. Irritable bowel syndrome-type symptoms in patients with inflammatory bowel disease: a real association or reflection of occult inflammation? *Am J Gastroenterol* 2010;**105**:1788–94. http://dx.doi.org/10.1038/ajg.2010.156

166. Nurmi E, Haapamaki J, Paavilainen E, Rantanen A, Hillila M, Arkkila P. The burden of inflammatory bowel disease on health care utilization and quality of life. *Scand J Gastroenterol* 2013;**48**:51–7. http://dx.doi.org/10.3109/00365521.2012.685750

167. Jelsness-Jorgensen LP, Bernklev T, Moum B. Calprotectin is a useful tool in distinguishing coexisting irritable bowel-like symptoms from that of occult inflammation among inflammatory bowel disease patients in remission. *Gastroenterol Res Pract* 2013;620707. http://dx.doi.org/10.1155/2013/620707

168. Ramadas AV, Gunesh S, Thomas GA, Williams GT, Hawthorne AB. Natural history of Crohn's disease in a population-based cohort from Cardiff (1986–2003): a study of changes in medical treatment and surgical resection rates. *Gut* 2010;**59**:1200–6. http://dx.doi.org/10.1136/gut.2009.202101

169. Mao R, Xiao YL, Gao X, Chen BL, He Y, Yang L, *et al.* Fecal calprotectin in predicting relapse of inflammatory bowel diseases: A meta-analysis of prospective studies. *Inflammatory Bowel Dis* 2012;**18**:1894–9. http://dx.doi.org/10.1002/ibd.22861

170. Canani RB, Terrin G, Rapacciuolo L, Miele E, Siani MC, Puzone C, *et al.* Faecal calprotectin as reliable non-invasive marker to assess the severity of mucosal inflammation in children with inflammatory bowel disease. *Digest Liver Dis* 2008;**40**:547–53.

171. Fagerberg UL, Loof L, Lindholm J, Hansson LO, Finkel Y. Fecal calprotectin: a quantitative marker of colonic inflammation in children with inflammatory bowel disease. *JPGN* 2007;**45**:414–20. http://dx.doi.org/10.1097/MPG.0b013e31810e75a9

172. Van Rheenen PF. Role of fecal calprotectin testing to predict relapse in teenagers with inflammatory bowel disease who report full disease control. *Inflamm Bowel Dis* 2012:**18**;18–25. http://dx.doi.org/10.1002/ibd.22896

173. Panaccione R, Hibi T, Peyrin-Biroulet L, Schreiber S. Implementing changes in clinical practice to improve the management of Crohn's disease. *J Crohns Colitis* 2012;**6**(Suppl. 2):235–42. http://dx.doi.org/10.1016/S1873-9946(12)60503-0

174. D'Haens G, Ferrante M, Vermeire S, Baert F, Noman M, Moortgat L, *et al.* Fecal calprotectin is a surrogate marker for endoscopic lesions in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:2218–24. http://dx.doi.org/10.1002/ibd.22917

175. Lees C, Clark A, Walkden A, Satsangi J, Arnott I. Clinical utility of faecal calprotectin in diagnosing IBD during first presentation to the gastroenterology clinic: A novel investigative algorithm. *J Crohns Colitis* 2010;**4**:S32. http://dx.doi.org/10.1016/S1873-9954(10)70062-8

176. National Institute for Health and Care Excellence (NICE). SeHCAT (Tauroselcholic [75Selenium] acid) for the investigation of bile acid malabsorption (BAM) and measurement of bile acid pool loss. *Diagnostics guidance 7*. 2012. URL: http://guidance.nice.org.uk/DG7 (accessed 19 July 2013).

177. Poullis A, Foster R, Shetty A, Fagerhol MK, Mendall MA. Bowel inflammation as measured by fecal calprotectin: a link between lifestyle factors and colorectal cancer risk. *Cancer Epidem Biomar* 2004;**13**:279–84. http://dx.doi.org/10.1158/1055-9965.EPI-03-0160

178. Hoff G, Grotmol T, Thiis-Evensen E, Bretthauer M, Gondal G, Vatn MH. Testing for faecal calprotectin (PhiCal) in the Norwegian Colorectal Cancer Prevention trial on flexible sigmoidoscopy screening: comparison with an immunochemical test for occult blood (FlexSure OBT). *Gut* 2004;**53**:1329–33. http://dx.doi.org/10.1136/gut.2004.039032

179. Alrubaiy L, Malik A, Rees I, Bowen D. Usefulness of fecal calprotectin in clinical practice in a district general hospital. *Inflamm Bowel Dis* 2012;**18**:S53–4. http://dx.doi.org/10.1097/00054725-201212001-00130

180. Taylor N, Hills E, Sheen C, Al-Bahrani A, Grellier L. An audit of faecal calprotectin testing in suspected inflammatory bowel disease in the under 45's. *J Crohns Colitis* 2013;**7**:S131. http://dx.doi.org/10.1016/S1873-9946(13)60326-8

181. Pavlidis P, Chedgy FJQ, Tibble JA. Diagnostic accuracy and clinical application of faecal calprotectin in adult patients presenting with gastrointestinal symptoms in primary care. *Scand J Gastroenterol* 2013:**48**:1103–4. http://dx.doi.org/10.3109/00365521.2013.816771

182. Foell D, Wittkowski H, Vogl T, Roth J. S100 proteins expressed in phagocytes: a novel group of damage-associated molecular pattern molecules. *J Leukoc Biol* 2007;**81**:28–37. http://dx.doi.org/10.1189/jlb.0306170

183. Kaiser T, Langhorst J, Wittkowski H, Becker K, Friedrich AW, Rueffer A, *et al.* Faecal S100A12 as a non-invasive marker distinguishing inflammatory bowel disease from irritable bowel syndrome. *Gut* 2007;**56**:1706–13. http://dx.doi.org/10.1136/gut.2006.113431

184. Turner D, Leach ST, Mack D, Uusoue K, McLernon R, Hyams J, *et al.* Faecal calprotectin, lactoferrin, M2-pyruvate kinase and S100A12 in severe ulcerative colitis: a prospective multicentre comparison of predicting outcomes and monitoring response. *Gut* 2010;**59**:1207–12. http://dx.doi.org/10.1136/gut.2010.211755

185. Roseth AG, Kristinsson J, Fagerhol MK, Schjonsby H, Aadland E, Nygaard K, *et al.* Faecal calprotectin: a novel test for the diagnosis of colorectal cancer? *Scand J Gastroenterol* 1993;**28**:1073–6. http://dx.doi.org/10.3109/00365529309098312

186. Tibble JA, Sigthorsson G, Foster R, Scott D, Fagerhol MK, Roseth A, *et al.* High prevalence of NSAID enteropathy as shown by a simple faecal test. *Gut* 1999;**45**:362–6. http://dx.doi.org/10.1136/gut.45.3.362

187. Mooiweer E, Fidder HH, Van Erpecum KJ, Siersema PD, Laheij RJ, Oldenburg B. Determination of faecal haemoglobin is equally effective as faecal calprotectin in identifying inflammatory bowel disease patients with active endoscopic inflammation. *J Crohns Colitis* 2013;**7**:S114. http://dx.doi.org/10.1016/S1873-9946(13)60283-4

188. Sherwood RA. Faecal markers of gastrointestinal inflammation. *J Clin Pathol* 2012;**65**:981–5. http://dx.doi.org/10.1136/jclinpath-2012-200901

189. Sutherland AD, Gearry RB, Frizelle FA. Review of fecal biomarkers in inflammatory bowel disease. *Dis Colon Rectum* 2008;**51**:1283–91. http://dx.doi.org/10.1007/s10350-008-9310-8

190. D'Inca R, Dal PE, Di LV, Ferronato A, Fries W, Vettorato MG, *et al.* Calprotectin and lactoferrin in the assessment of intestinal inflammation and organic disease. *Int J Colorectal Dis* 2007;**22**:429–37. http://dx.doi.org/10.1007/s00384-006-0159-9

191. Joishy M, Davies I, Ahmed M, Wassel J, Davies K, Sayers A, *et al.* Fecal calprotectin and lactoferrin as noninvasive markers of pediatric inflammatory bowel disease. *JPGN* 2009;**48**:48–54. http://dx.doi.org/10.1097/MPG.0b013e31816533d3

192. Nancey S, Boschetti G, Moussata D, Cotte E, Peyras J, Cuerq C, *et al.* Neopterin is a novel reliable fecal marker as accurate as calprotectin for predicting endoscopic disease activity in patients with

inflammatory bowel diseases. *Inflamm Bowel Dis* 2013;**19**:1043–52. http://dx.doi.org/10.1097/MIB.0b013e3182807577

193. Arasaradnam RP, Ouaret N, Thomas MG, Quraishi N, Heatherington E, Nwokolo CU, *et al.* A novel tool for noninvasive diagnosis and tracking of patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2013;**19**:999–1003. http://dx.doi.org/10.1097/MIB.0b013e3182802b26

194. University of British Columbia. *FOCUS: The Future of Fecal Calprotectin Utility Study for the Diagnosis and Management of Inflammatory Bowel Disease (IBD)*. 2013. URL: http://clinicaltrials.gov/ct2/results?term=NCT01676324&Search=Search (accessed 19 July 2013).

195. Colledge NR, Walker BR and Ralston SH. *Davidson's Principles and Practice of Medicine*. 21st edn. Edinburgh: Churchill Livingstone/Elsevier;2010.

196. Bruzzese E, Raia V, Gaudiello G, Polito G, Buccigrossi V, Formicola V, *et al.* Intestinal inflammation is a frequent feature of cystic fibrosis and is reduced by probiotic administration. *Aliment Pharmacol Therapeut* 2004;**20**:813–19. http://dx.doi.org/10.1111/j.1365-2036.2004.02174.x

197. Bunn SK, Bisset WM, Main MJ, Golden BE. Fecal calprotectin as a measure of disease activity in childhood inflammatory bowel disease. *JPGN* 2001;**32**:171–7. http://dx.doi.org/10.1097/00005176-200102000-00015

198. Canani RB, Passaro M, Puzone C. Diagnostic value of faecal calprotectin in paediatric age. *Med Bambino* 2009;**28**:239–42.

199. Costa F, Mumolo MG, Bellini M, Romano MR, Ceccarelli L, Arpe P, *et al.* Role of faecal calprotectin as non-invasive marker of intestinal inflammation. *Digest Liver Dis* 2003;**35**:642–7. http://dx.doi.org/10.1016/S1590-8658(03)00381-5

200. Eder P, Stawczyk-Eder K, Krela-Kazmierczak I, Linke K. Clinical utility of the assessment of fecal calprotectin in Lesniowski-Crohn's disease. *Pol Arch Med Wewn* 2008;**118**:622–6.

201. El Nuamani N, El-Gendi H, Zaki S, Elamin H, Tawfik M, Al-Sheshtawy F. Evaluation of the Diagnostic Yield of Faecal Calprotectin in Egyptian Children with Chronic Diarrhoea. *Arab J Gastroenterol* 2007;**8**:132–5.

202. Elkjaer M, Burisch J, Voxen HV, Deibjerg KB, Slott Jensen JK, Munkholm P. A new rapid home test for faecal calprotectin in ulcerative colitis. *Aliment Pharmacol Therapeut* 2010;**31**:323–30. http://dx.doi.org/10.1111/j.1365-2036.2009.04164.x

203. Erbayrak M, Turkay C, Eraslan E, Cetinkaya H, Kasapoglu B, Bektas M. The role of fecal calprotectin in investigating inflammatory bowel diseases. *Clinics* 2009;**64**:421–5. http://dx.doi.org/10.1590/S1807-59322009000500009

204. Flagstad G, Helgeland H, Markestad T. Faecal calprotectin concentrations in children with functional gastrointestinal disorders diagnosed according to the Pediatric Rome III criteria. *Acta Paediatr* 2010;**99**:734–7. http://dx.doi.org/10.1111/j.1651-2227.2010.01698.x

205. Grogan JL, Casson DH, Terry A, Burdge GC, El-Matary W, Dalzell AM. Enteral feeding therapy for newly diagnosed pediatric Crohn's disease: a double-blind randomized controlled trial with two years follow-up. *Inflamm Bowel Dis* 2012;**18**:246–53.

206. Guo H-B, Cao J-B, Wang Z-H, Li H-R. Fecal calprotectin in differential diagnosis between irritable bowel syndrome and inflammatory bowel disease. *World Chin J Dig* 2009;**17**:1152–5.

207. Jensen MD, Kjeldsen J, Nathan T. Fecal calprotectin is equally sensitive in Crohn's disease affecting the small bowel and colon. *Scand J Gastroenterol* 2011;**46**:694–700. http://dx.doi.org/10.3109/00365521.2011.560680

208. Kobelska-Dubiel N, Igny-acs I, Krauss H, Cichy W, Kobelski M. [Fecal calprotectin as an inflammatory marker in inflammatory bowel diseases in children.] *Pediatr Wspolczesna* 2007;**9**:172–5.

209. Langhorst J, Elsenbruch S, Koelzer J, Rueffer A, Michalsen A, Dobos GJ. Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel diseases: performance of fecal lactoferrin, calprotectin, and PMN-elastase, CRP, and clinical indices. *Am J Gastroenterol* 2008;**103**:162–9. http://dx.doi.org/10.1111/j.1572-0241.2007.01556.x

210. Meucci G, D'Inca R, Maieron R, Orzes N, Vecchi M, Visentini D, *et al.* Diagnostic value of faecal calprotectin in unselected outpatients referred for colonoscopy: A multicenter prospective study. *Digest Liver Dis* 2010;**42**:191–5. http://dx.doi.org/10.1016/j.dld.2009.07.002

211. Olafsdottir E, Aksnes L, Fluge G, Berstad A. Faecal calprotectin levels in infants with infantile colic, healthy infants, children with inflammatory bowel disease, children with recurrent abdominal pain and healthy children. *Acta Paediatr* 2002;**91**:45–50. http://dx.doi.org/10.1111/j.1651-2227.2002.tb01638.x

212. Perminow G, Beisner J, Koslowski M, Lyckander LG, Stange E, Vatn MH, *et al.* Defective paneth cell-mediated host defense in pediatric ileal Crohn's disease. *Am J Gastroenterol* 2010;**105**:452–9. http://dx.doi.org/10.1038/ajg.2009.643

213. Quail MA, Russell RK, Van Limbergen JE, Rogers P, Drummond HE, Wilson DC, *et al.* Fecal calprotectin complements routine laboratory investigations in diagnosing childhood inflammatory bowel disease. *Inflamm Bowel Dis* 2009;**15**:756–9. http://dx.doi.org/10.1002/ibd.20820

214. Ricanek P, Brackmann S, Perminow G, Lyckander LG, Sponheim J, Holme O, *et al.* Evaluation of disease activity in IBD at the time of diagnosis by the use of clinical, biochemical, and fecal markers. *Scand J Gastroenterol* 2011;**46**:1081–91. http://dx.doi.org/10.3109/00365521.2011.584897

215. Roseth AG, Aadland E, Jahnsen J, Raknerud N. Assessment of disease activity in ulcerative colitis by faecal calprotectin, a novel granulocyte marker protein. *Digestion* 1997;**58**:176–80.

216. Roseth AG, Schmidt PN, Fagerhol MK. Correlation between faecal excretion of indium-111-labelled granulocytes and calprotectin, a granulocyte marker protein, in patients with inflammatory bowel disease. *Scand J Gastroenterol* 1999;**34**:50–4.

217. Schoepfer AM, Beglinger C, Straumann A, Trummler M, Renzulli P, Seibold F. Ulcerative colitis: correlation of the Rachmilewitz endoscopic activity index with fecal calprotectin, clinical activity, C-reactive protein, and blood leukocytes. *Inflamm Bowel Dis* 2009;**15**:1851–8. http://dx.doi.org/10.1002/ibd.20986

218. Schoepfer AM, Beglinger C, Straumann A, Trummler M, Vavricka SR, Bruegger LE, *et al.* Fecal calprotectin correlates more closely with the Simple Endoscopic Score for Crohn's disease (SES-CD) than CRP, blood leukocytes, and the CDAI. *Am J Gastroenterol* 2010;**105**:162–9. http://dx.doi.org/10.1038/ajg.2009.545

219. Shastri YM, Povse N, Stein J. A prospective comparative study for new rapid bedside fecal calprotectin test with an established ELISA to assess intestinal inflammation. *Clin Lab* 2009;**55**:53–5.

220. Shulman RJ, Eakin MN, Czyzewski DI, Jarrett M, Ou CN. Increased gastrointestinal permeability and gut inflammation in children with functional abdominal pain and irritable bowel syndrome. *J Pediatr* 2008;**153**:646–50. http://dx.doi.org/10.1016/j.jpeds.2008.04.062

221. Sidhu R, Wilson P, Wright A, Yau CWH, D'Cruz FA, Foye L, *et al.* Faecal lactoferrin – a novel test to differentiate between the irritable and inflamed bowel? *Aliment Pharmacol Therapeut* 2010;**31**:1365–70. http://dx.doi.org/10.1111/j.1365-2036.2010.04306.x

222. Sipponen T, Karkkainen P, Savilahti E, Kolho KL, Nuutinen H, Turunen U, *et al.* Correlation of faecal calprotectin and lactoferrin with an endoscopic score for Crohn's disease and histological findings. *Aliment Pharmacol Therapeut* 2008;**28**:1221–9. http://dx.doi.org/10.1111/j.1365-2036.2008.03835.x

223. Sipponen T, Savilahti E, Kolho KL, Nuutinen H, Turunen U, Farkkila M. Crohn's disease activity assessed by fecal calprotectin and lactoferrin: correlation with Crohn's disease activity index and endoscopic findings. *Inflamm Bowel Dis* 2008;**14**:40–6.

224. Sipponen T, Haapamaki J, Savilahti E, Alfthan H, Hamalainen E, Rautiainen H, *et al.* Fecal calprotectin and S100A12 have low utility in prediction of small bowel Crohn's disease detected by wireless capsule endoscopy. *Scand J Gastroenterol* 2012;**47**:778–84. http://dx.doi.org/10.3109/00365521.2012.677953

225. Summerton CB, Longlands MG, Wiener K, Shreeve DR. Faecal calprotectin: a marker of inflammation throughout the intestinal tract. *Eur J Gastroenterol Hepatol* 2002;**14**:841–5. http://dx.doi.org/10.1097/00042737-200208000-00005

226. Tibble J, Teahon K, Thjodleifsson B, Roseth A, Sigthorsson G, Bridger S, *et al.* A simple method for assessing intestinal inflammation in Crohn's disease. *Gut* 2000;**47**:506–13. http://dx.doi.org/10.1136/gut.47.4.506

227. Uslu N, Baysoy G, Balamtekin N, Hizal G, Demir H, Saltik-Temizel IN, *et al.* [A noninvasive marker for pediatric inflammatory bowel disease: fecal calprotectin.] *Cocuk Sagligi Hast Derg* 2011;**54**:22–7.

228. Wassell J, Dolwani S, Metzner M, Losty H, Hawthorne A. Faecal calprotectin: a new marker for Crohn's disease? *Ann Clin Biochem* 2004;**41**:3–2. http://dx.doi.org/10.1258/000456304323019613

229. Xiang JY, Ouyang Q, Li GD, Xiao NP. Clinical value of fecal calprotectin in determining disease activity of ulcerative colitis. *World J Gastroenterol* 2008;**14**:53–7. http://dx.doi.org/10.3748/wjg.14.53

230. Johne B, Kronborg O, Ton HI, Kristinsson J, Fuglerud P. A new fecal calprotectin test for colorectal neoplasia. Clinical results and comparison with previous method. *Scand J Gastroenterol* 2001;**36**:291–6. http://dx.doi.org/10.1080/003655201750074618

231. Roseth AG, Fagerhol MK, Aadland E, Schjonsby H. Assessment of the neutrophil dominating calprotectin in feces. A methodologic study. *Scand J Gastroenterol* 1992;**27**:793–8. http://dx.doi.org/10.3109/00365529209011186

232. Bewtra M, Kaiser LM, TenHave T, Lewis JD. Crohn's disease and ulcerative colitis are associated with elevated standardized mortality ratios: a meta-analysis. *Inflamm Bowel Dis* 2013;**19**:599–613. http://dx.doi.org/10.1097/MIB.0b013e31827f27ae

233. British Medical Association and Royal Pharmaceutical Society of Great Britain. *British National Formulary*. No. 65, March 2013. London: BMA and RPS; 2013. www.bnf.org/bnf/index.htm

Appendix 1 Comparison of ulcerative colitis, Crohn's disease, irritable bowel syndrome and coeliac disease

TABLE 74 Comparison between UC, CD, IBS and coeliac disease

	UC	CD	IBS	Coeliac disease
Age	Any	Any	Any, but more in young	Any age Adults: peak onset in third or fourth decade
Gender	M = F	M = F	Female preponderance	In adults, females are affected twice as often as males
Population distribution	Incidence in Western countries: 10–20 per 100,000 Prevalence in Western countries: 100–200 per 100,000	Incidence in Western countries: 5–10 per 100,000 Prevalence in Western countries: 50–100 per 100,000	About 20% but only 10% consult GPs	Worldwide but more common in northern Europe Prevalence in the UK approximately 1% but 50% of them are asymptomatic (include both undiagnosed 'silent' cases and 'latent' cases, who later develop coeliac disease)
Ethnic group	Any	Any; more common in Ashkenazi Jews	Not reported	Not mentioned
Genetic factors	*HLA-DR103* associated with severe disease	*CARD 15/NOD-2* mutations predispose	Not reported	Genetically susceptible individuals intolerant to wheat gluten and similar proteins found in rye, barley and, to oats: Associated with other HLA-linked autoimmune disorders and with certain other diseases
Risk factors	More common in non-/ex-smokers; appendectomy protects	More common in smokers (RR = 3)	History of psychological stress	Not mentioned as risk factors but coeliac disease associated with other HLA-linked autoimmune disorders and with certain other diseases, such as IDDM, thyroid disease, primary biliary cirrhosis, IBD, Sjögren's syndrome, IgA deficiency, pernicious anaemia, sarcoidosis, myasthenia gravis, Down's syndrome, etc.
Diagnosis	Clinical confirmed by biopsy	Biopsy	Clinical. Diagnosis supported by symptoms for more than 6 months; worsened by stress; FBC and ESR normal	Endoscopic small bowel biopsy is the gold standard (villous atrophy): IgA antiendomysial antibodies by immunofluorescence: haematology (micro- and macrocytic anaemia) and biochemistry (low Ca, Mg, total protein, albumin and vitamin D)

	UC	CD	IBS	Coeliac disease
Anatomical distribution	Colon only; begins at anorectal margin with variable proximal extension Proctitis (rectum); proctosigmoiditis (rectum and sigmoid colon); pancolitis (whole colon)	Any part of GI tract; perianal disease common; patchy distribution – 'skip lesions' Sites involved (in order of frequency): terminal ileum and right side of colon, colon alone, terminal ileum alone, ileum and jejunum	Colonoscopy	Small bowel
Extraintestinal manifestations	Common	Common	Associated with other conditions such as dysmenorrhoea, non-ulcer dyspepsia, 'fibromyalgia'	Common
Presentation	Bloody diarrhoea *Proctitis* – rectal bleeding, mucus discharge, tenesmus *Proctosigmoiditis* – bloody diarrhoea with mucus; some develop fever, lethargy and abdominal discomfort *Extensive pancolitis* – bloody diarrhoea with passage of mucus. Severe case – anorexia, malaise, weight loss and abdominal pain, patient is toxic with fever, tachycardia and signs of peritoneal inflammation	Variable; pain, diarrhoea, weight loss all common *Ileal CD*: there may be subacute or even acute intestinal obstruction. Diarrhoea – watery but no blood or mucus *Crohn's colitis*: identical to UC but rectum spared and presence of perianal disease. Many presents with symptoms of both small bowel and colonic disease. In few, isolated perianal disease, vomiting from jejunal strictures and severe oral ulceration	Recurrent colicky abdominal pain or cramping, relieved by defecation Abdominal distension Episodes of diarrhoea but can have more of a constipation pattern Patients well, no weight loss	Depends on age of onset Infancy: diarrhoea, malabsorption and failure to thrive – symptoms starts after weaning on to cereals Older children: non-specific features like delayed growth; malnutrition, mild abdominal distension; growth and pubertal delay Adults: Highly variable, depending on the severity and extent of small bowel involvement. Some florid malabsorption, others non-specific symptoms, such as tiredness, weight loss, folate or iron-deficiency anaemia. Oral ulceration, dyspepsia and bloating; mild undernutrition and increased risk of osteoporosis
Histology	Inflammation limited to mucosa; crypt distortion, cryptitis, crypt abscesses, loss of goblet cells	Submucosal or transmural inflammation common; deep fissuring ulcers, fistulas; patchy changes; granulomas	Normal	Subtotal villous atrophy Sometime villous appears normal but there may be excess numbers of intraepithelial lymphocytes

Appendix 2 Search strategy

Calprotectin: diagnostic studies and economics

MEDLINE (Ovid)
Searched: 1946 to September 2012.

1. exp Inflammatory Bowel Diseases/di [Diagnosis]
2. exp Irritable Bowel Syndrome/di [Diagnosis]
3. crohn's disease.tw.
4. ulcerative colitis.tw.
5. inflammatory bowel disease*.tw.
6. irritable bowel syndrome*.tw.
7. (IBS or IBD).tw.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. calprotectin.tw.
10. 8 and 9

EMBASE (Ovid)
Searched: 1980 to September 2012.

1. crohn's disease.tw.
2. ulcerative colitis.tw.
3. inflammatory bowel disease*.tw.
4. irritable bowel syndrome*.tw.
5. (IBS or IBD).tw.
6. calprotectin.tw.
7. 1 or 2 or 3 or 4 or 5
8. 6 and 7
9. exp *Crohn disease/di [Diagnosis]
10. exp *ulcerative colitis/di [Diagnosis]
11. 9 or 10
12. 6 and 11
13. 8 or 12
14. (fecal or faecal).tw.
15. 13 and 14

Auto-alerts
Ran auto-alerts of the above searches in MEDLINE and EMBASE from September 2012 to March 2013 for studies added subsequent to the initial searches.

The Cochrane Library: all sections
Searched: September 2012.

Search terms: calprotectin and (inflammatory bowel disease* or irritable bowel syndrome or crohn's disease or ulcerative colitis)

Web of Science: Science Citation Index, Conference Proceedings Citation Index

Searched: 1980 to September 2012.

Search terms: calprotectin and (inflammatory bowel disease* or irritable bowel syndrome or crohn's disease or ulcerative colitis)

Cost-effectiveness searches

MEDLINE (Ovid) and EMBASE

MEDLINE (Ovid) searched:1996 to October 2012.

EMBASE searched: 1996 to October 2012.

1. exp Economics/
2. Health Status/
3. exp "Quality of Life"/
4. exp Quality-Adjusted Life Years/
5. exp Patient Satisfaction/
6. (pharmacoeconomic* or pharmaco-economic* or economic* or cost-effective* or cost-benefit*).tw.
7. (health state* or health status).tw.
8. (qaly* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or SF-36 or SF36 or SF-6D or HUI).tw.
9. (markov or time trade off or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.
10. (quality adj2 life).tw.
11. (decision adj2 model).tw.
12. (visual analog* scale* or discrete choice experiment* or health* year* equivalen*).tw.
13. ("resource use" or resource utili?ation).tw.
14. (well-being or wellbeing or satisfaction).ti.
15. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
16. exp Inflammatory Bowel Diseases/ec, px [Economics, Psychology]
17. exp Irritable Bowel Syndrome/ec, px [Economics, Psychology]
18. (inflammatory bowel disease* or crohn* disease or ulcerative colitis).ti.
19. irritable bowel.m_titl.
20. 17 or 18 or 19
21. 15 and 20
22. limit 21 to english language
23. limit 22 to yr = "1996 -Current"

Cochrane Library, Economic Evaluations Database, Issue 4 of 4, October 2012

Search terms: (inflammatory bowel or irritable bowel or crohn* or ulcerative colitis)

Other searches for calprotectin

- Searched the website of the journal *Gut*.
- Searched ECCO (European Crohn's and Colitis Organisation) 2012 and 2013 Congress abstracts.
- Checked reference lists of previous systematic reviews.
- Personal communication with experts for unpublished data.

Searches for adverse effects of colonoscopy

MEDLINE (Ovid)
Searched: 1946 to February 2013.

1. exp *Colonoscopy/ae [Adverse Effects]
2. (Colonoscopy or Sigmoidoscopy).m_titl.
3. (perforation* or perforated or complication*).tw.
4. 2 and 3
5. 1 or 4
6. limit 5 to english language
7. (case reports or comment or letter).pt.
8. 6 not 7
9. colonoscopy.m_titl.
10. 8 and 9

PubMed
Searched: All database up to February 2013.

Search terms: (colonoscopy and (perforat* or adverse or complication* or risk)) in title field

Natural history/progression of inflammatory bowel disease

1. exp Inflammatory Bowel Diseases/
2. (inflammatory bowel disease* or crohn* disease or ulcerative colitis).ti.
3. 1 or 2
4. (natural history or (disease adj course) or (clinical course) or progression or (disease adj2 progress*)).tw.
5. exp Disease Progression/
6. 4 or 5
7. 3 and 6
8. limit 7 to english language

MEDLINE (Ovid)
Searched: 1946 to October 2012.

1. (inflammatory bowel disease* or crohn* disease or ulcerative colitis).ti.
2. (natural history or (disease adj course) or (clinical course) or progression or (disease adj2 progress*)).ti.
3. 1 and 2
4. limit 3 to english language

Research in progress

Included only open studies and excluded studies with unknown status.

1. ClinicalTrials.gov
2. Current Controlled Trials
3. UK Clinical Trials Gateway
4. UK Clinical Research Network Study Portfolio
5. EU Clinical Trials Register website
6. EUDRACT European Clinical Trials Database
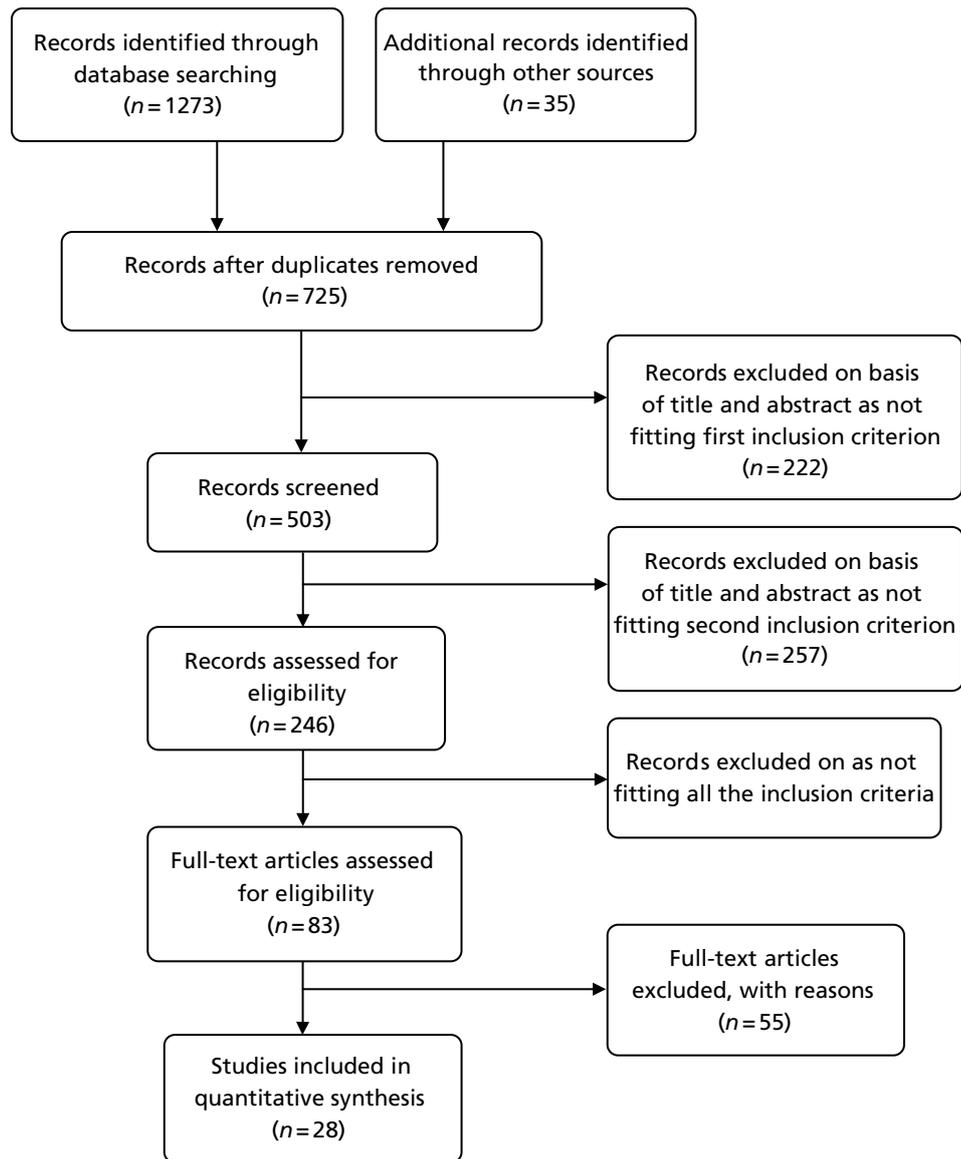7. WHO (World Health Organization) Clinical Trials Search Portal

FIGURE 27 Flow diagram.

TABLE 75 Reasons for exclusion

First author and year	Reason for exclusion
Aomatsu 2011[64]	Previously diagnosed – and does not say how long since diagnosis
Bremner 2005[93]	Only 12/43 (28%) were newly diagnosed
Bruzzese 2004[196]	Not enough detail about the 15 patients with active IBD
Bunn 2001[71]	Only 9/22 (41%) are newly diagnosed – remainder under review on treatment (from data in table 1)
Bunn 2001[197]	Not newly diagnosed, 13/22 (59%) under review on treatment
Canani 2004[198]	Could not derive a 2 × 2 table; includes patients in remission (proportion not reported)
Canani 2008[170]	States that patients had previously established diagnosis of IBD
Costa 2003[199]	Does not state what percentage of patients not previously diagnosed – first paragraph of p. 645 describes patients with clinically active disease and those with quiescent disease, i.e. mixture of patients
D'Inca 2007[190]	Patients with IBD had known diagnosis – referred for active disease or surveillance
Dolci 2012[96]	Letter – no information on patients – could not get 2 × 2 data – laboratory based setting
Eder 2008[200]	Confirmed diagnosis of CD – mean duration 5 years
El Nuamani 2007[201]	Requested full text but library unable to locate it
Elkjaer 2010[202]	Patients with known CD – but maybe useful to compare POCT vs. ELISA?
Erbayrak 2009[203]	Previously diagnosed – mean disease period 5 years
Fagerberg 2007[171]	Comprises a mixture of children with suspected or previously confirmed IBD – but does not give us the numbers of each
Flagstad 2010[204]	Spectrum bias as only looked at children with functional GI disorders – organic disease was excluded
Grogan 2012[205]	Spectrum bias – excluded children with only large bowel disease
Guo 2009[206]	Chinese language. Appeared from abstract that patients had a known diagnosis
Hessels 2012[98]	Comprised a mixture of patients with suspected (40/85) or relapse of IBD
Hornung 2011[110]	Health economics study, no 2 × 2 data available
Jensen 2011[207]	Possible spectrum bias; also, does not give the diagnosis of 43 of the 83 patients who were not CD, i.e. we do not know how many of them had inflammatory vs. non-inflammatory
Joishy 2009[191]	Includes new cases and those in relapse but does not say how many of each
Keohane 2010[165]	Patients included only if in clinical remission
Kobelska-Dubiel 2007[208]	Polish language – not able to get translation
Komraus 2012[94]	Patients had a known diagnosis of various types of IBD; could not get data for a 2 × 2 table
Koulaouzidis 2011[161]	Spectrum bias – all patients had previously had negative bidirectional endoscopies but continuing suspicion of CD and referred for small bowel capsule endoscopy
Langhorts 2008[209]	Only patients with previously diagnosed IBD were included
Meucci 2010[210]	Not all new onset. High proportion of IBD patients had clinically quiescent IBD – did not report FC data for active IBD separately – so could not derive 2 × 2 table
Mindemark 2012[111]	Health economics paper – no 2 × 2 data available
Moum 2010[43]	Article looked at reproducibility of tests – no 2 × 2 data available
Olafsdottir 2002[211]	Majority of patients not have the reference standard of colonoscopy
Perminow 2010[212]	Not enough detail to extract sensitivity and specificity data

TABLE 75 Reasons for exclusion (*continued*)

First author and year	Reason for exclusion
Quail 2009[213]	Measured at diagnosis but not enough detail reported to derive data for a 2 × 2 table
Ricanek 2011[214]	Does not give the manufacturer of the test and does not give a cut-off value for the test – does not report enough data to construct a 2 × 2 table
Roseth 1997[215]	Most patients were previously diagnosed – only 9/62 = 15% underwent first diagnostic examination – median disease duration is 5 years
Roseth 1999[216]	Patients appeared to be in remission and were receiving maintenance treatment only
Schoepfer 2007[74]	Not newly presenting patients in a differential diagnosis situation – 19 of 24 patients with CD in remission – so had no symptoms
Schoepfer 2009[217]	Previously confirmed diagnosis of UC referred for colonoscopy'
Schoepfer 2010[218]	Previously confirmed diagnosis of CD referred for colonoscopy'
Shastri 2009[219]	Poster – not enough details about patients or data for 2 × 2 tables
Shaoul 2012[92]	Not enough data for a 2 × 2 table
Shulman 2008[220]	Only includes children with functional abdominal pain (IBS)
Sidhu 2010[221]	Includes patients with known diagnosis – about 44% active disease and 66% inactive disease
Sipponen 2008[222]	Includes patients with an established CD diagnosis
Sipponen 2008[223]	Not newly diagnosed – patients referred for ileocolonoscopy had a disease duration of mean 9.2 years
Sipponen 2012[224]	Time between index test and reference test greater than 3 months. Spectrum bias – other tests had already been done. Small bowel only
Summerton 2002[225]	Most had a known diagnosis, so not new. No data on IBS or IBD in relation to FC cut-off levels
Sydora 2012[99]	Patients have had a known diagnosis prior to the study
Tibble 2000[226]	Patients appeared to be a subgroup of those included in Tibble *et al.*[49]
Tomkins 2012[105]	A mixture of previously diagnosed and newly presenting patients but numbers of each not known – personal communication
Usl 2011[227]	Turkish language – not able to get translation
Vestergaard 2008[100]	Patients not newly diagnosed – aim of the study was to compare Rapid and ELISA tests, patients had known diagnosis
Wassell 2004[228]	Do not know duration of disease in CDs. The IBS had been diagnosed in previous year – but 1 year post diagnosis in IBS group could mean that they had been treated so FC might have been higher at onset
Wassell 2012[95]	No details of patients providing samples for FC testing
Xiang 2008[229]	Patients already had known diagnosis of CD

Appendix 3 Description of different tests

Enzyme-linked immunosorbent assay

PhiCal enzyme-linked immunosorbent assay test

This is a quantitative ELISA kit manufactured by Calpro (Oslo, Norway). The kit can be used to measure increased concentration of calprotectin in plasma, cerebrospinal fluid, synovial fluid, urine and stool. This test has Food and Drug Administration (FDA) approval and is marketed in the USA for determination of calprotectin level in stools. It has a conformité Européenne (European Conformity) (CE) mark.

A polyclonal rabbit antibody is used. The manufacturer states that the affinity of the antibody to six different epitopes of calprotectin makes this ELISA test more robust and less likely to give false results compared with the test using monoclonal antibody, which has affinity for single epitope.

The cut-off value is 50 mg/kg and a FC level of > 50 mg/kg is regarded as positive. Previously the cut-off level was 10 mg/l.

URL: www.phical.com/uploads/PhiCaltestperformance.pdf (last accessed 19 July 2013).

URL: www.phical.com/uploads/Instructions.pdf (last accessed 19 July 2013).

PhiCal calprotectin enzyme-linked immunosorbent assay kit K6927

This is a quantitative ELISA test manufactured by Immundiagnostik AG (Bensheim, Germany). The kit is supplied by Biohit in the UK. Indications for using this kit include: marker of acute inflammation; estimation of degree of GI inflammation for monitoring Morbus CD, colitis ulcerosa or the patient status after removal of polyps and discrimination between patients with IBD and IBS.

An older version is K6937. Both versions have CE marks.

The assay uses two monoclonal antibodies that bind to human calprotectin. The sample can be stored but the manufacturer advises against storing samples for > 48 hours at 2–8 °C. Stool samples can be stored for longer periods at –20 °C.

A limitation of the test, not relevant to this review, is that stool samples with calprotectin level higher than the upper standard value need to be diluted and re-assayed. A FC level of above 50 mg/kg is regarded as positive, for adults and children aged 4–17 years. The manufacturer however recommends the laboratory to establish their own normal range.

URL: www.immundiagnostik.com/fileadmin/pdf/PhiCalCalpro_Stuhl_1h_K6927.pdf (last accessed 19 July 2013).

EK-CAL

This ELISA kit is manufactured by Bühlmann Laboratories (Schönenbuch, Switzerland). It is used for extraction and quantitative determination of FC levels. It has a CE mark.

A monoclonal antibody is used.

The assay can be performed in two different ways, based on the expected FC levels. The low-range ELISA procedure can be used for FC levels up to 600 μg/g (range 10–600 μg/g) and the extended range ELISA procedure for FC levels of up to 1800 μg/g (range 30–1800 μg/g).

The cut-off level is 50 μg/g for both adults and children aged between 4 and 17 years.

URL: www.buhlmannlabs.ch/files/documents/core/Inflammation/ifu/ek-cal-ifu-ce-121120.pdf (last accessed 19 July 2013).

Calprest

This is an ELISA kit developed by Eurospital Spa (Trieste, Italy).

CE mark – not mentioned in NICE scoping documents.

The cut-off level is 50 mg/kg. The manufacturer suggests retesting after a short period of time in patients with FC levels of between 50 and 100 mg/kg. The FC level of above 50 mg/kg is considered as positive.

URL: www.calprotectintest.com/english/calprest.html (last accessed 19 July 2013).

Calpro Calprotectin ELISA test (ALP)

This quantitative method was developed by Calpro AS (Lysaker, Norway).

Based on two studies (Johne *et al.* 2001;[230] Roseth *et al.* 1992[231]), calprotectin values of < 50, > 50, 350 and 200–40,000 mg/kg represented normal value, positive value, median value in patients with symptomatic colorectal cancers and active, symptomatic IBD, respectively.

One limitation of the test is that repeated freeze–thaw cycles of the specimen may affect the accuracy of the test results. The manufacturer cautions against a diagnosis based on a single stool test.

URL: www.calpro.no/products/calprotectin-elisa-test-alp (last accessed 19 July 2013).

Rapid test

Quantum Blue

This is a rapid test manufactured by Bühlmann Laboratories. There are two types of Bühlmann rapid tests: (1) the lower range Quantum Blue LF CAL (30–300 μg/g) and (2) the high-range Quantum Blue LF-CHR (100–1800 μg/g). The LF-CAL is designed for distinguishing between OBD and non-organic, or to exclude IBD. The cut-off value of this test is 50 μg/g. The LF-CHR test is follow-up of IBD patients during their therapy.

There are two parts: the test cartridge (to load the stool sample) and the reader. The reader is used to read quantitative concentration of FC. The results are available within 12–15 minutes – calprotectin μg/g.

The manufacturer recommends re-testing samples if results are between 30 and 70 μg/g. This zone is regarded as 'grey zone' and corresponds to the 2.5th–97.5th percentile of imprecision around the cut-off of 50 μg/g.

URL: www.buhlmannlabs.ch/core/quantum-blue/ (last accessed 19 July 2013).

PreventID CalDetect

This is a semiquantitative immunochromatographic rapid test manufactured by Preventis, GmbH (Bensheim, Germany).

The result is interpreted in about 10 minutes.

If there is a solid red control (C) line then it indicates the test has run correctly. The next test bands (T1, T2, T3) will depend on the concentration of calprotectin. If C, T1 and T2 are visible then it indicates a calprotectin concentration of between 15 and 60 µg/g. If all of the test bands (C, T1, T2 and T3) are visible then it indicates a calprotectin concentration of > 60 µg/g.

If the control band (C) remains blue or only a test band (T) is visible then the test is invalid.

URL: www.preventis-online.de/fileadmin/pdf/checksEngl/CalDetect_engl.pdf (last accessed 19 July 2013).

Prevista

This is a chromatographic immunoassay manufactured by GmbH & Co KG (Munich, Germany). The test device has two lines: a test and a control. The test line contains anti-calprotectin antibodies, whereas the control line contains anti-immunoglobulin antibodies, both dried on the membrane.

The results are read within the next 5 minutes. (Details taken from Damms and Bischoff[58] – no webpage found.)

EliA platform (details based on correspondence with manufacturer)

EliA platform is a fully automated calprotectin stool test, manufactured by the Immunodiagnostics Division of Thermo Fisher (TF IDD) [previously manufactured by Phadia but the company was acquired by Thermo Fisher in 2011].

No details, such as CE mark, were available from the NICE scoping documents.

The test was formally launched in November 2012, and is being currently used across seven sites in the UK. The test is a fully quantitative test, which gives results in milligrams per kilogram. Four different types of instruments are available, namely Phadia 100, 250, 2500 and 5000. They all vary in size and capacity, and are designed to meet the requirement of different laboratories. The most commonly used platform in the UK are Phadia 250 and Phadia 100. The test is run as a single test and does not need to be repeated – an advantage over other ELISA tests. The platform is fully automated. The Phadia solution can be added to the existing Phadia systems without the need for further readers and plate washers. The fully automated system reduces laboratory technician workload, time and cost.

Appendix 4 Quality assessment tables

TABLE 76 Quality assessment of all the included studies

Study	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Whole sample verified using reference standard?	Same reference standard used?	Were the patients newly diagnosed?	Reference standard results blinded?	Index test results blinded?	Same clinical data as used in test results available in practice?	Uninterpretable/intermediate test results reported?	Withdrawals explained?
Ashorn 2009[89]	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes
Bharathi 2005[88]	Yes	Yes	Unclear	No	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes
Basumani 2012[83]	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes
Burri 2012[85]	Yes	Yes	Unclear	Unclear	No	Yes	Yes	Yes	Yes	Yes	Yes
Canani 2006[65]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Carroccio 2003[87]	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Damms 2008[58]	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes
Diamanti 2010[86]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dolwani 2004[48]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes
El-Badry 2010[81]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes
Fagerberg 2005[91]	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Garcia 2006[59]	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes	Unclear
Henderson 2012[30]	Yes	Yes	Unclear	No	Unclear	Yes	Unclear	Unclear	Yes	Yes	Yes
Kok 2012[39]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lee 2013[90]	Yes	Unclear	Yes	No	Unclear	Yes	No	Yes	Yes	Yes	Yes

Study	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Whole sample verified using reference standard?	Same reference standard used?	Were the patients newly diagnosed?	Reference standard results blinded?	Index test results blinded?	Same clinical data as used in test results available in practice?	Uninterpretable/ intermediate test results reported?	Withdrawals explained?
Li 2006[60]	Unclear	Yes	Yes	Yes	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Unclear
Licata 2012[61]	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Limburg 2000[78]	Unclear	Yes	Unclear	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Manz 2012[84]	Yes	Yes	Unclear	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Otten 2008[73]	Yes	Yes	Yes	No	Unclear	Yes	Yes	Yes	Yes	Yes	Yes
Schoepfer 2008[75]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Schroder 2007[77]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shitrit 2007[62]	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes
Sidler 2008[82]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Tibble 2002[49]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes
Tomas 2007[66]	Unclear	Yes	Unclear	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes
Turvill 2012[46]	Yes	Unclear	Unclear	No	Yes	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Van de Vijver 2012[80]	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

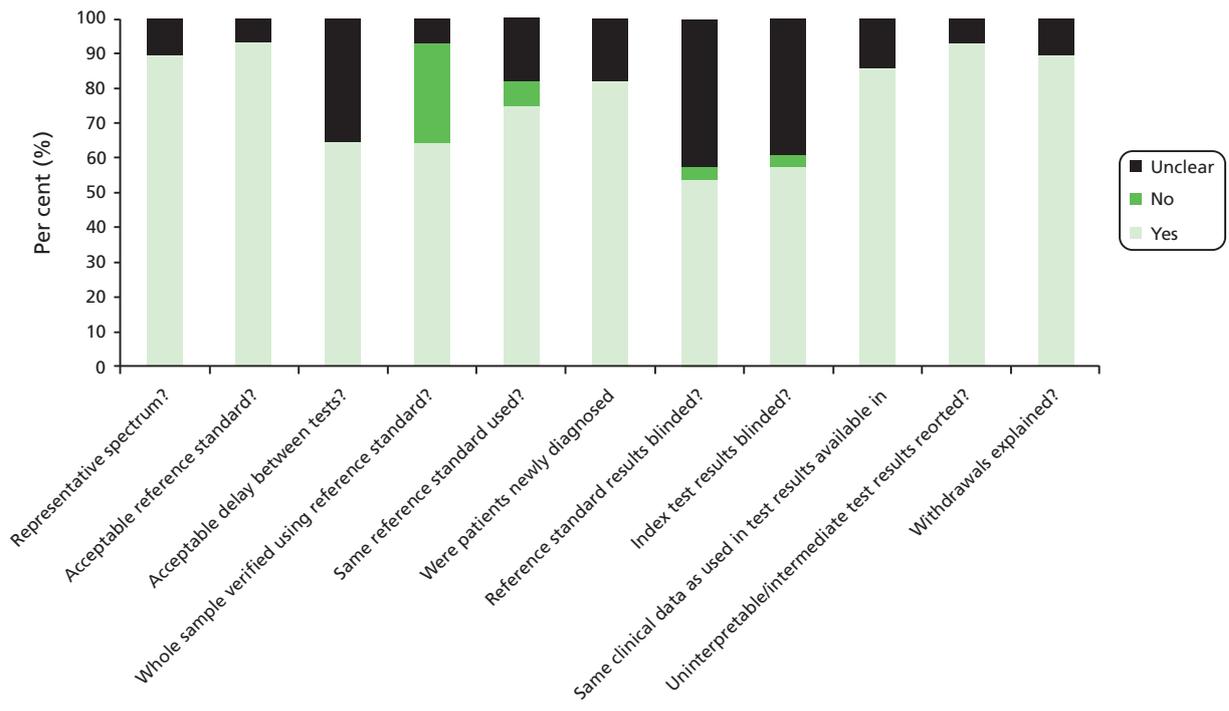**TABLE 76** Quality assessment of all the included studies (*continued*)

Study	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Whole sample verified using reference standard?	Same reference standard used?	Were the patients newly diagnosed?	Reference standard results blinded?	Index test results blinded?	Same clinical data as used in test results available in practice?	Uninterpretable/ intermediate test results reported?	Withdrawals explained?
Ashorn 2009[89]	✓	✓	?	✓	✓	✓	?	?	✓	✓	✓
Bharathi 2005[88]	✓	✓	?	x	✓	✓	?	?	?	✓	✓
Basumani 2012[83]	✓	✓	?	✓	✓	✓	?	?	?	✓	✓
Burri 2012[85]	✓	✓	?	?	x	✓	✓	✓	✓	✓	✓
Canani 2006[65]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Carroccio 2003[87]	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	✓
Damms 2008[58]	✓	✓	✓	✓	✓	?	?	✓	✓	✓	✓
Diamanti 2010[86]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dolwani 2004[48]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
El-Bady 2010[81]	✓	✓	✓	?	✓	✓	?	?	?	✓	✓
Fagerberg 2005[91]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Garcia 2006[59]	✓	✓	✓	✓	✓	✓	?	?	✓	✓	?
Henderson 2012[30]	✓	✓	?	x	?	?	?	?	✓	✓	✓
Kok 2012[39]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lee 2013[90]	✓	?	✓	x	?	✓	x	✓	✓	✓	✓

Study	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Whole sample verified using reference standard?	Same reference standard used?	Were the patients newly diagnosed?	Reference standard results blinded?	Index test results blinded?	Same clinical data as used in test results available in practice?	Uninterpretable/intermediate test results reported?	Withdrawals explained?
Li 2006[60]	?	✓	✓	✓	?	?	?	?	✓	✓	?
Licata 2012[61]	✓	✓	?	✓	✓	?	✓	✓	✓	✓	✓
Limburg 2000[78]	?	✓	?	✗	✓	✓	✓	✓	✓	✓	✓
Manz 2012[84]	✓	✓	?	✓	✗	✓	✓	✗	✓	✓	✓
Otten 2008[73]	✓	✓	✓	✗	?	✓	✓	✓	✓	✓	✓
Schoepfer 2008[75]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Schroder 2007[77]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Shitrit 2007[62]	✓	✓	✓	✓	✓	?	?	?	✓	✓	✓
Sidler 2008[82]	✓	✓	✓	✓	✓	✓	?	✓	✓	✓	✓
Tibble 2002[49]	✓	✓	✓	✓	?	✓	?	?	✓	✓	✓
Tomas 2007[66]	?	?	?	✓	✓	✓	?	?	✓	?	?
Turvill 2012[46]	✓	✓	?	✗	✓	✓	?	?	✓	?	?
Van de Vijver 2012[80]	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓

Appendix 5 Protocol

Protocol for assessment of faecal calprotectin

8th October 2013.

HTA reference number 2012/48.

Title: faecal calprotectin in the differential diagnosis of chronic bowel disease.

Aim: to review the clinical accuracy and cost-effectiveness of faecal calprotectin testing for distinguishing between inflammatory and non-inflammatory bowel disease in people with chronic lower GI symptoms.

NB. This protocol may evolve in the course of the assessment.

Assessment group – Warwick Evidence
Project lead: Norman Waugh
   Professor of public health medicine and health technology assessment
   Warwick Evidence
   Division of Health Sciences
   Warwick Medical School
   Coventry CV4 7AL

   norman.waugh@warwick.ac.uk 02476 151585

Project team:

Dr Ewen Cummins, health economist

Dr Pamela Royle, senior research fellow

Dr Deepson Shyangdan, research fellow

Dr Ramesh Arasaradnam, consultant gastroenterologist, UHCW and associate professor of gastroenterology, Warwick Medical School

Dr Ngianga-Bakwin Kandala, principal research fellow

Introduction

Chronic abdominal pain or discomfort, accompanied by diarrhoea or constipation, is common and the symptoms can be due to a number of different conditions, some more serious than others.

The conditions include irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD). The commonest forms of the latter are ulcerative colitis and Crohn's disease, which make up about 90% (to be checked) of IBD.

Irritable bowel syndrome

The most common symptoms of IBS include recurrent colicky abdominal pain or cramping felt in the lower abdomen and relieved by defecation. There may be abdominal distension (bloating) and altered bowel habit – episodes of diarrhoea and constipation. Features supporting a diagnosis of IBS: symptoms > 6 months, associated with other, non-GI problems, stress worsens symptoms. IBS is very common – perhaps 15% of the UK population, though many people who have it never consult their GPs about it. It is commonest in young women. The underlying mechanism is an alteration in the functioning of the muscle in the wall of the bowel. People who have it are constitutionally well and do not lose weight. It is a troublesome but not serious condition, in the sense that it does not lead to serious adverse events.

The cause of IBS is not known in most people, but it sometimes follows an episode of infectious gastroenteritis ("food poisoning").

Inflammatory bowel disease

Ulcerative colitis is characterised by inflammation of the colon, sometimes intense, with bloody diarrhoea, but often much milder.

Crohn's disease can present in different ways. It is also called "regional enteritis" but this is somewhat misleading because Crohn's disease can affect any part of the GI tract.

Both UC and Crohn's can cause autoimmune disorders in other parts of the body, including the eye (uveitis), the joints (arthritis), the skin (erythema nodosum) and the bile ducts (sclerosing cholangitis). The onset of Crohn's can be less obvious than that of UC. In children the first sign may be failure to grow.

Colorectal cancer may also cause inflammation.

The key point to note is that distinguishing amongst inflammatory and non-inflammatory diseases by purely clinical means – signs and symptoms – can be difficult. So many patients are referred to gastroenterology.

NICE Clinical Guideline 61 'Irritable Bowel Syndrome' recommends that people presenting with abdominal pain or discomfort, bloating or change in bowel habit for at least six months should be asked if they have any red flag indicators such as unexplained weight loss. They should also be clinically tested for red flag indicators including anaemia, rectal masses, inflammatory biomarkers for IBD (FC is not specifically mentioned) and late onset (> 60 years) change in bowel habits. Presence of any of these indicators should result in a referral to secondary care for further investigation.

Therefore, patients presenting with symptoms/test results indicative of IBD are referred to secondary care for specialist investigation (most likely to a gastroenterology clinic).

If there are no red flag indicators to cause concern, the guideline states that patients who meet the IBS diagnostic criteria should receive the following laboratory tests to exclude other diagnoses:

- Full blood count (FBC)
- Erythrocyte sedimentation rate (ESR) or plasma viscosity
- C-reactive protein (CRP)
- Antibody testing for coeliac disease (endomysial antibodies [EMA] or tissue transglutaminase [TTG]).

Calprotectin

Calprotectin is a compound released by white blood cells. In people with bowel conditions that cause inflammation, the increased number of white blood cells in the bowel leads to an increase in faecal calprotectin (FC). There are now tests to detect or measure the level of calprotectin in faeces.

The proposed role of FC testing is, in people with lower gastrointestinal symptoms (pain, bloating, diarrhoea, change in bowel habit), to distinguish between those with inflammatory conditions and those with no inflammation. Many of those with inflammation will have IBD, but others may have cancer or other conditions. Most of those with no inflammation will have IBS.

Knowledge of the presence or absence of inflammation may affect the decision on referral for further investigation. The absence of inflammation may lead to a presumption of IBS, to be managed in primary care. The presence of inflammation would be likely to trigger referral to gastroenterology for further investigation, likely to include colonoscopy or sigmoidoscopy.

If calprotectin testing is cost-effective, the likely effect from testing being made more widely available would that it would become part of the primary care work-up pre-referral. So the main focus of this appraisal is expected to be on use of FC testing by primary care staff. However use of FC testing in secondary care will also be considered.

Hence there could be two benefits. Those with IBS would not be referred and might therefore escape further investigations especially colonoscopy. Those with inflammation might be referred and diagnosed sooner and receive appropriate treatment earlier.

Decision problem

As stated in the scope for this appraisal, the objective of the evaluation is to assess the clinical and cost-effectiveness of faecal calprotectin testing in distinguishing inflammatory from non-inflammatory diseases of the bowel. Scoping workshop feedback suggests that the following questions should be taken into account in guiding this evaluation:

- Is an FC test result a reliable way of identifying inflammation of the bowel?
- How do the different cut-off values used to interpret the results of quantitative FC tests affect their cost-effectiveness? What are the optimal cut-offs?
- What is the cost-effectiveness of the point-of-care tests? How does this compare to the fully quantitative FC tests?
- How will the performance of FC tests be affected when used in primary care, given the paucity of data on the use of these tests in primary care?
- How does performance of FC testing vary amongst primary and secondary care groups?

Methods

Population

Patients with lower gastrointestinal symptoms that are chronic, defined as persisting for at least 6- 8 weeks. All ages will be included. At the scoping workshop it was felt that a lower age of 12 years might be used, but preliminary investigation by the Warwick Evidence team suggests that studies in children and adolescents do not report results separately for the under 12s and over 12s. So we will have no lower age limit. The scope suggests an upper age limit of 60.

In adults, symptoms include abdominal pain or discomfort, bloating, diarrhoea or constipation.

The main focus will be those presenting in primary care but studies of hospital groups will also be included.

Paediatric and adult populations will be analysed separately.

Patients with red flag symptoms (as listed above) will be excluded since they should be referred without delay because such symptoms may be due to cancer.

Intervention

Faecal calprotectin tests. These are of two types;

- Laboratory testing, mostly using ELISA methods.
- Point-of-care tests (POCT), which can be used in primary care or in laboratories.

Lab methods are quantitative. POCT tests may be quantitative or semi-quantitative.

Some POCT testing may be used in smaller laboratories where throughput does not justify ELISA equipment

The scope envisages that the lab-based calprotectin tests can be treated as a group. We will seek expert advice from Biochemistry on this. We may provide a narrative description of these tests in an appendix. We note that differences in extraction buffers might be important.

As the scope reports, cut-offs for FC may be a single point, such as 50 µg/g, so that values below indicate no inflammation and values equal to and above indicate inflammation is present, or multiple cut-offs may be used, with results classed as;

- no inflammation
- indeterminate result (likely resulting in the individual being re-tested at a later date)
- inflammation confirmed.

The scope cites anecdotal evidence suggesting that as many as 85–90% of individuals investigated using an FC test in a gastroenterology clinic will have an FC level of less than 50 µg/g (no inflammation). Of the remaining 10–15%, half will have an indeterminate result (50–200 µg/g). Some clinics use 50–100 µg/g and one study found that most of this group had no abnormal findings, so there may be a case for 100 being the cut-off.

The review will seek to determine the best cut-offs. However, it should be noted that decisions will not be made only on calprotectin levels, but on the whole clinical picture. This raises the question of whether there should be different cut-offs for different patient groups according to symptoms.

One question will be the role of POC testing. Our starting assumption is that a definitely negative POCT need not be checked by a quantitative lab method, but that borderline and positive ones will be re-tested by a lab method. The scope envisages repeat testing after borderline results. After positive testing and referral to gastroenterology, we will assume that repeat testing by quantitative method (ELISA) will be done, partly as a baseline for future monitoring.

Comparators

In primary care, GPs suspecting inflammation can use ESR and CRP, which can indicate inflammation but not localise it. If GPs have access to faecal calprotectin testing, they would use that in people with suspected IBS. So FC would replace ESR and CRP testing.

However it might be more useful to compare pathways of care. A set of possible pathways is shown in appendix 2 in which the options include;

- No FC testing available. Clinical assessment and simple tests in primary care followed by decision on referral or symptomatic treatment/therapeutic trial
- Lab testing available to GPs. Lab provides result.
- "Lab plus" where GP provides clinical details along with test request and gastroenterologist or clinical biochemist provides commentary and advice
- POCT available in primary care.

Outcomes

Depending on data availability, these may include:

- Referral rates
- Numbers of colonoscopies with/without FC testing
- Proportion of colonoscopies with no abnormal findings
- Duration from onset of symptoms
- Costs
- Adverse events such as complications of colonoscopy, late presentation of Crohn's disease
- Quality of life
- QALYs.

Acceptance of the test will not be universal, and may vary amongst primary and secondary care – i.e. some patients might decline to produce a sample of faeces for their GP, but might possibly for a gastroenterologist if the alternative is colonoscopy.

Methods

General approach

A framework of six stages has been used to describe the process of evaluating a diagnostic technology (Fryback and Thornbury, 1991):

1. Technical quality of test information – feasibility and optimisation.
2. Diagnostic accuracy
3. Diagnostic thinking impact – change in referring physician's diagnosis.
4. Therapeutic choice impact – change in patient management plan
5. Patient outcome impact
6. Societal impact – change in costs and benefits
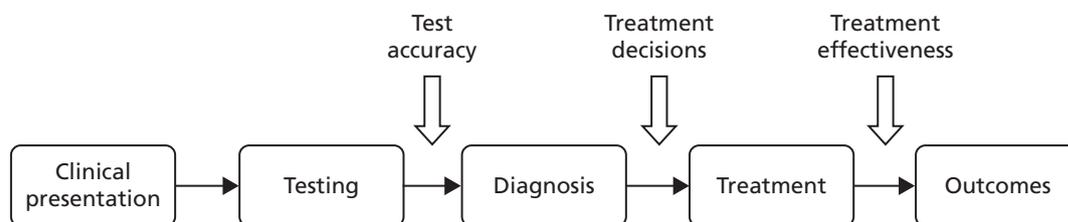
FIGURE 1 Determinants of the clinical effectiveness of a diagnostic technologies (Medical Services Advisory Committee (MSAC), 2005).

We will use a similar approach. The key finding will not be whether the tests reliably measure faecal calprotectin, but whether FC testing improves patient outcomes.

Searches.

Our starting point will be the previous review by the Centre for Evidence-based Purchasing. This review will update that.

We will search MEDLINE, EMBASE, SCI and all sections of The Cochrane Library, for systematic reviews (including any previous health technology assessments) and primary studies.

The search strategy below will be used for MEDLINE and adapted as appropriate for other databases. Searches will be not restricted to English language, in order to provide an impression of the total volume of literature. Some studies not in English may be translated if they look particularly useful, and if translation is available, but most will not.

1. exp Inflammatory Bowel Diseases/di [Diagnosis]
2. exp Irritable Bowel Syndrome/di [Diagnosis]
3. crohn's disease.tw.
4. ulcerative colitis.tw.
5. inflammatory bowel disease*.tw.
6. irritable bowel syndrome*.tw.
7. (IBS or IBD).tw.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. calprotectin.tw.
10. 8 and 9
11. limit 10 to english language

Searches will retrieve studies reporting for adverse events, especially associated with colonoscopy. We will seek data specific to diagnostic colonoscopy since rates are higher after therapeutic colonoscopy.

Selection of studies

Inclusion criteria: studies comparing faecal calprotectin as a guide to inflammation of the lower intestine, ideally with histology as the reference test. Initial searches reveal two problems with this. Firstly, some studies give numbers of patients with CD and UC, but do not give details on whether this is based on biopsy and histology. Secondly, some studies report colonoscopy as normal without further data on whether biopsy is done for e.g. microscopic colitis.

We will also seek follow-up studies of patients diagnosed as IBS.

Since the aim of the appraisal is to assess the usefulness of calprotectin for distinguishing between inflammatory and non-inflammatory bowel disease (in practice mostly between IBS and IBD), to be included studies should have;

A mixed group of patients with symptoms but not yet diagnosed, and ideally a mix reflecting case mix in primary care. It is assumed that coeliac disease has been excluded by TTG testing.

Calprotectin testing before, or blinded to the results of, endoscopy

Endoscopy (usually colonoscopy but sigmoidoscopy only studies would not be ruled out in UC) for all patients, with the endoscopist blinded to the results of calprotectin testing

The reference test of histology of biopsies taken during endoscopy

Hierarchies of evidence

1. The best evidence will come from studies in which FC is carried out in patients with symptoms of recent onset, lasting for at least 6 weeks, where the diagnosis is uncertain, and where colonoscopy is performed to provide a definitive histological diagnosis. We note from preliminary searches that some studies report colonoscopy but do not mention whether biopsies were obtained for histology. Some of these studies give details of UC and CD, and it is likely that histology was available but is simply not mentioned. Depending on numbers of studies, we may carry out a sensitivity analyses with and without studies with no mention of histology.

2. If at colonoscopy, the bowel appears normal, biopsies may still be taken, to check for microscopic colitis. However there may be cases where endoscopy is negative because in about 10% of cases of CD, it is limited to those parts of the small bowel that cannot be reached by colonoscopy or gastroscopy. There will also be cases where endoscopy is deemed to be too invasive. We will accept the following as proof of IBD:

   ○ Wireless capsule endoscopy with score indicating mild or worse activity (score > 134)
   ○ Small bowel capsule biopsy
   ○ Radiological evidence of thickening of the bowel wall
   ○ Ultrasound evidence of thickening of the bowel wall

3. Some studies are simply series of patients with known IBS or IBD with FC test results but no recent endoscopy, or no endoscopy at all if IBS has been diagnosed on purely clinical grounds. We may use these as guides to thresholds in symptomatic cases of recent onset (within 6 months). We will exclude long-standing (over 12 months) cases of IBD and patients with IBD in remission.

4. Some studies report FC results in patients that have had multiple investigations without a definitive diagnosis. This could cause a problem of spectrum bias which is likely to mean that the patients are not representative of those with symptoms of recent onset presenting in primary care. Any such studies will be used only for assessing the value of FC testing in specialist care, or in sensitivity analysis

5. If data permit, we may carry out a sensitivity analysis using only studies that have more than 50% of patients with non-inflammatory conditions, as a guide to NPV and negative LR of calprotectin in primary care. (For adults and children separately.)

Exclusions

- Studies of faecal calprotectin for monitoring activity, or response to treatment, in people with known IBD.
- Patients with IBD in remission will be excluded by absence of symptoms.
- Studies of serum calprotectin.
- Short duration of symptoms (< 6 weeks).
- Patients with symptoms following an acute infectious episode, lasting for less than 3 months.
- Patients over 60.
- Studies with more than 3 months interval between FC and colonoscopy.
- Studies where it is not clear whether symptoms are of recent onset.
- Patients taking NSAIDS or any other drug likely to results in raised FC levels. Low dose (75mg) aspirin will be allowed.
- Studies of patients with mix of long and short duration of symptoms may be useful if the majority (70–80%) are of short duration, or if the short duration group is reported separately. We may consider a sensitivity analysis including/excluding studies.

Where possible, data will be extracted from diagnostic studies for 2 × 2 tables, with FC as screening test and bowel histology as the reference test.

If data for 2 × 2 tables are not available, we will report what screening parameters are provided in studies.

We will rely mainly on studies published in full but may use those available only in other forms abstracts for some purposes, such as identifying emerging research.

Assessment of methodological quality

We will use the QUADAS tool (see appendix 2), possibly modified.
(www.bris.ac.uk/quadas/quadas-2)

Data collection, analysis and synthesis

We will use Review Manager, which now has a section for diagnostic reviews, and can generate coupled forest plots and ROC curves. We will also use MedCalc for producing figures. RevMan cannot do all the statistical analysis that is likely to be required and the statistical software package Stata will be used for more complex analysis.

If the main value of calprotectin testing is to rule out conditions causing inflammation, the key parameters will be negative predictive value (NPV) and negative likelihood ratio. Note that more than one test may be used, so if an initial test was negative but symptoms suggestive of IBD continued despite treatment for IBS, it could be repeated.

Results will need to take account of country of origin since the prevalences of CD and UC vary.

Heterogeneity will initially be examined by visual inspection of coupled forest plots of sensitivity and specificity using the reference standard of endoscopy, ideally with histology.

More variability among diagnostic accuracy study results is to be expected than with randomized trials. Some of this variability is due to chance, as many diagnostic studies have small sample sizes. The remaining heterogeneity may be due to differences in study populations, but differences in study methods are also likely to result in differences in accuracy estimates. Test accuracy studies with design deficiencies can produce biased results.

As recommended in Leeflang et al. 2009, we will investigate and identify potential sources of bias and to limit the effects of these biases on the estimates and the conclusions of the test accuracy.

To address these sources of bias, we will use are sensitivity analysis, subgroup analysis or meta-regression analysis. The STATA software will be used since meta-regression cannot be performed using Review Manager.

We will also report statistics used in diagnostic test accuracy studies: the sensitivity and the specificity, the positive and negative predictive value, the likelihood ratios for the respective test results, or the Receiver Operating Characteristic (ROC) curve and quantities which are can be performed in STATA.

We will also explore two newly developed approaches to fitting random effects in hierarchical models overcome existing limitations: the hierarchical summary ROC model and the bivariate random effects model.

Both models give a valid estimation of the underlying ROC curve and the average operating point. Addition of covariates to the models, or application of separate models to different subgroups enables exploration of heterogeneity. Both models can be fitted with statistical STATA software that fits mixed models.

Cost-effectiveness analysis

This will include the following stages:

- Cost analysis. We note that the NHS Technology Adoption Centre (NTAC) calprotectin pilots are collecting data on referral rates, and that a cost-consequence analysis will be performed by NTAC. It is important that this analysis is available for this appraisal. We will also seek costs from other sources including University Hospital for Coventry and Warwickshire.
- Cost-effectiveness. We will start with the approach used by Hutton and colleagues in the CEP economic assessment, and summarised in their figure 1. However we expect to add another branch for indeterminate or borderline results. In addition, their analysis was largely a cost-consequence analysis, rather than a cost-effectiveness one. It is possible that introducing a calprotectin service for GPs would lead to better outcomes and cost savings, in which case a cost-minimisation analysis would be adequate. However if there are false negatives and false positives, we may need to analyse the trade-offs from adjusting sensitivity and specificity through cost-effectiveness modelling. The CEP report concluded that POCT dominated lab-based testing, but noted that fewer IBD cases were correctly identified.
- The relative cost-effectivenesses of different cut-off points will also be consideration.
- Final decisions on approach will be made in the light of the clinical effectiveness findings.

Subgroups

- Children (under 14) vs adults
- IBD affecting only large bowel
- IBD affecting only small bowel
- Primary care vs secondary care groups as reflected in high proportions with IBS
- UC vs CD

Information from manufacturers

NICE will provide contact details for manufacturers and direct contact will be made as required. We note that there are several versions of some tests. When required, we will ask manufacturers to confirm which versions will continue to be marketed.

Data from manufacturers will not be used if received after 31st December.

Timelines

Progress report to NICE and NETSCC 7/1/13

Draft assessment report to NICE 21/2/13

Final assessment report to NICE 4/4/13

First AC meeting 8/5/13

Appendix 1

Quality assessment items derived from tool (Whiting 2003)

1. Was the spectrum of patients representative of the patients who will receive the test in practice? (representative spectrum)

2. Is the reference standard likely to classify the target condition correctly? (acceptable reference standard)

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? (acceptable delay between tests)

4. Did the whole sample or a random selection of the sample, receive verification using the intended reference standard? (partial verification avoided)

5. Did patients receive the same reference standard irrespective of the index test result? (differential verification avoided)

6. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)? (incorporation avoided)

7. Were the reference standard results interpreted without knowledge of the results of the index test? (index test results blinded)

8. Were the index test results interpreted without knowledge of the results of the reference standard? (reference standard results blinded)

9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (relevant clinical information)

10. Were uninterpretable/intermediate test results reported? (uninterpretable results reported)

11. Were withdrawals from the study explained? (withdrawals explained)

The term "quality assessment" is preferred to the more traditional "risk of bias" term because the latter, as used in systematic reviews such as Cochrane ones, is more associated with assessing internal validity of RCTs. We need to assess external validity through items such as spectrum bias.

Appendix 2 Possible service options

This is just for illustration and other options may be added.
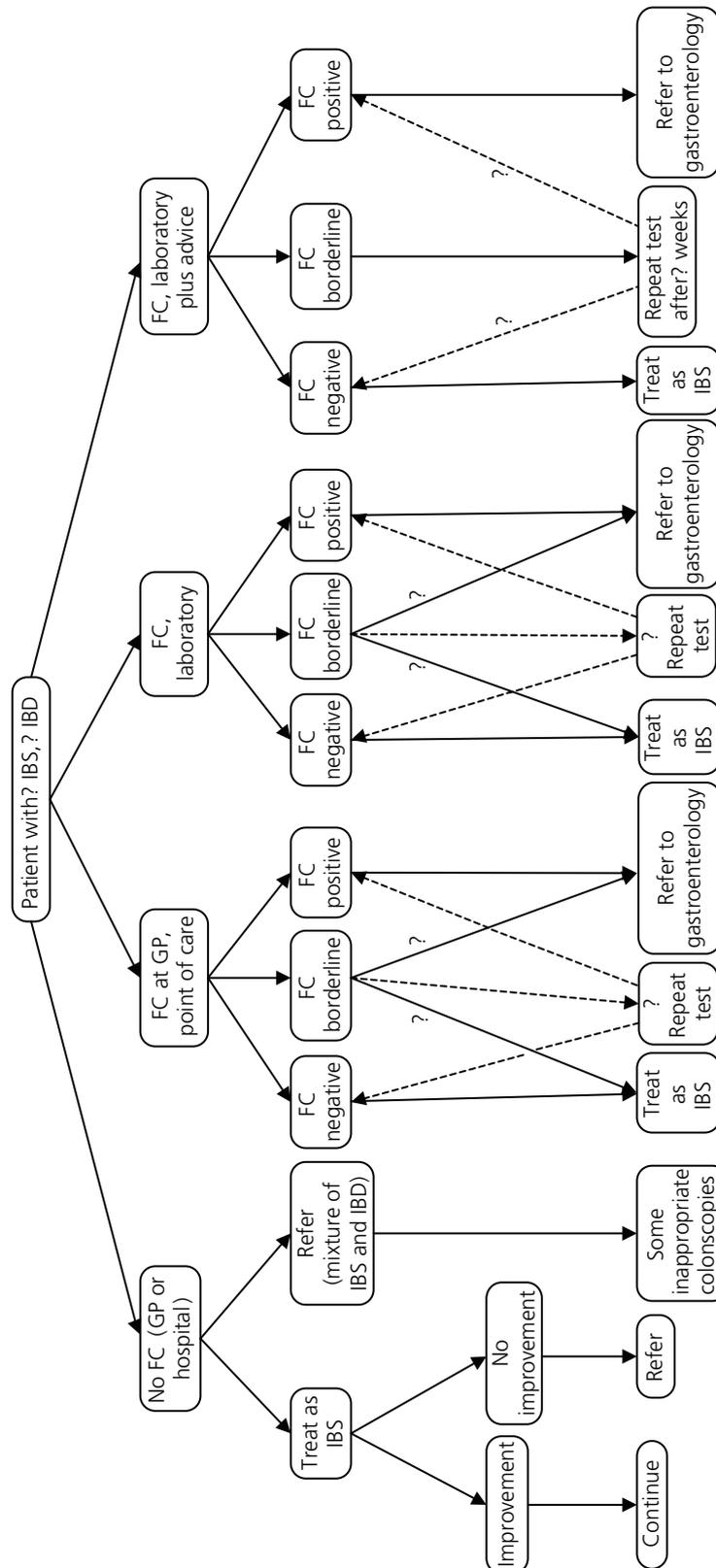
References

Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11(2):88–94.

Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008;149(12):889–97.

Appendix 6 Baseline characteristics of all the included studies

TABLE 77 Baseline characteristics: table A

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Ashorn 2009[89]	To examine the association of FC with serological markers in children and adolescents with IBD	Hospital for children and adolescents in Helsinki, Finland: May 2005 to November 2006	55: IBD (including indeterminate colitis) = 44; non-IBD = 11	Median (range) 13.8 (2.7 to 19.9)	Case series	34/36
Basumani 2012 (poster and personal communication)[83]	To test whether negative calprotectin results act as a screening tool to exclude the need of colonoscopy	District General Hospital, South Yorkshire: June 2009 to May 2011	119: IBD (CD = 6; UC = 6); IBS = 98; others = 9 (non-specific colitis = 2; adenocarcinoma = 1; tubulovillous adenoma = 4; microscopic colitis = 1; lymphocytic colitis = 1)	Adults (ages not reported)	Case series	Not reported
Bharathi 2005 (meeting abstract)[88]	To assess the NPV of FC in excluding bowel pathology in young patients with suspect IBS	Not reported	58: IBS = 42; non-intestinal pathology = 2; miscellaneous = 2; no significant pathology = 12	Not reported	Case series	Not reported
Burri 2012[85]	To compare three different assays in their ability to identify patients with organic intestinal disease	Department of Gastroenterology of the University Hospital Basel in Switzerland: July 2005 to August 2006	405: organic intestinal disease; significant findings = 143; non-significant findings = 262	Median (range) 63 (18 to 97)	A post-hoc analysis of a prospective case series study (Manz 2012)	179/226
Canani 2006[65]	To prospectively evaluate the diagnostic accuracy of FC, ASCA/pANCA, IP, and BWUS, independently and in combination	Paediatric Gastroenterology Unit, Naples, Italy: January-December 2003	45: IBD (CD = 17; UC = 10); non-IBD = 18 [functional disorders = 8, food allergy-related intestinal diseases = 5 (coeliac disease = 5, eosinophilic enterocolitis = 2, acute intermittent IgE-mediated disease = 1), infectious enterocolitis = 4, and Mediterranean familial fever = 1]	Median (IQR) CD group = 14.5 (5.1) UC group = 11.0 (5.0) Non-IBD group = 11.0 (3.3)	Prospective case series	CD = 6/11 UC = 6/4 Non-IBD = 12/6

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Carroccio 2003[87]	To evaluate the positive and NPVs of the FC assay in identifying the organic causes of chronic diarrhoea	Outpatient clinics of the Division of Internal Medicine of the University Hospital and of the Paediatric Division of 'Di Cristina' Hospital, both in Palermo, Italy: January to June 2001	120: (70 adults; 50 children) Children: IBS = 15; organic diarrhoea = 35; cow's milk intolerance = 15; coeliac disease = 13; multiple food intolerances = 5; intestinal giardiasis = 2 Adults: IBS = 40; organic diarrhoea = 30; coeliac disease = 10; multiple food intolerances = 2; colorectal cancer/adenomatous polyps = 3; microscopic colitis = 2; diverticulosis/diverticulitis = 4; CD = 9	Median (range) Adult group = 35 (18–72) Paediatric group = 3.5 (8 months to 10 years)	Prospective case series	Adults = 30/40 Children = 20/30
Damms 2008[58]	To evaluate the diagnostic accuracy of the new calprotectin rapid test compared with an established ELISA test	Gastroenterological departments of three hospitals and three outpatient gastroenterologies, based in Stuttgart, Germany	84: diverticulosis = 18; adenoma = 29; carcinoma = 8; active IBD = 18; intestinal infections = 11	Mean (range) 58 (20 to 85)	Open multicentre case–control study	62/78
Diamanti 2010[86]	To assess the diagnostic precision of a FC assay, compared with histology, as a stool-screening biomarker for IBD	Gastroenterology and Nutrition Unit of 'Bambino Gesu' Children's Hospital, Rome: January 1999 to December 2007	197: IBS (normal mucosa) = 28 aspecific colitis and benign lymphoid hyperplasia = 52 IBD = 117 (UC = 58; CD = 49)	Median (range) Group B (normal mucosa) = 9 (2 to 18) Group C (aspecific colitis and benign lymphoid hyperplasia) = 11 (1 to 18) Group D (IBD) = 12 (1 to 18)	Prospective case series	Group B = 15/13 Group C = 29/23 Group D = 65/52

TABLE 77 Baseline characteristics: table A (*continued*)

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Dolwani 2004[48]	To compare the utility of a single FC estimation to BaFT in exclusion of intestinal inflammation	Majority of cases were recruited from a gastroenterology outpatient clinic with a special interest in IBD at the University of Wales Hospital	63: small intestinal Crohn's disease = 6, jejunal lipoma = 1, Roux loop ulceration = 1, caecal carcinoma = 1; Crohn's colitis = 5; coeliac disease = 1; normal BaFT = 48	Median (range) 47 (17 to 86)	Comparative case series	20/43
El-Badry 2010[81]	To evaluate the sensitivity and diagnostic accuracy of FC assay at different cut-off values in discriminating between functional and organic GI disorders	Internal Medicine Department in Cairo University Hospital: November 2008 to August 2009	29: IBS = 20 IBD = 9 (UC = 8; CD = 1)	Mean (SD) IBS group = 39.4 (15.9) IBD group = 34.3 (15.75)	Prospective case series	IBS = 9/11 IBD = 6/3
Fagerberg 2005[91]	To determine if FC can be used as a diagnostic test of colonic inflammation to identify those children who require colonoscopy	Department of Gastroenterology, Astrid Lindgren Children s Hospital, Stockholm, and the Department of Pediatrics, Central Hospital, Vasteras, Sweden	36: IBD = 22 (CD = 10; UC = 7; indeterminate colitis = 3; juvenile colonic polyposis = 1; unspecified proctitis = 1); functional bowel disorder = 5; food intolerant = 4; spirochetosis = 1, one teenager had nutritional iron-deficiency anaemia = 1, improved spontaneously = 3	Median (range) inflamed group = 13.6 (6.7 to 17.8) Non-inflamed group = 14.2 (6.5 to 17.3)	Prospective comparative case series	Inflamed = 11/11 Non-inflamed = 6/8
Garcia 2006 (translated from Spanish)[59]	To assess the usefulness of FC to predict the presence of pathological colonoscopy	Hospital Universitario Reina Sofia Córdoba, Spain	190: normal colonoscopy = 117 (of which IBS = 9); colon adenoma (polyps) = 28; colorectal cancer = 20; IBD = 25 (UC = 9; CD = 16)	Mean normal colonoscopy = 59.5 Colon adenoma (polyps) = 60.2; colorectal cancer = 71.8; IBD = 35.8	Prospective case series	Normal colonoscopy = 69/48 Colon adenoma (polyps) = 9/19 colorectal cancer = 13/17 IBD = 11/14

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Henderson 2012[30]	To describe the differences in FC levels between IBD types (CD, UC and IBD-U) and non-IBD disease categories	The paediatric gastroenterology department Royal Hospital for Sick Children in Edinburgh: January 1 2005 to December 31 2010	190: IBD group = 91 (CD = 62; UC = 21; IBD-U = 8) Control (non-IBD) group = 99 (IBS = 32)	IBD group = 12.6 (9.5 to 14.0) Control group = 9.3 (5.2 to 12.7)	Retrospective case control	IBD group = 56/35 Non-IBD control group = 55/44
Kok 2012[39]	To quantify the diagnostic accuracy of 3 biomarker tests for the inclusion or exclusion of OBD in patients with persistent lower-abdomen complaints	Primary care. Data from the CEDAR study in 170 general practices in two regions of the Netherlands: July 2009 to January 2011	382: OBD = 99 [adenoma = 53.5% of OBD [adenoma > 1 cm = 30% of adenomas; ≤ 1 cm = 70% of adenomas); IBD = 19 (19%) of OBD]; Non-OBD = 283	Median (range) 60 (18 to 91)	Data from the CEDAR study, an ongoing, prospective, cross-sectional, diagnostic study	175/211
Lee 2013 (abstract)[90]	To determine whether the manufacturer's cut-off levels for referral are useful in diagnosis in patients with chronic diarrhoea	The tests done in primary and Secondary care during a 3-month period from October to December 2011	122: FC tests Organic disease = 23 (IBD = 9) Functional bowel disorder = 99	Not reported	Case series	Not reported
Li 2006 (translated from Chinese)[60]	To assess the value of FC in differential diagnosis of IBS	Both outpatients and hospitalised patients of Peking University Third Hospital: 2004–5	240: IBS = 60; Chronic inflammation group = 60 (UC = 33; CD = 15); colorectal cancer group = 60; healthy controls = 60	Mean (range) IBS = 51 (18) Chronic inflammation = 42 (16)	Prospective comparative study	IBS = 30/30 Chronic inflammation = 26/34
Licata 2012[61]	To assess the diagnostic performance of FC as a stool-screening biomarker for organic intestinal disease	Gastroenterology outpatient department (Gastroenterology and Hepatology Unit, University of Palermo) March 2004 to May 2009	346: No inflammation = 204 (IBS = 197; diverticular disease = 7); inflammation = 142 [IBD = 82 (CD = 56; UC = 26; microscopic colitis = 6; diverticular disease = 4; polyps = 6; ischaemic colitis = 1; nonspecific colitis = 22; IBS = 21]	Median (range) No inflammation = 38 (18 to 87) Inflammation = 41 (17 to 80)	Prospective case series	No inflammation = 74/130 Inflammation = 71/71

TABLE 77 Baseline characteristics: table A (*continued*)

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Limburg 2000[78]	To assess and compare calprotectin and Hb as stool screening biomarkers for colorectal inflammation	The Mayo Clinic (Rochester, MN) November 1996 to July 1998	110: inflammation = 29 (Crohn's or UC = 16; microscopic or collagenous colitis = 11; other inflammatory conditions = 2); no inflammation = 81 [histologically-confirmed normal mucosa = 49; macroscopically normal mucosa (without biopsy) = 11; polyps (all, 1 cm in diameter) but otherwise normal-appearing mucosa = 21]	Mean (SD) = 57 (16)	Prospective case series	With colorectal inflammation = 10/19; without colorectal inflammation = 30/51
Manz 2012[84]	To prospectively investigate the value of FC as a biological marker for the diagnosis of intestinal organic disease in symptomatic patients	Department of Gastroenterology of the University Hospital Basel in Switzerland	538: No clinically significant finding = 326 [normal finding = 314; hyperplastic polyps = 12]; clinically significant finding = 212 [esophagitis = 31; erosive gastritis/duodenitis = 22; gastric ulcers = 11; gastric carcinomas = 3; colitis/ ileitis = 53 (infectious colitis = 8; CD = 10; UC = 16; diverticulitis = 13; microscopic colitis = 5; ischaemic colitis = 1); adenomatous polyps = 50; colorectal cancers = 17]	Mean (IQR) Referred for colonscopy = 63 (53 to 71) Referred for EGD = 55 (42 to 65)	Prospective case series	Referred for colonoscopy = 173/218 Referred for EGD = 75/147

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Otten 2008[73]	To evaluate the diagnostic accuracy of two new rapid calprotectin and lactoferrin faecal tests in assessing colonic inflammation	Gelderse Vallei Hospital, The Netherlands: May to June 2007	114: IBS = 91; IBD = 23 (CD = 6; UC = 5; unspecified colitis = 12)	Mean IBS = 52.3 IBD = 44.5	Cross-sectional design	IBS = 42/49 IBD = 11/12
Schoepfer 2008[75]	To measure the accuracy of faecal markers, CRP, blood leucocytes, and IBD antibodies for discriminating IBD from IBS	Outpatients and inpatients from the Departments of Gastroenterology of the University Hospital Bern and Kantonsspital Lucerne: April 2005 to October 2006	94: IBD = 64 (CD = 36; UC = 28; IBS = 30	Mean (SD) [range] CD = 41 (18) [20 to 78] UC = 45 (14) [23 to 72] IBS = 40 (19) [20 to 79]	Prospective case series	Overall IBD = 32/64 CD = 17/19 UC = 15/13 IBS = 8/22
Schroder 2007[77]	To evaluate calprotectin, lactoferrin and polymorphonuclear neutrophil-elastase in faeces to detect active GI inflammation	J.W. Goethe-University, Frankfurt am Main, Germany (1st Department of Internal Medicine) August 2002 and January 2004	76: IBD = 45 (CD = 25; UC = 20); IBS = 31	Median (range) CD = 40 (25 to 59) UC = 38 (24 to 75) IBS = 43 (20 to 72)	Prospective case series	Total = 33/43 CD = 7/18 UC = 15/5 IBS = 11/20
Shitrit 2007[62]	To assess the predictive value of FC in organic colonic disease	Department of Gastroenterology, Shaare Zedek Medical Center, Israel	69: Abnormal histology = 28 [IBD = 11 (CD = 7); carcinoma = 12; polyps = 5]; normal histology = 41	Not reported	Case series	Not given
Sidler 2008[82]	To define the appropriate roles for faecal S100A12 and calprotectin in the initial investigations of children with suspected IBD	Gastroenterology Outpatient Clinic at Sydney Children's Hospital, Randwick, Australia	61: IBD = 31 (CD = 30; UC = 1); non-IBD = 30	Mean (SD) [range] IBD = 11.9 (3.3) [2.4 to 16] Non-IBD = 10.3 (3.6) [2.2 to 15.5]	Prospective case series	IBD = 19/12 Non-IBD = 17/13

TABLE 77 Baseline characteristics: table A (*continued*)

Study ID	Aim of study	Setting and dates	No. of participants	Age (years)	Study design	Gender (M/F)
Tibble 2002[49]	To determine if the use of FC and IP are useful in differentiating between patients with organic and non-organic disease	Gastroenterology outpatient department of a teaching hospital in South London	602: Organic disease = 263 [small bowel (CD = 84; Coeliac disease = 12; infective diarrhoea = 9; small bowel enteropathy = 21; diabetic diarrhoea = 50); colonic (CD = 18; UC = 87; microscopic colitis = 5; collagenous colitis = 1; diverticular disease = 14; cancer = 7)] non-organic disease = 339 (IBS = 275; IBS + non-ulcer dyspepsia = 38; IBS + other = 26)	Median (range) 40 (18 to 90)	Prospective case series	231/371
Tomas 2007[66]	To evaluate FC in paediatric patients with signs and symptoms suggestive of IBD	Patients with GI symptoms who had been referred to the Paediatric Gastroenterology Unit of Son Dureta University Hospital (Palma de Mallorca), Spain: 2003–2005	43: Functional pathology = 13; organic disease = 30 [IBD = 15]	Mean (range) 10.1 (3 months to 15.3 years)	Retrospective case series	Not reported
Turvill 2012[46]	To determine the NPV of a normal FC in excluding organic intestinal disease in patients with intestinal symptoms	The Department of Gastroenterology, York Hospital: January 2004 to May 2007	630: Cohort 1 (normal FC) = 500; cohort 2 (raised FC) = 130 organic disease = 109; non-organic disease = 521	Mean (range) 41 (16 to 60)	Retrospective cohort study from patient records.	Cohort 1 (FC < 50): 145/355

Cohort 2 (FC > 50): 43/87 |
| Van de Vijver 2012[80] | To determine a diagnostic strategy to minimise the number of patients with negative endoscopy results without missing any cases of IBD | Paediatric outpatient clinics of six general hospitals and one tertiary care hospital in the northern region of the Netherlands: February 2009 to June 2010 | 117: IBD = 42 (CD = 24; UC = 16; IBD unclassified = 2); non-IBD = 75 | Mean (range) 14 (6 to 18) | Prospective diagnostic accuracy study | Confirmed IBD = 19/26

Non-IBD = 38/37 |

BWUS, bowel wall ultrasonography measurement; IP, intestinal permeability.

TABLE 78 Baseline characteristics: table B

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Ashorn 2009[89]	Suspicion of IBD	Not reported	PhiCal Test, produced by NovaTec Immundiagnostica, Dietzenbach, GmBH, Germany – ELISA ≥ 100 µg/g	Not reported	Upper GI biopsies using the modified Sydney system. All underwent upper and lower endoscopies. Diagnosis of IBD made on histopathological criteria	Paediatric Research Foundation. No disclosures
Basumani 2012[83] (poster and personal communication)	Patients were all first presentation referrals with chronic diarrhoea to single clinic. Patients had to have been tested with both FC and histology	Calprotectin level of < 8 and no biopsy; coeliac disease; faecal elastase level of < 200; no calprotectin	Immunodiagnostic by Biohit – ELISA 8, 25, 50, 75, 100 and 150 µg/g	Not reported	Histology	No disclosures. No funding sources mentioned
Bharathi 2005[88] (meeting abstract)	Presented with abdominal pain and/or loose stools	Not reported	PhiCal – ELISA 60 µg/g	Not reported	In 38/42 patients the diagnosis of IBS made after appropriate, targeted investigations (organic gastrointestinal disease, flexible sigmoidoscopy, colonoscopy, ultrasound, small bowel studies, blood tests)	No disclosures. No funding sources mentioned
Burri 2012[85]	Patients undergoing endoscopy of the GI tract for abdominal discomfort at the Department of Gastroenterology of the University Hospital Basel in Switzerland	Patients younger than 18 years were excluded	Calprotectin (EK-CAL, Bühlmann Laboratories, Switzerland) – ELISA; PhiCal – ELISA Both used cut-off levels of ≥ 50 µg/g (manufacturer's recommended cut-off)	Samples collected at home 24 hours prior to bowel preparation for endoscopy – delivered on day investigation. Stored in a refrigerator – transferred to laboratory within 48 hours for analysis	Final diagnosis independently adjudicated by two gastroenterologists on the basis of all the patient's available medical records (clinical data, laboratory values, endoscopy report, histology report). 70 patients (17.3%) also received esophagogastroduodenoscopy	Independent funding. Bühlmann Laboratories AG provided the assays. All authors declared no conflict of interest

TABLE 78 Baseline characteristics: table B (*continued*)

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Canani 2006[65]	All children referred to the paediatric gastroenterology unit for initial assessment of suspected IBD	Patients with symptoms or signs (right lower quadrant mass or perianal disease or haematochezia) mandating a complete work-up for IBD	Calprest, Eurospital Spa, Trieste, Italy – ELISA 95.3 mg/g	Stool samples were collected before endoscopy, stored at –20 °C and thawed at room temperature before testing	Three expert paediatric gastroenterologists made a final diagnosis of IBD or non-IBD according to the presence or absence of previously reported clinical, radiographic, endoscopic and histopathological criteria. Bowel wall ultrasonography performed in all patients within 24 hours of admission	No disclosures. No funding sources mentioned
Carroccio 2003[87]	A history of chronic diarrhoea of unknown origin, lasting for more than 4 weeks, with or without abdominal pain	Previous evaluation for chronic diarrhoea; overt GI bleeding, sigmoidoscopy or colonoscopy during the previous 2 years for any cause; familial adenomatous polyposis and hereditary on polyposis; colorectal cancer	Calprest; Eurospital – ELISA 50 µg/g and 100 µg/g (< 50 = negative; 50–100 = borderline; > 100 = positive)	One stool sample collected and returned within 1 week of first visit. Samples stored at –20 °C. Two aliquots from a single stool sample from each participant were assayed within 4 weeks, and the mean of the two measurements recorded	All patients evaluated by the Rome criteria for IBS and haematology and chemistry tests. Adults also had biopsy with sigmoidoscopy or colonoscopy. If positive at the first-step, then underwent a diagnostic work-up. Children with positive occult blood in the stool or with serum indices of inflammation had colonoscopy with biopsy	No disclosures. No funding sources mentioned
Damms 2008[58]	At least 18 years old and underwent colonoscopy according to a medical indication (GI disorders) or for CRC screening– preventative medical check-up	Known extraintestinal inflammatory diseases. Patients whose medical history included NSAID drugs or anticoagulants	Bühlmann Laboratories AG, Schönenbuch, Switzerland – ELISA Prevista GmbH & Co KG, Munich, Germany – semiquantitative rapid test 50 µg/g (manufacturer's recommended cut-off)	A cold chain was maintained for all samples throughout. Specimens preserved at –20 °C, and assayed within the next 3 months. Patients provided a single stool sample before colonoscopy	Patients underwent colonoscopy according to a medical indication (GI disorders) or for colorectal cancer screening-preventative medical check-up	No disclosures. No funding sources mentioned. Support for analytical kits from Prevista GmbH, Munich

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Diamanti 2010[86]	Recurrent abdominal pain and altered bowel habit, associated with one or more symptoms including rectal bleeding; loss of weight; abnormality of physical examination; delay of pubertal development; positive clinical history; altered blood tests	Not given	Calprest, Eurospital, Trieste, Italy – 100 µg/g and 160 µg/g	Two stool samples were collected before purgation and an average of the two calprotectin measurements was recorded. FC was measured on frozen (–20°) stool specimens	Colonoscopies including intubation of terminal ileum. Mucosal biopsy samples were taken from the terminal ileum, the caecum, the ascending colon, the transverse colon, the descending colon, and the rectum. An experienced paediatric histopathologist, assessed each biopsy specimen	No disclosures. No funding sources mentioned
Dolwani 2004[48]	Consecutive patients undergoing small bowel BaFT examination as part of their clinical work-up after presenting with abdominal pain and or diarrhoea	Patients with: known malignancy, on NSAIDs, coeliac diseases, severe cardiopulmonary, renal or hepatic impairment, significant	ELISA-based method (supplied by Calprotech, London, UK) – brand not specified 60 µg/g (determined by comparing results and using coordinates of ROC curves)	One faecal sample within 7 days before, or 7–10 days after, BaFT. A portion of stool frozen at –20°C and then samples were assayed into batches. On days 0, 1, 3 and 7 a sample of stool was taken and frozen for later analysis	Patients undergoing BaFT also underwent rigid sigmoidoscopy and stool cultures as part of their workup and those with abnormal rigid sigmoidoscopy or positive stool cultures were excluded	Funded by grant from the Wales office of Research Department No disclosures
El Badry 2010[81]	Presence of symptoms for at least 6 months suggestive of organic pathology, including intense abdominal pain, chronic diarrhoea, weight loss and/or anorectal bleeding. Also, to have had an endoscopic and/or intestinal radiological procedure at the initial hospital visit	Regular intake of NSAIDs, aspirin, and/ or anti-coagulants, or the concomitant presence of other non-GI diseases, e.g. rheumatoid arthritis, other connective tissue inflammatory diseases or liver cirrhosis	PhiCal, produced by Nycomed Pharma – ELISA >50 µg/g and >100 µg/g (normal reference value was <15µg/g and the manufacturer established the margin values between 50 and 100 µg/g)	A single stool sample was collected from each patient and stored in a suitable container at –20°C until assayed for calprotectin. The faecal sample was delivered 3 days before colonoscopy	All patients evaluated using ROME III criteria for IBS. Also, had full medical history with thorough clinical examination; stool analysis and culture to exclude infections; complete blood picture, abdominal ultrasound examination; and complete colonoscopy with intubation of the terminal ileum including multiple biopsies from the lesions for histopathological evaluation	No disclosures. No funding sources mentioned

TABLE 78 Baseline characteristics: table B (*continued*)

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Fagerberg 2005[91]	Children with GI symptoms who were scheduled for colonoscopy to rule out IBD	Had no bacterial gastroenteritis detectable by faecal culture or serology and did not have any other chronic inflammatory disease	Calprest, Eurospital SpA, Trieste, Italy – ELISA ≥ 50 µg/g	The stool samples were prepared and analysed for calprotectin according to the manufacturer's instructions. Stool was sent the same day or next day by mail to the laboratory. After extraction, the supernatant was collected and frozen at –20 °C	Decision to perform a colonoscopy based on the child's medical history, physical examination and blood tests. In general, complete ileocolonoscopies were performed. Histopathological evaluations of colorectal inflammation were made for diagnosis of IBD. All children with non-inflamed colonic mucosa had a complete colonoscopy, as did most of the children with inflammation	County Council of Bvastmanland, Swedish Society of Medicine, Mayflower Charity Foundation for Children, Sweden. No disclosures
Garcia 2006[59] (translated from Spanish)	Consecutive individuals who underwent colonoscopy for medical indications	Patients with severe cardiopulmonary disease, kidney or liver disease, coeliac disease, known malignancy and patients with other organic processes in different colonoscopy polyps, colorectal cancer or IBD	Calprest, Eurospital, Trieste, Italy – ELISA 217 mg/kg	All patients collected a stool sample 1 day before the colon preparation for the determination of the FC	All patients had a complete colonoscopy. The diagnosis of IBD was based on clinical criteria, endoscopic and histological findings	Funded by Eurospital and Schering Plough No disclosures
Henderson 2012[30]	All patients potentially undergoing endoscopy before 18 years of age in South East Scotland. IBD group: All incident cases of PIBD diagnosed by standard criteria and had FC measured. Control patients identified from hospital records and departmental endoscopy lists	Aged < 1 year or > 18 years of age on the endoscopy date; a greater than 6-month delay between the FC sample and the endoscopy date; FC sample taken after endoscopy; any previously known, hospital diagnosed, GI disease; and previous upper or lower GI endoscopy	PhiCal Test > 50, 100, 200, 300, and 800 µg/g (< 50 µg/g: normal; 51–100 µg/g: possible GI inflammation; 101–200 µg/g: GI inflammation; > 200 µg/g: active GI inflammation)	FC measured according to manufacturer's instructions	IBD patients: all incident cases diagnosed by standard clinical, histological and radiological findings. Non-IBD (control) patients: all had undergone both upper and lower endoscopy for the clinical suspicion of PIBD	Medical Research Council project grant; Chief Scientist Office in Scotland, and Cure Crohn's Colitis. No conflict of interest

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Kok 2012[39]	Patients consulting their GPs for persistent lower-abdominal complaints; high risk of OBD (as lower abdomen complaints for > 2 weeks, plus one or more of the following: rectal bleeding, altered defecation pattern, abdominal pain, fever, diarrhoea, weight loss, sudden onset in the elderly, physical findings suggestive of OBD	< 18 years old, unable to give informed consent, previously diagnosed with OBD, or positive on triple faeces test, not requiring endoscopy	Quantum Blue Bühlmann Laboratories – POCT EK-CAL (Bühlmann) – ELISA > 50 µg/g (manufacturer's recommended cut-off)	Faecal sample collected directly following inclusion into the study and kept refrigerated at all times. Samples processed directly in 38% of cases and initially frozen in 62%	OBD determined at endoscopy as performed by experienced gastroenterologists, taking biopsies if required according to routine clinical practice. All patients with an inconclusive diagnostic reference procedure followed for 3 months to establish a definite diagnosis. Colonoscopy performed in 91.9% of patients, sigmoidoscopy in 5.5%, and other bowel examinations in 10 2.6%. 89.9% of OBD was confirmed by histology	Funded by Netherlands Organisation for Health Research and Development. Alere Health provided iFOB point-of-care tests. Bühlmann provided the calprotectin POC and ELISA test. No conflict of interest
Lee 2013 (abstract)[90]	Patients presenting with chronic diarrhoea without a pre-existing diagnosis of IBD	Not reported	Not reported ≤ 60 µg/g	FC measured according to manufacturer's instructions	Not clear	Funded by CLRN, NIHR. Disclosures: one author a speaker with Shire Pharmaceuticals; other authors: no disclosures
Li 2006[60] (translated from Chinese)	IBS group: all confirmed by Rome II criteria; chronic inflammation group – confirmed by colonoscopy or operation or pathological diagnosis; Control group: people who were cured from a polyp of the intestinal tract by endoscopic therapy	Absence of overt upper GI symptoms or stomach/small intestine disease; severe disease of the heart, lung, liver, kidney, nerve, or mental disorder. Colorectal adenomas excluded	PhiCal Test (Bio-Rad 550) made by Biohit, Finland – ELISA 50 mg/kg (set according to references as well as the recommended value from the instructions of the assay kit)	Faecal samples collected within 1 week of endoscopy or before surgical operation, transferred to hospital within 2 hours of collection, sealed and stored under −20°C, in preparation for the FC assay	The IBS group all confirmed by Rome II criteria, and no abnormality was seen by colonoscopy or colon contrast. The chronic inflammation group were confirmed by colonoscopy/operation/pathological diagnosis	No funding sources mentioned. No disclosures

TABLE 78 Baseline characteristics: table B (*continued*)

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Licata 2012[61]	Age of at least 18 years; consecutive patients referred to evaluate chronic (≥ 4 weeks) non-bloody diarrhoea of unknown origin	Overt GI bleeding, known colorectal or gastric neoplasia, familial adenomatous polyposis and hereditary non-polyposis colorectal cancer syndrome, history of colonic surgery, recent intake of NSAIDs, aspirin or anticoagulants	Calprest, Eurospital SpA, Trieste, Italy – ELISA > 150 µg/g	Stool samples collected and returned by each participant in a disposable device to avoid toilet water artefacts and to simplify laboratory sampling. Upon receipt, samples were aliquoted for immediate assay or stored at –20 °C until assay	All patients underwent colonoscopy with biopsies. Histopathological evaluations of colorectal inflammation were made by experienced GI histopathologists. Histological inflammation was defined by histological standard criteria and subtyped as: CD, UC, microscopic colitis	No funding sources mentioned. No disclosures
Limburg 2000[78]	Adult patients who had been referred for colonoscopy to evaluate chronic diarrhoea (≥ 4-week duration) of unknown origin or chronic colitis of unknown activity	Abnormalities on GI radiographs, overt GI bleeding, GI endoscopy performed within the preceding 2 weeks, known colorectal neoplasia, familial adenomatous polyposis, and hereditary non-polyposis colorectal cancer syndrome	PhiCal, Nycomed Pharma, Oslo, Norway – ELISA 100 µg/g (preset threshold values supplied by the manufacturer (< 50 mg/g of stool = negative, 50–100 mg/g of stool = weakly positive, and > 100 mg/g of stool = strongly positive)	One stool specimen was collected and returned by each participant using a disposable device to avoid toilet water artefact and simplify laboratory sampling. Upon receipt, samples were aliquoted for immediate essay or stored at –70°C until assay performance	Colonoscopies were performed by experienced staff gastroenterologists, unaware of the faecal assay results. Caecal intubation, coupled with ≥ 90% mucosal surface visualisation, constituted a complete examination. Mucosal abnormalities were recorded by anatomic subsite and biopsies were obtained when clinically indicated	NIH Grant plus a grant from Nycomed Pharma, Oslo, Norway (the manufacturers of the PhiCal ELISA). No disclosures

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Manz 2012[84]	Patients undergoing endoscopy of the GI tract for abdominal discomfort – was defined as any sensation of any quality and intensity of abdominal pain. If several symptoms were present, abdominal discomfort had to be the main symptom	Patients younger than 18 years	EK-CAL, Bühlmann Laboratories AG, Schönenbuch, Switzerland – ELISA 50 µg/g (manufacturer's recommended cut-off)	A single stool sample was collected from each participant 24 hours prior to bowel preparation for endoscopy. Samples delivered on the day of the investigation and stored in a refrigerator before transfer to the study laboratory within 48 hours of analysis	All patients underwent standard endoscopies performed by four senior gastroenterologists. Biopsies were collected if appropriate as decided by the endoscopist. Patients with no significant lesion but FC levels > 50 µg/g on initial endoscopy were further investigated with either EGD or colonoscopy. The endoscopists performing the follow-up endoscopy were aware of the reason for the investigation (positive test)	Researchers were independent of funding. One author is an employee of Bühlmann Laboratories. All authors declared no conflict of interest
Otten 2008[73]	Consecutive patients with lower GI abdominal complaints, including bloating, change in defecation frequency or consistency, or blood and mucus in the faeces, referred for endoscopy or sigmoidoscopy to the endoscopy unit	Younger than 18 years of age, a history of colonic surgery and those with iron deficiency	PreventID CalDetect, Preventis, Bensheim, Germany – point-of-care test 15, 50 and 60 µg/g (in line with the instructions of the manufacturer, the test was evaluated positive when at least the second test line appeared)	On the day of endoscopy patients returned their faecal samples. Faecal rapid tests were performed at the endoscopy unit before the procedure. Faecal samples were stored at –20 °C for a maximum of 1 month for ELISA test at a later stage	All patients underwent colonoscopy or sigmoidoscopy according to routine procedure. According to routine clinical practice, the diagnosis was based on the endoscopic picture; biopsies were taken if necessary to confirm the diagnosis. In half of the patients with IBD, biopsies were taken	No disclosures. No funding sources mentioned. Orange Medical supplied the calprotectin tests

TABLE 78 Baseline characteristics: table B (*continued*)

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Schoepfer 2008[75]	Age 18–80 years, complete colonoscopy with intubation of terminal ileum including biopsies, faecal samples delivered from 3 to 1 days before colonoscopy and after the evaluation, an established diagnosis of bowel disease	Incomplete ileocolonoscopy, microscopic colitis, infectious ileocolitis, colorectal cancer, colorectal polyps, history of colorectal or small bowel surgery, regular intake of aspirin and/or an NSAID	PhiCal Test delivered by Calpro AS (Oslo, Norway) – ELISA 15, 50 and 60 µg/g (in line with the instructions of the manufacturer, the test was evaluated positive when at least the second test line appeared)	Patients collected faecal samples in three faecal tubes. Samples collected from the outpatients were sent by urgent mail from Monday to Thursday so that no samples arrived on weekends. All samples were processed within 48 hours after collection	IBD diagnosis based on symptoms, clinical examination, endoscopic findings, histological analysis, radiological work-up, and laboratory tests. All IBS patients fulfilled the Rome II criteria and had an endoscopy of the upper GI tract. The decision to perform an examination of the jejunum and proximal ileum was left to the judgement of the treating gastroenterologist	Funded by Swiss National Science Foundation. No disclosures
Schroder 2007[77]	History of chronic diarrhoea of unknown origin, lasting for more than 4 weeks, with or without abdominal pain	Previous evaluation for chronic diarrhoea, overt GI bleeding, sigmoidoscopy or colonscopy during the previous 2 months performed for any cause, familial adenomatous polyposis and hereditary non-polyposis colorectal cancer syndrome and pregnancy	Immundiagnostik AG, Bensheim Germany – ELISA 15 µg/g (manufacturer's recommended cut-off)	Fresh stool samples provided for determination within 1 week prior to colonoscopy. Samples received within 24 hours of defecation. Upon receipt, stool was aliquoted for immediate assay or stored at –20 °C until assay performance	Colonoscopies with biopsies from each segment of the colon. Mucosal abnormalities recorded by anatomic location. Inflammation was defined and graded by standard histological criteria and subtyped by endoscopic and histological features	Partly funded by the Else Kröner–Fresenius Foundation, Germany. The authors declared no conflict of interest

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Shitrit 2007[62]	Patients referred to the Department of Gastroenterology for colonoscopic examination of various indications, including screening	Intake of NSAIDs and/or antibiotics during the 3 months preceding the study, concomitant serious illness, pregnancy, alcohol abuse, and evidence of a respiratory tract infection	Calprest, Eurospital, Trieste, Italy – ELISA 150 mg/kg (normal level described as 25 mg/kg)	Stool samples collected before colonoscopic preparation. The samples were stored in a household freezer and they were brought on the day of examination. The samples were frozen at –20 °C until assayed	Colonoscopy and histopathology	No disclosures. No funding sources mentioned
Sidler 2008[82]	Age between 2 and 18 years, and presenting with GI symptoms suggestive of an organic gut disease and required further investigation based on clinical assessment. Symptoms included chronic diarrhoea over more than 1 month, bloody stools, and abdominal pain occurring for at least 1 month	Children with a previously established diagnosis of an organic GI disease; infectious gastroenteritis was excluded by at least two negative stool cultures; having used NSAIDs, antibiotics, or corticosteroids in the preceding 2 weeks	PhiCal test, Nycomed Pharma, Oslo Norway – ELISA 50 mg/kg	Prior to admission for GI endoscopy and colonscopy, children provided a stool sample collected at home before the bowel preparation in a sterile collection vessel, stored briefly at –20 °C (home freezer), and transported frozen to the laboratory, where samples were stored at –80 °C until analysis	All children underwent upper GI endoscopy and complete ileocolonoscopy. Multiple tissue samples were assessed by an experienced paediatric histopathologist. Final diagnosis based upon standard diagnostic criteria, including clinical, endoscopic, histological and imaging findings	Partly funded by the Foundation Eugenio Litta, Geneva, and Freiwillige Akademische Gesellschaft, Basel, Switzerland. No disclosures

TABLE 78 Baseline characteristics: table B (*continued*)

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Tibble 2002[49]	Referred to a gastroenterology outpatient department by GPs. All patients had clinical symptoms suggestive of organic intestinal disease or IBS that had not responded to therapy instituted by primary care and were of sufficient severity for further investigation to exclude organic pathology	Previously known diagnosis of IBD, colorectal carcinoma, and serious cardiopulmonary, hepatic, renal, neurological, and psychiatric disease, or referred for investigation of symptoms of oesophageal reflux, symptoms associated with gastroesophageal pathology, or dyspepsia	In-house ELISA – not a commercially available (< 10 mg/l – converted to 50 µg/g)	Patients provided a single stool sample for measurement of calprotectin that was submitted within 48 hours	Patients were classified investigator as having positive or negative Rome I criteria. Each patient underwent one or more invasive diagnostic imaging procedures of the GI tract as the gold standard, appropriate to their symptoms	Supported by NHS Executive South Thames Regional Office. No disclosures
Tomas 2007[66]	New patient referrals from primary care, aged 16–60 years, with intestinal symptoms, defined as any of change of bowel habit, abdominal pain, bloating, mucorrhoea, bleeding, tenesmus or urgency	Not reported	Calprest, Eurospital, Trieste, Italy – 50, 100, 150, 200 µg/g	A single stool sample was collected from each patient in a plastic container, which was sent to the laboratory in < 48 hours; samples were then frozen at –70°C until analysed	Diagnosis based on clinical criteria, laboratory, image and endoscopic test results, in relation with their evolution, and complied with the functional pathology criteria established in the Rome II meeting	No disclosures. No funding sources mentioned

Study ID	Inclusion criteria	Exclusion criteria	FC test and cut-off used	Details of FC testing	Reference standard	Funding source or conflict of interest
Turvill 2012[46]	New patient referrals from primary care, aged 16–60 years, with intestinal symptoms. These were defined as any of change of bowel habit, abdominal pain, bloating, mucorrhoea, bleeding, tenesmus or urgency	Patients with fast-track colorectal symptoms	PhiCal – ELISA 50 µg/g (manufacturer's recommended cut-off)	On request of a FC, a stool sample was delivered by the patient either to the hospital or to their primary care provider. It was then forwarded internally to laboratory	Cohort 1: 43% of patients had a full evaluation of the colon by colonoscopy or barium enema and 60% had supportive histology; 53% had no investigations Cohort 2: extensive intestinal investigation, including colonoscopy, barium enema, barium meal, Computed tomographic enterography, capsule endoscopy and supportive histology in 83% of patients	No funding sources mentioned. The authors declared no conflict of interest
Van de Vijver 2012[80]	Aged between 6 and 18 years of age with abdominal complaints and with a clinical suspicion of IBD who fulfilled the clinical criteria for IBD	Younger children (who have higher normal values of FC)	Calpro, Calpro AS, Lysaker, Norway – ELISA 50 µg/g (manufacturer's recommended cut-off)	After the first presentation at the outpatient clinic, all patients provided a stool sample collected at home	Paediatricians, referred 68 of 117 patients for endoscopy on the basis of a high index of suspicion for IBD. Majority did not need endoscopy to exclude IBD, and had other tests, including stool analyses for bacteria, ova and parasites, gastroscopy, abdominal ultrasound, CT scan, Meckel scan, serology and dietary measurements leading to the diagnosis	No funding sources mentioned. The authors declared no conflict of interest

Appendix 7 Cost-effectiveness model inputs

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Patient characteristics						
Primary care	Age	25	–	–	Deterministic	CG61 assumption[35]
	Female	50.0	–	–	Deterministic	CG61 assumption[35]
Secondary care	Age	16	–	–	Deterministic	Assumption
	Female IBD	38.5	35	91	Beta	Henderson et al. 2012[30]
	Female non-IBD	44.4	44	99	Beta	Henderson et al. 2012[30]
Disease prevalences and standardised morbidity ratios						
Primary care	IBD prevalence	6.3	7	111	Beta	Durham data[50]
	CD : UC	40.4	365	904	Beta	Shivananda et al. 1996[29]
Secondary care	IBD prevalence	47.9	91	190	Beta	Henderson et al. 2012[30]
	CD–UC split	74.7	62	83	Beta	Henderson et al. 2012[30]
CD standardised morbidity ratio		1.38	1.23	1.55	Log-normal	Bewtra, M. et al. 2013[232]
UC standardised morbidity ratio		1.19	1.00	1.35	Log-normal	Bewtra, M. et al. 2013[232]
IBS standardised morbidity ratio		1.00	–	–		Assumption
Test accuracies						
GP	Sensitivity	100.0	7	7	Beta	Otten et al. 2008[73]
	Specificity	78.8	82	104	Beta	Otten et al. 2008[73]
CalDetect	15 µg/g sensitivity	100.0	23	23	Beta	Otten et al. 2008[73]
	15 µg/g specificity	94.5	86	91	Beta	Otten et al. 2008[73]
	60 µg/g sensitivity	60.9	14	23	Beta	Otten et al. 2008[73]
	60 µg/g specificity	97.8	89	91	Beta	Otten et al. 2008[73]
	15 µg/g sensitivity	95.7	22	23	Beta	Hessells et al. 2012[98]
	15 µg/g specificity	53.2	33	62	Beta	Hessells et al. 2012[98]
	60 µg/g sensitivity	87.0	20	23	Beta	Hessells et al. 2012[98]
	60 µg/g specificity	74.2	46	62	Beta	Hessells et al. 2012[98]

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Quantum Blue	30 μg/g sensitivity	95.7	22	23	Beta	Hessells et al. 2012[98]
	30 μg/g specificity	69.4	43	62	Beta	Hessells et al. 2012[98]
	40 μg/g sensitivity	91.3	21	23	Beta	Hessells et al. 2012[98]
	40 μg/g specificity	83.9	52	62	Beta	Hessells et al. 2012[98]
	50 μg/g sensitivity	87.0	20	23	Beta	Hessells et al. 2012[98]
	50 μg/g specificity	83.9	52	62	Beta	Hessells et al. 2012[98]
	60 μg/g sensitivity	78.3	18	23	Beta	Hessells et al. 2012[98]
	60 μg/g specificity	87.1	54	62	Beta	Hessells et al. 2012[98]
ELISA	50 μg/g sensitivity IBD vs. IBS	93.0	83.0	97.0	Gamma detriment	EAG figure 6
	50 μg/g specificity IBD vs. IBS	94.0	73.0	99.0	Gamma detriment	EAG Figure 6
	50 μg/g sensitivity IBD vs. non-IBD	93.0	83.0	97.0	Gamma detriment	EAG Figure 13
	50 μg/g specificity IBD vs. non-IBD	94.0	73.0	99.0	Gamma detriment	EAG Figure 13
	100 μg/g sensitivity IBD vs. non-IBD	93.0	83.0	97.0	Gamma detriment	EAG Figure 16
	100 μg/g specificity IBD vs. non-IBD	94.0	73.0	99.0	Gamma detriment	EAG Figure 16
	50 μg/g sensitivity IBD vs. IBS	100.0	12	12	Beta	Basumani et al. 2012[83]
	50 μg/g specificity IBD vs. IBS	60.2	59	98	Beta	Basumani et al. 2012[83]
	100 μg/g sensitivity IBD vs. IBS	91.7	11	12	Beta	Basumani et al. 2012[83]
	100 μg/g specificity IBD vs. IBS	81.6	80	98	Beta	Basumani et al. 2012[83]
	150 μg/g sensitivity IBD vs. IBS	83.3	10	12	Beta	Basumani et al. 2012[83]
	150 μg/g specificity IBD vs. IBS	85.7	84	98	Beta	Basumani et al. 2012[83]
	50 μg/g sensitivity IBD vs. non-IBD	100.0	12	12	Beta	Basumani et al. 2012[83]
	50 μg/g specificity IBD vs. non-IBD	60.2	59	98	Beta	Basumani et al. 2012[83]
	100 μg/g sensitivity IBD vs. non-IBD	91.7	11	12	Beta	Basumani et al. 2012[83]
	100 μg/g specificity IBD vs. IBS	81.6	80	98	Beta	Basumani et al. 2012[83]

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Colonoscopy	Sensitivity	95.0	–	–	Deterministic	Expert opinion
	Specificity	100.0	–	–	Deterministic	Expert opinion
IBS and IBD false negatives						
IBS	Response to dietary advice	47.0	33.0	57.0	Log-normal	Mearin 2004[155]
	Subsequence non-responders to medication	5.0	–	–	Deterministic	YHEC expert opinion[50]
IBD	False-negative response to dietary advice	0	–	–	Deterministic	Assumption
	False-negative subsequence non-responders to medication	100.0	–	–	Deterministic	Assumption
Time to representation among non-responders, weeks		12	–	–	Deterministic	Expert opinion
Treatment effectiveness CD						
CD induction therapy						
Prednisolone	Duration, weeks	8	–	–	Deterministic	CG152[108] assumption
	Response rate	67.13	–	–	Look-up table	CG152[108]
	Withdrawal rate	11.83	–	–	Look-up table	CG152[108]
Prednisolone + azathioprine	Duration, weeks	8	–	–	Deterministic	CG152[108] assumption
	Response rate	65.74	–	–	Look-up table	CG152[108]
	Withdrawal rate	9.77	–	–	Look-up table	CG152[108]
Adalimumab	Duration, weeks	6	–	–	Deterministic	CG152[108] assumption
	Response rate	62.34	–	–	Look-up table	CG152[108]
	Withdrawal rate	10.47	–	–	Look-up table	CG152[108]
CD induction therapy responders to no treatment		20.0	–	–	Deterministic	Expert opinion

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
CD maintenance therapy						
Azathioprine	Duration (repeated/cycle length), weeks	8			Deterministic	CG152[108] assumption
	Withdrawal to no treatment (conservative)	0.00			Deterministic	CG152[108] assumption
	Relapse to induction (conservative)	5.33	–	–	Look-up table	CG152[108]
CD post maintenance induction therapy						
Prednisolone	Duration, weeks	8	–	–	Deterministic	CG152[108] assumption
	Response rate	59.19	–	–	Look-up table	CG152[108]
Prednisolone + azathioprine	Duration, weeks	8	–	–	Deterministic	CG152[108] assumption
	Response rate	59.31	–	–	Look-up table	CG152[108]
Adalimumab	Duration, weeks	6	–	–	Deterministic	CG152[108] assumption
	Response rate	55.81	–	–	Look-up table	CG152[108]
	Maintenance of response early thereafter (weekly)	87.12	–	–	Look-up table	CG152[108]
	Maintenance of response subsequent (weekly)	96.37	–	–	Look-up table	CG152[108]
Treatment effectiveness UC						
UC induction therapy						
Low-dose ASA	Duration, weeks	8	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	10.9	–	–	Look-up table	UC guideline[109]
	Response rate I no withdrawal	33.4	–	–	Look-up table	UC guideline[109]
High-dose ASA	Duration, weeks	8	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	9.1	–	–	Look-up table	UC guideline[109]
	Response rate I no withdrawal	44.1	–	–	Look-up table	UC guideline[109]
High-dose ASA + topical	Duration, weeks	4	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	11.1	–	–	Look-up table	UC guideline[109]
	Response rate I no withdrawal	52.1%	–	–	Look-up table	UC guideline[109]

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
High-dose ASA + beclometasone	Duration, weeks	4	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	2.3	–	–	Look-up table	UC guideline[109]
	Response rate I no withdrawal	67.6	–	–	Look-up table	UC guideline[109]
Prednisolone	Duration, weeks	8	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	0.0	–	–	Look-up table	UC guideline assumption[109]
	Response I no withdrawal	52.4	–	–	Look-up table	UC guideline[109]
Inpatient	Duration, weeks	1	–	–	Deterministic	UC guideline assumption[109]
	Withdrawal	0.0	–	–	Deterministic	UC guideline assumption[109]
	Response I no withdrawal	91.0	–	–	Deterministic	UC guideline assumption[109]
Surgery	Duration, weeks	1	–	–	Deterministic	UC guideline assumption[109]
	Duration response	100.00	–	–	Deterministic	UC guideline assumption[109]
UC induction therapy responders to no treatment		20.0	–	–	Deterministic	Expert opinion
UC maintenance therapy						
Maintenance duration (repeated/cycle length) weeks		8	–	–	Deterministic	UC guideline assumption[109]
LASA	Withdrawal	9.7	–	–	Look-up table	UC guideline[109]
	LASA loss or remission I no withdrawal	6.8	–	–	Look-up table	UC guideline[109]
Azathioprine loss of remission		5.9	–	–	Look-up table	UC guideline[109]
No therapy, loss of remission		13.5	–	–	Look-up table	UC guideline[109]
UC induction post maintenance therapy						
As for induction therapy		–	–	–	Look-up table	UC guideline[109]

Variable	Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Administration costs per course (£)					
Prednisolone — Derived from multiple NHS reference costs	164	–	–	Log-normals	NHS reference costs 2011–12[154]
Azathioprine — Derived from multiple NHS reference costs	164	–	–	Log-normals	NHS reference costs 2011–12[154]
Adalimumab — Derived from multiple NHS reference costs, 6 weeks	241	–	–	Log-normals	NHS reference costs 2011–12[154]
Adalimumab — Derived from multiple NHS reference costs, 8 weeks	280	–	–	Log-normals	NHS reference costs 2011–12[154]
LASA induction	95			Log-normals	NHS reference costs 2011–12[154]
HASA induction	95			Log-normals	NHS reference costs 2011–12[154]
HAST induction	50			Log-normals	NHS reference costs 2011–12[154]
HASB induction	50			Log-normals	NHS reference costs 2011–12[154]
LASA maintenance	24			Log-normals	NHS reference costs 2011–12[154]
Azathioprine maintenance	24			Log-normals	NHS reference costs 2011–12[154]
Drug costs per course					
Prednisolone — Derived from multiple NHS reference costs	38.10			Deterministic	BNF March 2013[233]
Azathioprine — Derived from multiple NHS reference costs	42.92			Deterministic	BNF March 2013[233]
Adalimumab — Derived from multiple NHS reference costs, 6 weeks	1408.56			Deterministic	BNF March 2013[233]
Adalimumab — Derived from multiple NHS reference costs, 8 weeks	1760.70			Deterministic	BNF March 2013[233]
LASA induction	74.24			Deterministic	BNF March 2013[233]
HASA induction	152.07			Deterministic	BNF March 2013[233]
HAST induction (4 weeks)	180.57			Deterministic	BNF March 2013[233]
HASB induction (4 weeks)	132.59			Deterministic	BNF March 2013[233]
LASA maintenance	74.24			Deterministic	BNF March 2013[233]
Azathioprine maintenance	42.92			Deterministic	BNF March 2013[233]

Variable		Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Utilities						
CD	Remission	0.890	SD 0.130, n = 129		Log-normal	Stark et al. 2010[124]
	No remission	0.610	SD 0.290, n = 97		Log-normal	Stark et al. 2012[124]
UC	Remission	0.910	SD 0.140, n = 138		Log-normal	Stark et al. 2012[124]
	No remission	0.710	SD 0.180, n = 81		Log-normal	Stark et al. 2012[124]
IBS	Remission–no remission increment	0.071	0.020	0.147	Log-normal	CG61[35]
	No remission	0.662	–	–	Deterministic	Brazier et al. 2004[117]
Test staff timings and costs						
Staff time, GP nurse	CalDetect, minutes	15.00	–	–	Deterministic	Expert opinion
Staff time, Grade 6/7	Quantum Blue, minutes	12.50	–	–	Deterministic	Expert opinion
Staff time, Grade 6/7	ELISA, minutes	11.75	–	–	Deterministic	Expert opinion
Staff time, GP nurse	CalDetect, cost	8.32	–	–	Deterministic	Expert opinion
Staff time, Grade 6/7	Quantum Blue, cost	8.65	–	–	Deterministic	Expert opinion
Staff time Grade 6/7	ELISA, cost	8.13	–	–	Deterministic	Expert opinion
GP per appointment		36.00			Deterministic	PSSRU unit costs[107]

Variable	Mean (%)	Low CI/Q/n	Upper CI/Q/n	PSA	Source
Other costs					
301 outpatient Consultant first face to face non-admitted	164	113	194	Log-normal	NHS reference costs 2011–12[154]
301 outpatient Consultant follow-up face to face non-admitted	115	79	142	Log-normal	NHS reference costs 2011–12[154]
301 outpatient Nurse follow-Up face to face non-admitted	85	65	101	Log-normal	NHS reference costs 2011–12[154]
DAP823 Haematology	3.09	1.76	4.18	Log-normal	NHS reference costs 2011–12[154]
DAP831 Virology	7.75	5.25	9.99	Log-normal	NHS reference costs 2011–12[154]
DAP841 Biochemistry	1.23	0.80	1.46	Log-normal	NHS reference costs 2011–12[154]
FZ51Z outpatient colonoscopy no biopsy	276.32	219.30	306.16	Log-normal	NHS reference costs 2011–12[154]
FZ52Z outpatient colonoscopy with biopsy	316.92	283.18	309.29	Log-normal	NHS reference costs 2011–12[154]
FZ51Z Day-case colonoscopy no biopsy	527.24	413.47	577.85	Log-normal	NHS reference costs 2011–12[154]
FZ52Z Day-case colonoscopy with biopsy	570.45	449.52	657.86	Log-normal	NHS reference costs 2011–12[154]
FZ54Z outpatient sigmoidoscopy no biopsy	174.05	110.68	224.91	Log-normal	NHS reference costs 2011–12[154]
FZ55Z outpatient sigmoidoscopy with biopsy	169.84	91.79	214.97	Log-normal	NHS reference costs 2011–12[154]
FZ54Z Day-case sigmoidoscopy no biopsy	445.88	335.10	508.42	Log-normal	NHS reference costs 2011–12[154]
FZ55Z Day-case sigmoidoscopy with biopsy	480.96	373.35	544.25	Log-normal	NHS reference costs 2011–12[154]
FZ37G Inpatient non-surgical elective	4092.94	1648.26	5197.01	Log-normal	NHS reference costs 2011–12[154]
FZ37H Inpatient non-surgical elective	2570.84	1546.94	3173.01	Log-normal	NHS reference costs 2011–12[154]
FZ37I Inpatient non-surgical elective	2574.02	1405.67	3106.13	Log-normal	NHS reference costs 2011–12[154]
FZ37J Inpatient non-surgical elective	1981.62	1327.26	2341.05	Log-normal	NHS reference costs 2011–12[154]
FZ74A Inpatient surgical	8281.13	6646.52	9666.87	Log-normal	NHS reference costs 2011–12[154]
FZ74A Inpatient surgical	6127.84	5148.80	6855.26	Log-normal	NHS reference costs 2011–12[154]

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library