Clinical practice guideline

Irritable bowel syndrome in adults: Diagnosis and management of irritable bowel syndrome in primary care

National Collaborating Centre for Nursing and Supportive Care
commissioned by
National Institute for Health and Clinical Excellence

August 2007

Draft full guideline for consultation

National Collaborating Centre for Nursing and Supportive Care

This guideline was developed by the National Collaborating Centre for Nursing and Supportive Care (NCCNSC) on behalf of the National Institute for Health and Clinical Excellence (NICE). The guideline was commissioned and funded by NICE and developed in accordance with NICE processes and methodologies.

Based at the Royal College of Nursing, the NCCNSC is a partnership of organisations brought together for the purposes of supporting the development of NICE clinical practice guidelines. The partnership is comprised of representatives from the following organisations:

- Centre for Evidence-Based Medicine, University of York
- Clinical Effectiveness Forum for Allied Health Professions
- Healthcare Libraries, University of Oxford
- Health Economics Research Centre, University of Oxford
- Royal College of Nursing
- UK Cochrane Centre.

Disclaimer

As with any clinical practice guideline, the recommendations contained in this guideline may not be appropriate in all circumstances. A limitation of a guideline is that it simplifies clinical decision-making (Shiffman, 1997). Decisions to adopt any particular recommendations must be made by practitioners in the context of:

- Available resources
- Local services, policies and protocols
- The circumstances and wishes of the patient
- Available personnel and devices
- Clinical experience of the practitioner
- Knowledge of more recent research findings

Contents

Guideline development group membership and acknowledgements

Mr Joe Blanchard-Smith
Patient/carer representative
IBS Network

Dr Ian Bullock
Director
NCCNSC

Dr Jamie Dalrymple
Medical Practitioner, Drayton & St Faith's Medical Practice, Norwich
Primary Care Society for Gastroenterology

Ms Sarah Davis
Senior Health Economist
NCCNSC

Ms Jackie Eastwood
Specialist Pharmacist in Gastroenterology, St. Mark's Hospital, Harrow
Royal Pharmaceutical Society of Great Britain

Dr Charles Foster
Senior Researcher, University of Oxford
British Heart Foundation, Health Promotion Research Group

Dr Jenny Gordon
Research and Development Fellow
NCCNSC

Mrs Jenny Gough
Nurse Consultant, Camden Primary Care Trust, London
Royal College of Nursing

Prof Elspeth Guthrie
Professor of Psychological Medicine & Medical Psychotherapy, Manchester Royal Infirmary
Royal College of Psychiatrists

Dr Miranda Lomer
Post-Doctoral Research Fellow, The Rayne Institute, St. Thomas' Hospital, London

British Dietetic Association

Ms Marion Saunders
Patient/carer representative

Ms Theresa Shaw (Chair)
Chief Executive
Foundation of Nursing Studies

Dr Richard Stevens
General Practitioner, East Oxford Health Centre
Royal College of General Practitioners

Dr Maggie Westby
Senior Research and Development Fellow
NCCNSC

Prof Peter Whorwell
Professor of Medicine and Gastroenterology, Wythenshawe Hospital, Manchester
British Society of Gastroenterology

Dr Mary Wilson
Nurse Specialist, Westwood Hospital, North Humberside
Royal College of Nursing

Staff at the NCCNSC who contributed to this guideline

Dr Anayo Akunne
Health Economist

Ms Nahara Martinez-Gonzales
Systematic Reviewer

Mr Paul Miller
Information Specialist

Ms Emma Nawrocki
Administrator

Mr Edward Weir
Centre Manager

The NCCNSC would like to thank the following people for their assistance with developing the guideline

Dr Yoon Loke

Dr Giok Ong

Dr Jacoby Patterson

Dr Joanne Lord

The NCCNSC also acknowledges the expert opinions of the following people

Kevin Whelan
Nutritional Sciences Division, King's College London

Professor E Quigley

Yvonne McKenzie, Liane Reeves, Annabel Drake, Anna Brian, Carolyn Johnstone, Adele Nichol, Kristina Zaremba
Gastroenterology Specialist Group of the British Dietetic Association

Scottish Nutrition & Diet Resources Initiative (SNDRI)

Terminology

Where the term 'carer' is used, this refers to unpaid carers as opposed to paid careworkers.

Abbreviations

ARR: Absolute Relative Risk

BNF: British National Formulary

CAM: Complementary and Alternative Medicine

CBT: Cognitive Behavioural Therapy

CEAC: Cost-effectiveness acceptability curve

CI: Confidence interval

CRP: C-reactive protein **-** is used mainly as a marker of inflammation

ESR: Erythrocyte Sedimentation Rate - is a non-specific measure of inflammation that is commonly used as a medical screening test.

EMA: Anti-endomysium antibodies, inflammatory markers used in the diagnosis of Coeliac disease

FBC: Full blood count

FOB: Faecal occult blood

GDG: Guideline development group

GI: Gastrointestinal

GP: General practitioner

GRADE: Guidelines Recommendations Assessment Development Evaluation

HRQoL: Health-related quality of life

IBD: Inflammatory Bowel Disease, A general term for any disease characterized by inflammation of the bowel. Examples include colitis and Crohn's disease. Symptoms include abdominal pain, diarrhea, fever, loss of appetite and weight loss.

IBS: Irritable Bowel Syndrome

IBS-A: Irritable Bowel Disease with alternating symptoms of diarrhoea and constipation

IBS-C: Irritable Bowel Disease with constipation as primary bowel dysfunction

IBS-D: Irritable Bowel Disease with diarrhoea as the primary bowel dysfunction

ICER: Incremental cost-effectiveness ratio

LY: Life-year

NHS: National Health Service

NICE: National Institute for Health and Clinical Excellence

NNT: Number needed to treat

OR: Odds ratio

PCT: Primary Care Trust

PEG: polyethylene glycol (macrogol)

PSA: Probabilistic sensitivity analysis

PSS: Personal Social Services

QALY: Quality-adjusted life-year

RCT: Randomised controlled trial

RR: Relative risk

SSRI: selective serotonin re-uptake inhibitors

TGTT: Total gut transit time.

TTG: Anti-transglutaminase antibodies, inflammatory markers used in the diagnosis of Coeliac disease

Organisations

DoH	Department of Health
NCCNSC	National Collaborating Centre for Nursing and Supportive Care
NICE	National Institute for Health and Clinical Excellence
RCN	Royal College of Nursing

General glossary

Absolute risk reduction (Risk difference): The difference in event rates between two groups (one subtracted from the other) in a comparative study.

Abstract: Summary of a study, which may be published alone or as an introduction to a full scientific paper.

Acupuncture: An ancient Chinese technique involving the insertion of fine needles just under the skin in specific locations in order to relieve pain and treat a wide variety of complaints. Historically, acupuncture is one component of an overall program of Chinese medicine that includes theory, practice, diagnosis, physiology, and the use of herbal preparations.

Adjustment: A statistical procedure in which the effects of differences in composition of the populations being compared (or treatment given at the same time) have been minimised by statistical methods.

Algorithm (in guidelines): A flow chart of the clinical decision pathway described in the guideline, where decision points are represented with boxes, linked with arrows.

Allocation concealment: The process used to prevent advance knowledge of group assignment in a RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants.

Applicability : The degree to which the results of an observation, study or review are likely to hold true in a particular clinical practice setting.

Arm (of a clinical study): Subsection of individuals within a study who receive one particular intervention, for example placebo arm.

Association: Statistical relationship between two or more events, characteristics or other variables. The relationship may or may not be causal.

Baseline: The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.

Bias: Systematic (as opposed to random) deviation of the results of a study from the 'true' results that is caused by the way the study is designed or conducted.

Biofeedback: a technique in which an individual learns to consciously control involuntary responses such as heart rate, brain waves, and muscle contractions. Information about a normally unconscious physiologic process is relayed back to the patient as a visual, auditory, or tactile signal. These responses are electronically monitored and noted through beeps, graphs, or on a computer screen, which are seen and heard by the participant.

Blinding (masking): Keeping the study participants, caregivers, researchers and outcome assessors unaware about the interventions to which the participants have been allocated in a study

Bloating: fullness or swelling in the abdomen that often occurs after meals

Borborygmus: the rumbling noise produced by the movement of gas through the intestines. The plural of this word is borborygmi.

Bristol Stool chart: A validated, illustrated tool used to define stool type and consistency developed by Dr K W Heaton, Reader in Medicine at the University of Bristol. Copyright Norgine Ltd 2000.

Carer (caregiver): Someone other than a health professional who is involved in caring for a person with a medical condition.

Case-control study: A study in which the amount of exposure to a potentially causative factor in a group of patients (cases) who have a particular condition is compared with the exposure in a

similar group of people who do not have the clinical condition (the latter is called the control group).

Clinical efficacy: The extent to which an intervention is active when studied under controlled research conditions.

Clinical effectiveness: The extent to which an intervention produces an overall health benefit in routine clinical practice.

Clinical impact: The effect that a guideline recommendation is likely to have on the treatment or treatment outcomes, of the target population.

Clinical question: In guideline development, this term refers to the questions about treatment and care that are formulated to guide the development of evidence-based recommendations.

Clinician: A healthcare professional providing healthcare, for example doctor, nurse or physiotherapist.

Cochrane Library: A regularly updated electronic collection of evidence-based medicine databases, including the Cochrane Database of Systematic Reviews.

Cochrane Review: A systematic review of the evidence from randomised controlled trials relating to a particular health problem or healthcare intervention, produced by the Cochrane Collaboration. Available electronically as part of the Cochrane Library.

Coeliac Disease: Coeliac disease (also called celiac disease, non-tropical sprue, c(o)eliac sprue and gluten intolerance) is an autoimmune disorder characterised by damage to all or part of the villi lining the small intestine. This damage is caused by exposure to gluten and related proteins found in wheat, rye, malt and barley, and to a lesser degree in oats.

Cohort study: A retrospective or prospective follow-up study. Groups of individuals to be followed up are defined on the basis of presence or absence of exposure to a suspected risk factor or intervention. A cohort study can be comparative, in which case two or more groups are selected on the basis of differences in their exposure to the agent of interest.

Co-morbidity: Coexistence of more than one disease or an additional disease (other than that being studied or treated) in an individual.

Comparability: Similarity of the groups in characteristics likely to affect the study results (such as health status or age).

Compliance: The extent to which a person adheres to the health advice agreed with healthcare professionals. May also be referred to as 'adherence'.

Confidence interval (CI): The range of numerical values within which we can be confident that the population value being estimated is found. Confidence intervals indicate the strength of evidence; where confidence intervals are wide they indicate less precise estimates of effects.

Confounding: In a study, confounding occurs when the effect of an intervention on an outcome is distorted as a result of an association between the population or intervention or outcome and another factor (the 'confounding variable') that can influence the outcome independently of the intervention under study.

Consensus methods: Techniques that aim to reach an agreement on a particular issue. Formal consensus methods include Delphi and nominal group techniques, and consensus development conferences. In the development of clinical guidelines, consensus methods may be used where there is a lack of strong research evidence on a particular topic. Expert consensus methods will aim to reach agreement between experts in a particular field.

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

Consultation: The process that allows stakeholders and individuals to comment on initial versions of NICE guidance and other documents so their views can be taken into account when the final version is being produced.

Cost-benefit analysis: A type of economic evaluation, which estimates the net benefit to society of an intervention as the incremental (difference in) benefit of the intervention minus the incremental (difference in) cost, with all benefits and costs measured in monetary units. If benefits exceed costs, the evaluation would be a basis for recommending the intervention.

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

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

Cost impact: The total cost to the person, the NHS or to society.

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

Costing study: The simplest form of economic study, measuring only the costs of given interventions.

Cost-utility analysis: A form of cost-effectiveness analysis in which the units of effectiveness are quality-adjusted life-years (QALYs).

Crohn's Disease: a chronic inflammatory disease of the digestive tract and it 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 auto-immune disorders outside the bowel, such as rheumatoid arthritis.

Cross sectional study: Examination of the relationship between disease and other variables of interest as they exist in a defined population assessed at a particular time.

Data extraction tables: Tabulated presentation of data collected from individual studies.

Decision problem: A clear specification of the interventions, patient populations and outcome measures and perspective adopted in an evaluation, with an explicit justification, relating these to the decision which the analysis is to inform.

Decision analytic techniques: A way of reaching decisions, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees that direct the clinician through a succession of possible scenarios, actions and outcomes.

Delphi technique: is a systematic interactive forecasting method based on independent inputs of a panel of selected experts over two or more rounds. Questions are usually formulated as hypotheses and experts are asked to comment. Each round of questioning is followed with the

feedback on the preceding round of replies, usually presented anonymously. Thus the experts are encouraged to revise their earlier answers in light of the replies of other members of the group. It is believed that during this process the range of the answers will decrease and the group will converge towards the "correct" answer. After several rounds the process is complete and the median scores determine the final answer.

Deterministic analysis: A deterministic analysis is one in which the best estimate for each parameter has been used to give a single estimate of cost-effectiveness. It is the opposite of a probabilistic sensitivity analysis (See sensitivity analysis)

Diarrhoea: is a condition in which the sufferer has frequent and watery or loose bowel movements (from the ancient Greek word διαρροή = leakage; lit. "to run through").

Differential Diagnosis: Distinguishing between two or more diseases and conditions with similar symptoms by systematically comparing and contrasting their clinical findings, including physical signs, symptoms, as well as the results of laboratory tests and other appropriate diagnostic procedures. See also Red Flags.

Discounting: Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.

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

Dosage: The prescribed amount of a drug to be taken, including the size and timing of the doses.

Drop-out: A participant who withdraws from a clinical trial before the end.

Economic evaluation: Comparative analysis of alternative courses of action in terms of both their costs and consequences.

Effect (as in effect measure, treatment effect, estimate of effect, effect size): The observed association between interventions and outcomes or a statistic to summarise the strength of the observed association.

Effectiveness: See “Clinical effectiveness”

Efficacy: See “Clinical efficacy”

Endoscopy: A procedure that uses an endoscope to diagnose or treat a condition. There are many types of endoscopy; examples include colonoscopy, sigmoidoscopy, gastroscopy, enteroscopy, and esophogealgastroduodenoscopy (EGD).

Epidemiological study: A study which looks at how a disease or clinical condition is distributed across populations, e.g. across geographical areas or over time, or between age groups.

Evidence: Information on which a decision or guidance is based. Evidence is obtained from a range of sources including randomised controlled trials, observational studies, expert opinion (of clinical professionals and/or patients).

Evidence table: A table summarising the results of a collection of studies which, taken together, represent the evidence supporting a particular recommendation or series of recommendations in a guideline.

Exclusion criteria (literature review): Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.

Exclusion criteria (clinical study): Criteria that define who is not eligible to participate in a clinical study.

Expert consensus: See ‘Consensus methods’.

Extra-colonic symptoms: IBS symptoms that are not directly associated with the GI tract but are not uncommon features of IBS e.g. low back pain, bladder symptoms, thigh pain, gynaecological symptoms

Extrapolation: In data analysis, predicting the value of a parameter outside the range of observed values.

False positive: Positive test diagnostic result in a subject who does not possess the attribute for which the test is conducted. The incorrect labelling of a healthy person following screening.

Flatus: Gas or wind produced in the intestines, mostly as a result of the normal activity of bacteria in the bowel.

Follow-up: Observation over a period of time of an individual, group or population whose relevant characteristics have been assessed in order to observe changes in health status or health-related variables.

Functional Bowel Disorder: In medicine, the term functional bowel disorder refers to a group of disorders which are characterised by chronic abdominal complaints without a structural or biochemical cause that could explain symptoms. Functional bowel disorders include: * Functional dyspepsia* Non-cardiac chest pain (NCCP)* Chronic abdominal pain* Functional constipation* Irritable bowel syndrome (IBS)

Generalisability: The extent to which the results of a study based on measurement in a particular patient population and/or a specific context hold true for another population and/or in a different context. In this instance, this is the degree to which the guideline recommendation is applicable across both geographical and contextual settings. For instance, guidelines that suggest substituting one form of labour for another should acknowledge that these costs might vary across the country.

Generic name: The general non-proprietary name of a drug or device.

Global Improvement: a research study outcome measuring an overall improvement in a group of defined IBS symptoms (e.g. pain, bowel habit, quality of life). Each symptom is given a score and the aggregate of the scores from each symptom forms the global improvement score.

Global improvement score: an aggregate score of groups of IBS symptoms used to measure changes in severity and frequency of symptoms before, during and after treatment interventions.

Gold standard: A method, procedure or measurement that is widely accepted as being the best available, to which a new method is compared.

Good Practice Points: Recommended good practice based on the clinical experience of the Guideline Development Group.

Grey literature: Reports that are unpublished or have limited distribution, and are not included in the common bibliographic retrieval systems.

Gut motility: A term referring to the contractions of the gastrointestinal tract (peristalsis). These contractions cause food to be pushed through the GI tract in a controlled fashion

Harms: Adverse effects of an intervention.

Health professional: Includes nurses, allied health professionals and doctors.

Health economics: The study of the allocation of scarce resources among alternative healthcare treatments. Health economists are concerned with both increasing the average level of health in the population and improving the distribution of health

Health technology assessment: The process by which evidence on the clinical effectiveness and the costs and benefits of using a technology in clinical practice is systematically evaluated.

Health-related quality of life: A combination of an individual's physical, mental and social well-being; not merely the absence of disease.

Hydrogen Breath Test: test for lactose intolerance that measures breath samples for too much hydrogen.

Hypnotherapy: a deep state of relaxation is achieved through focused attention. While in this trance-like state, the unconscious mind is highly receptive to new perspectives and ideas. The use of imagery and positive suggestions at this time can help a client imagine and actually experience herself in the future, as she desires to be. This may make the desired changes happen much faster and with less resistance, as a result of the hypnosis experience

Hypothesis: A supposition made as a starting point for further investigation.

Idiopathic Constipation: Constipation is termed idiopathic when it cannot be explained by any anatomical, physiological, radiological or histological abnormalities. The exact aetiology is not fully understood but it is generally accepted that a combination of factors may contribute to the condition.

Implementation: Introducing the use of the guidance recommendations in practice.

Incidence: The number of new cases of illness commencing, or of persons falling ill during a specified time period in a given population.

Inclusion criteria (literature review): Explicit criteria used to decide which studies should be considered as potential sources of evidence.

Incremental analysis: The analysis of additional costs and additional clinical outcomes with different interventions.

Incremental cost: The mean cost per patient associated with an intervention minus the mean cost per patient associated with a comparator intervention

Incremental cost effectiveness ratio (ICER): The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest.

Incremental net benefit (INB): The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 x QALYs gained) – Incremental cost

Inflammatory Bowel Disease: general term for any disease characterized by inflammation of the bowel. Two of the most common Inflammatory Bowel Diseases are ulcerative colitis and Crohn's disease. Note: Not to be confused with Irritable Bowel Syndrome.

Intervention: Healthcare action intended to benefit the patient, for example, drug treatment, surgical procedure, psychological therapy.

Indication (specific): The defined use of a technology as licensed by the Medicines and Healthcare products Regulatory Agency (MHRA).

Intention-to-treat analysis (ITT analysis): An analysis of the results of a clinical study in which the data are analysed for all study participants as if they had remained in the group to which they were randomised, regardless of whether or not they remained in the study until the end, crossed over to another treatment or received an alternative intervention

Internal validity: The degree to which the results of a study are likely to approximate the 'truth' for the participants recruited in a study (that is, are the results free of bias?). It refers to the integrity of the design and is a prerequisite for applicability (external validity) of a study's findings.

Intrinsic: Factors present within the individual.

Licence: An authorisation from the MHRA to market a medicinal product.

Life-years gained: Average years of life gained per person as a result of the intervention.

Logistic regression model: A data analysis technique to derive an equation to predict the probability of an event given one or more predictor variables. This model assumes that the natural logarithm of the odds for the event (the logit) is a linear sum of weighted values of the predictor variable. The weights are derived from data using the method of maximum likelihood.

Meta-analysis: A statistical technique for combining (pooling) the results of a number of studies that address the same question and report on the same outcomes to produce a summary result. The aim is to derive more precise and clear information from a large data pool. It is generally more reliably likely to confirm or refute a hypothesis than the individual trials.

Multivariate model: A statistical model for analysis of the relationship between two or more predictor (independent) variables and the outcome (dependent) variable.

Narrative summary: Summary of findings given as a written description.

Negative predictive value: The proportion of individuals with a negative test result who do NOT have the disease.

Nominal group technique: a methodology for achieving team consensus quickly when the team is ranking several options or alternatives or selecting the best choice among them. The method basically consists of having each team member come up with his or her personal ranking of the options or choices, and collation of everyone's rankings into the team consensus.

Number needed to treat: The number of patients that who on average must be treated to prevent a single occurrence of the outcome of interest.

Observational study: Retrospective or prospective study in which the investigator observes the natural course of events with or without control groups; for example, cohort studies and case–control studies.

Odds ratio: A measure of treatment effectiveness. The odds of an event happening in the treatment group, expressed as a proportion of the odds of it happening in the control group. The 'odds' is the ratio of non-events to events.

Off-label: A drug or device used treat a condition or disease for which it is not specifically licensed.

Opportunity cost: The opportunity cost of investing in a healthcare intervention is the other healthcare programmes that are displaced by its introduction. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.

Outcome: Measure of the possible results that may stem from exposure to a preventive or therapeutic intervention. Outcome measures may be intermediate endpoints or they can be final endpoints.

Pain score: a research study outcome measuring changes in pain using an aggregate score of pain type, duration, frequency and severity. Scales used vary.

p values: The probability that an observed difference could have occurred by chance, assuming that there is in fact no underlying difference between the means of the observations. If the probability is less than 1 in 20, the p value is less than 0.05; a result with a p value of less than 0.05 is conventionally considered to be 'statistically significant'.

Peer review: A process where research is scrutinised by experts that have not been involved in the design or execution of the studies.

Peristalsis: Synchronized or coordinated contraction of the muscles that propel food content through the gastrointestinal (GI) tract to facilitate normal digestion and the absorption of nutrients. Peristalsis is dependent upon the coordination between the muscles, nerves, and hormones in the digestive tract.

Placebo: An inactive and physically identical medication or procedure used as a comparator in controlled clinical trials.

Positive predictive value: The proportion of individuals with a positive test result who actually have the disease

Prevalence: The proportion of persons with a particular disease within a given population at a given time.

Prognosis: A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.

Proprietary name: The brand name given by the manufacturer to a drug or device it produces.

Psychotherapy: a set of techniques intended to cure or improve psychological and behavioural problems in humans. The commonest form of psychotherapy is direct personal contact between therapist and patient, mainly in the form of talking.

Qualitative research: Research concerned with subjective outcomes relating to social, emotional and experiential phenomena in health and social care.

Quality of life: See "Health-related quality of life"

Quality adjusted life years (QALYs): An index of survival that is adjusted to account for the patient's quality of life during this time. QALYs 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 QALYs gained are the mean QALYs associated with one treatment minus the mean QALYs associated with an alternative treatment.

Quick reference guide (for a guideline): An abridged version of NICE guidance, which presents the key priorities for implementation and summarises the recommendations for the core clinical audience.

Randomisation: Allocation of participants in a research study to two or more alternative groups using a chance procedure, such as computer-generated random numbers. This approach is used in an attempt to ensure there is an even distribution of participants with different characteristics between groups and thus reduce sources of bias.

Randomised controlled trial (RCT): A comparative study in which participants are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between the groups. The random allocation eliminates bias in the assignment of treatment to patients and establishes the basis for the statistical analysis.

Recurrent: A symptom and/or sign that resolves then returns at least once.

'Red Flag' symptoms: A warning term used to indicate further investigation of specific symptoms is warranted to identify potential differential diagnosis.

Reference standard (or gold standard): An agreed standard, for example for a test or treatment, against which other interventions can be compared.

Refractory IBS: people with IBS who do not respond to first line therapies after 12 months and who develop a continuing symptom profile.

Relative risk: The number of times more likely or less likely an event is to happen in one group compared with another (calculated as the risk of the event in group A/the risk of the event in group B).

Reliability/repeatability: The degree of agreement exhibited when a measurement is repeated under identical conditions. Reliability refers to the degree to which the results obtained by a measurement procedure can be replicated.

Remit: The brief given by the Department of Health and Welsh Assembly Government at the beginning of the guideline development process. This defines core areas of care that the guideline needs to address.

Resource implication: The likely impact in terms of finance, workforce or other NHS resources.

Retrospective cohort study: A study in which a defined group of persons with an exposure that occurred in the past and an appropriate comparison group who were not exposed are identified at a time later than when they were exposed and followed from the time of exposure to the present, and in which the incidence of disease (or mortality) for the exposed and unexposed are assessed.

Review of the literature: An article that summarises the evidence contained in a number of different individual studies and draws conclusions about their findings. It may or may not be systematically researched and developed.

Secondary benefits: Benefits resulting from a treatment in addition to the primary, intended outcome.

Selection bias (also allocation bias): A systematic bias in selecting participants for study groups, so that the groups have differences in prognosis and/or therapeutic sensitivities at baseline. Randomisation (with concealed allocation) of patients protects against this bias.

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 (of a search): The proportion of relevant studies identified by a search strategy expressed as a percentage of all relevant studies on a given topic. It describes the comprehensiveness of a search method (that is, its ability to identify all relevant studies on a given topic). Highly sensitive strategies tend to have low levels of specificity and vice versa.

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.

Multi-way simple sensitivity analysis (scenario analysis): two or more parameters are varied at the same time and the overall effect on the results is evaluated.

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

Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (For example, Monte Carlo simulation).

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.

Stakeholder: Those with an interest in the use of a technology under appraisal or a guideline under development. Stakeholders include manufacturers, sponsors, healthcare professionals, and patient and carer groups.

Statistical power: The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.

Stool: solid waste that pass through the rectum as bowel movements. Stools are undigested foods, bacteria, mucus, and dead cells

Stool score: a research study outcome measuring changes in bowel habit using an aggregate score of stool type,stool consistency, stool frequency, complete evacuation. Scales used vary.

Syndrome: a combination of signs and/or symptoms that forms a distinct clinical picture indicative of a particular disorder

Synthesis of evidence: A generic term to describe methods used for summarising (comparing and contrasting) evidence into a clinically meaningful conclusion in order to answer a defined clinical question. This can include systematic review (with or without meta-analysis), qualitative and narrative summaries.

Systematic review: Research that summarises the evidence on a clearly formulated question according to a predefined protocol using systematic and explicit methods to identify, select and appraise relevant studies, and to extract, collate and report their findings. It may or may not use statistical meta-analysis.

Total gastrointestinal transit time (TGTT): the length of time food takes to pass through the gastrointestinal tract from ingestion to excretion. It is estimated using radio opaque markers and

can define three types of delay: right colon (colonic inertial), left colon and recto sigmoid. The exact type of delay may be an important basis for treatment.

Time horizon: The time span used in the NICE appraisal which reflects the period over which the main differences between interventions in health effects and use of healthcare resources are expected to be experienced, and taking into account the limitations of supportive evidence.

Treatment allocation: Assigning a participant to a particular arm of the trial.

Treatment options: The choices of intervention available.

User: Any one using the guideline.

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.

1 EXECUTIVE SUMMARY

The National Institute for Health and Clinical Excellence ('NICE' or 'the Institute') commissioned the National Collaborating Centre for Nursing and Supportive Care (NCC-NSC) to develop guidelines on irritable bowel syndrome (IBS). This follows referral of the topic by the Department of Health and Welsh Assembly Government. This document describes the methods for developing the guidelines and presents the evidence and consensus based recommendations. It is the source document for the NICE (abbreviated version for health professionals); Understanding NICE Guidance, and; Quick Reference Guide versions of the guidelines which will be published by NICE. The guidelines were produced by a multidisciplinary guideline development group and the development process was undertaken by the NCC-NSC.

The main areas examined by the guideline were during the:

- IBS Positive Diagnosis
- Red flags for suspected cancer and other morbidities
- IBS Management focussed on lifestyle advice relating to diet and physical activity, drug and behavioural therapies.
- Referral and follow-up.

This guideline covers areas relevant to the diagnosis and management of IBS reflecting the complete patient journey, from the person presenting with IBS symptoms, positive diagnosis and management, targeted at symptom control. The guideline incorporates Cochrane reviews, published NICE clinical and public health guidance, Health Technology Assessment reports, systematic and health economic reviews produced by the National Collaborating Centre for Nursing and Supportive Care. Recommendations are based on clinical and cost effectiveness evidence, and where this is insufficient, the GDG used all available information sources and experience to make consensus recommendations using nominal group technique.

The care pathway reflects a logical sequencing to what is, in effect, tracking the progress of the patient from entry to primary care through to lifestyle adaptation and therapeutic intervention, enabling the person with IBS to learn to live with this chronic condition. The partnership that the person with IBS forms with their primary care clinician/team is key to this being a positive experience where shared decision making feature strongly in aiming for symptom control. This sequencing has enabled the Guideline Development Group (GDG), supported by the technical team, to look at the evidence reviews, understand the clinical context and consider the patient voice when shaping guidance. Patient experience is at the heart of development. Evidence published after June 2007 was not considered.

Healthcare professionals should use their clinical judgement and consult with patients when applying the recommendations. Recommendations aim to reduce variations in practice, thus improving patient outcomes related to both the diagnosis and continuous management of IBS.

This guidance is intended to be the source document for primary care local policy development. Its success is dependent on the primary health care team and patients working in partnership in implementing key recommendations. The algorithm provides healthcare professionals, patients and carers to visualise the care pathway, summarising clinical and cost effective evidence and consensus decisions.

Key recommendations that are priorities for implementation (numbering corresponds to the abbreviated (NICE) version of the guideline)

1.1.1.1 Primary care clinicians should consider assessment for IBS if the patient reports having had any of the following symptoms for at least 6 months:

- change in bowel habit
- abdominal pain/discomfort
- bloating.

1.1.1.2 Patients should be asked if they have any of the following 'red flag' symptoms:

- unintentional and unexplained weight loss
- rectal bleeding
- familial history of bowel cancer.

Patients should be assessed for:

- anaemia
- abdominal masses
- rectal masses.

Identification of any of the above should result in referral into secondary care for further investigation (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27; www.nice.org.uk/CG027).

1.1.1.3 For a positive diagnosis of IBS to be made, the person must complain of abdominal pain or discomfort which is either relieved by defaecation, or associated with altered bowel frequency or altered stool form. This must be accompanied by at least two of the following four symptoms:

- altered stool passage (straining, urgency, incomplete evacuation)
- abdominal bloating (less common in men), distension, tension or hardness
- symptoms made worse by eating
- passage of mucus.

It should be noted that other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and can be used to support the diagnosis.

1.1.2.1 In people who meet the IBS diagnostic criteria, it is recommended that the following tests should be undertaken to exclude other diagnostic possibilities:

- 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]).

1.1.2.2 The following tests should not be done to confirm diagnosis in people who meet the IBS diagnostic criteria:

- ultrasound
- rigid/flexible sigmoidoscopy
- colonoscopy; barium enema
- thyroid function test
- faecal ova and parasite test
- faecal occult blood
- hydrogen breath test (for lactose intolerance and bacterial overgrowth).

1.2.1.1 People with IBS should be given information that explains the importance of self-help in effectively managing their IBS. This should include information on general lifestyle, physical activity, diet and symptom-targeted medication.

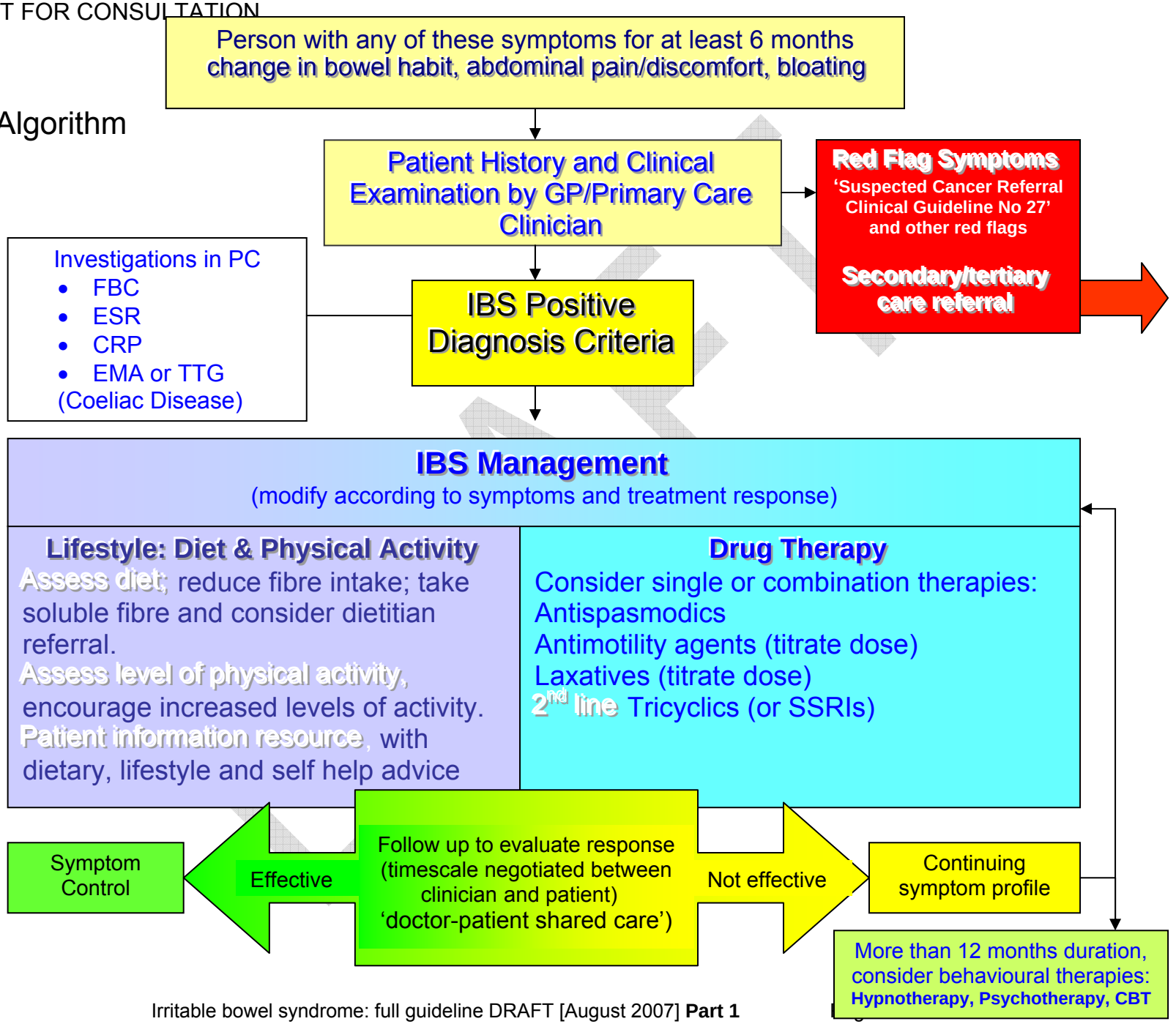
1.2.1.5 Primary care clinicians should review the fibre intake of a person with IBS, adjusting (usually decreasing) it according to effect while monitoring symptoms. People with IBS should be actively discouraged from taking insoluble fibre (bran). If an increase in dietary fibre is advised, this should be soluble fibre (such as ispaghula powder) or foods high in soluble fibre (for example, oats).

1.2.2.4 Primary care clinicians should advise people with IBS how to adjust laxative or antimotility agent doses according to the clinical response. The dose should be titrated according to the stool consistency with the aim of achieving a soft, well formed stool (corresponding to Bristol Stool Chart type 4).

1.2.2.5 Primary care clinicians should consider the benefit of prescribing tricyclics as second-line treatment for people with IBS. Treatment should be initiated at a low starting dose (5–10 mg equivalent of amytriptyline), once at night, which should be reviewed regularly. The dose can subsequently be increased, but does not usually need to exceed 30 mg.

The IBS algorithm demonstrates the importance of positive diagnosis in providing an effective platform for both the person presenting with IBS symptoms and primary care clinician to work towards symptom control. It importantly identifies red flag symptoms, meaning in practice that the person would leave this guideline and be referred to secondary/tertiary care for further investigation.

IBS Algorithm



2 PRINCIPLES OF PRACTICE

The principles outlined below describe the ideal context in which to implement the recommendations contained in this guideline.

These have been adapted from the NICE clinical practice guideline: *Assessment and prevention of falls in older people* (2004).

2.1 Person-centred care

- People who may have Irritable bowel syndrome (IBS) should be made aware of the guideline and its recommendations, and should be referred to the *Understanding NICE Guidance* version of the guideline.
- People who may have IBS should be involved in shared decision–making about individualised IBS management strategies.
- Healthcare professionals are advised to respect and incorporate the knowledge and experience of people who have been self managing this condition.
- People who may have IBS should be informed about any potential risks and/or associated complications with IBS.

2.2 Collaborative interdisciplinary approach to care

- All members of the interdisciplinary team should be aware of the guidelines and all care should be documented in the patient's health care records.
- A collaborative, multi-disciplinary approach should be provided by appropriately trained professionals.
- The roles of parents/carers and health professionals in implementing the guideline recommendations should be sensitively negotiated.

2.3 Organisational issues

- There should be an integrated approach to the diagnosis and management of IBS in Primary Care with a clear strategy and policy supported by management.
- Care should be delivered in a context of continuous quality improvement, where improvements to care following guideline implementation are the subject of regular feedback and audit.
- The health care team should have received appropriate training and have demonstrated their competence in the diagnosis and management of IBS.
- Commitment to and availability of education and training are required to ensure that all staff, regardless of their profession, are given the opportunity to update their knowledge, and are able to implement the guideline recommendations.

- People who have IBS should be cared for by personnel who have undergone appropriate training and who know how to initiate and maintain appropriate management of IBS. Staffing levels and skill mix should reflect the needs of patients.

2.4 Background to the current guideline

In January 2006, The National Collaborating Centre for Nursing and Supportive Care (NCC-NSC) was commissioned by NICE to develop a clinical guideline on the diagnosis and management of Irritable Bowel Syndrome (IBS) for use in Primary Care in England and Wales.

2.5 Clinical need for the guideline

Irritable bowel syndrome (IBS) is one of the most common functional gastrointestinal disorders. It is a chronic, relapsing and often life-long disorder, characterised by the presence of abdominal pain/discomfort associated with defaecation, a change in bowel habit together with disordered defaecation (constipation or diarrhoea or both), the sensation of abdominal distension, and may include associated non-colonic symptoms. These morbidities may cause dehydration, lack of sleep, anxiety and lethargy which may lead to time off work, avoidance of stressful or social situations and significant reduction in quality of life.

People may present with differing symptom profiles, most commonly 'diarrhoea predominant', 'constipation predominant', and alternating symptoms. Clinical management will inevitably be directed by the presenting symptoms, but different symptom types may have differing prognoses that assist in determining the type and urgency of investigations and subsequent management. Symptoms sometimes overlap with other gastrointestinal (GI) disorders such as non-ulcer dyspepsia, or with coeliac disease.

There are three possible diagnostic approaches which may be used; a diagnosis by excluding organic disease which may involve multiple investigative procedures; a diagnosis based on positive symptom criteria, resulting in a minimum of diagnostic tests; a diagnosis combining positive symptom based criteria with investigations to exclude 'red flag' symptoms. In practice diagnosis has been predominantly by exclusion of organic disease which has led to patients being subjected to investigations and tests which are not required to confirm IBS.

Diagnosis and management of IBS can be frustrating for patients and clinicians. Both parties need to have a clear understanding of the current state of knowledge of IBS and recognition of the chronic nature of the condition. The implication is that the management of this condition may involve a long-term therapeutic partnership between the person with IBS and the primary care clinician. There may be many contributing factors to be taken into consideration. Associated non-colonic problems include functional urinary and gynaecological problems, gallbladder and stomach symptoms, back pain, migraine and depression. It has previously been shown that if a non-colonic feature of IBS is especially severe (for example, a gynaecological symptom) the

patient may be referred to the wrong speciality. This may result in unnecessary and sometimes costly investigations and/or delayed treatment. IBS is associated with a disproportionately high prevalence of abdominal and pelvic surgery, although the cause of this has not been established.

IBS most commonly affects people between the ages of 20 and 30 years and is twice as common in women as in men. The prevalence of the condition in the general population is estimated to lie somewhere between 10 and 20%. Recent trends indicate that there is also a significant prevalence of IBS in older people; therefore, IBS diagnosis should be a consideration when an older person presents with unexplained abdominal symptoms. The true prevalence of IBS in the whole population may be higher than estimated, because it is thought that many people with IBS symptoms do not seek medical advice; NHS Direct online data suggest that 75% of people using this service rely on self-care. In England and Wales, the number of people consulting for IBS is extrapolated to between 1.6 and 3.9 million. Evidence suggests that age and race have no consistent effect on the incidence of symptoms.

Causes of IBS have not been adequately defined, although gut hypersensitivity, disturbed colonic motility, post-infective bowel dysfunction or a defective antinociceptive (anti-pain) system are possible causes. Stress commonly aggravates the disorder and around half of IBS outpatients attribute the onset of symptoms to a stressful event. Lactose, gluten or other food intolerance is also identified as an antecedent. Colonic flora may be abnormal in IBS patients. People with IBS tend to alter their diet to alleviate symptoms of IBS, often this is self directed or guidance is sought from inadequately qualified nutritionists. Excluding individual foods or complete food groups without appropriate supervision can readily lead to inadequate nutrient intakes and ultimately malnutrition. In addition, symptoms often remain unresolved leading to further inappropriate dietary restriction.

Primary care investigations may include: routine blood tests such as full blood count, urea and electrolytes, and liver function tests; tests for thyroid function, tissue transglutaminase anti-endomysial antibodies (to exclude coeliac disease); inflammatory markers, stool microscopy; urinary screen for laxatives; and lactose tolerance testing. Other investigations such as gut transit studies (radiological tests to measure the time required for food to move through the digestive tract) and sigmoidoscopy (endoscopy of the lower part of the bowel) are routinely performed in secondary care.

Patients are likely to be referred to a secondary care specialist if symptoms are atypical (for example, patients over 40 years with change in bowel habit and/or rectal bleeding), if GI cancer is suspected on clinical examination, or if there is a family history of GI cancer.

2.6 Management Issues

The aetiology of IBS has not yet been established and as a result management focuses on the relief of symptoms. The symptom profile, as previously described, may vary and may require a combination of different modalities to achieve effective relief. These include diet and lifestyle interventions, patient education and self help, pharmacological interventions, behavioural and psychological therapies, complementary and alternative therapies. No single drug will alleviate the multiple symptoms often present in people with IBS. Management should focus on the predominant symptom which may require concomitant use of medications and other therapeutic interventions. This guideline will review the different therapies commonly used in the management of IBS.

3 SUMMARY OF RECOMMENDATIONS

1. GUIDANCE

1.1 DIAGNOSIS OF IBS

The positive diagnosis of IBS is a key aspect of this guideline. In exploring the multiple features of the syndrome, primary care clinicians should establish symptom profiles, with pain/discomfort being a key symptom. In establishing the quantity and quality of pain/discomfort, clinicians need to identify the site (which can be anywhere in abdomen) and whether it varies. This distinguishes IBS from cancer-related pain/discomfort, which typically has a fixed site.

When establishing the person's bowel habit, the Bristol Stool Chart (see appendix I) helps with description, particularly when determining quality and quantity of stool. People presenting with IBS symptoms commonly report incomplete evacuation/rectal hypersensitivity and urgency, which is increased in diarrhoea-predominant IBS. About 20% of people experiencing incontinence choose not to disclose this unless asked directly. Primary care clinicians should ask people who present with symptoms of IBS open questions to get a feel for the multiple features of the syndrome.

1.1.1 Initial assessment

1.1.1.1 Primary care clinicians should consider assessment for IBS if the patient reports having had any of the following symptoms for at least 6 months:

- change in bowel habit
- abdominal pain/discomfort
- bloating.

1.1.1.2 Patients should be asked if they have any of the following 'red flag' symptoms:

- unintentional and unexplained weight loss
- rectal bleeding

- familial history of bowel cancer.

Patients should be assessed for:

- anaemia
- abdominal masses
- rectal masses.

Identification of any of the above should result in referral into secondary care for further investigation (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27; www.nice.org.uk/CG027).

1.1.1.3 For a positive diagnosis of IBS to be made, the person must complain of abdominal pain or discomfort which is either relieved by defaecation, or associated with altered bowel frequency or altered stool form. This must be accompanied by at least two of the following four symptoms:

- altered stool passage (straining, urgency, incomplete evacuation)
- abdominal bloating (less common in men), distension, tension or hardness
- symptoms made worse by eating
- passage of mucus.

It should be noted that other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and can be used to support the diagnosis.

1.1.2 Diagnostic tests

1.1.2.1 In people who meet the IBS diagnostic criteria, it is recommended that the following tests should be undertaken to exclude other diagnostic possibilities:

- 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]).

1.1.2.2 The following tests should not be done to confirm diagnosis in people who meet the IBS diagnostic criteria:

- ultrasound
- rigid/flexible sigmoidoscopy
- colonoscopy; barium enema
- thyroid function test
- faecal ova and parasite test
- faecal occult blood
- hydrogen breath test (for lactose intolerance and bacterial overgrowth).

1.2 CLINICAL MANAGEMENT OF IBS

1.2.1 Dietary and lifestyle advice

1.2.1.1 People with IBS should be given information that explains the importance of self-help in effectively managing their IBS. This should include information on general lifestyle, physical activity, diet and symptom-targeted medication.

1.2.1.2 Primary care clinicians should give lifestyle advice, encouraging people with IBS to make the most of their available leisure time and ensuring that they create relaxation time.

1.2.1.3 Primary care clinicians should assess the physical activity levels of people with IBS using the General Practice Physical Activity Questionnaire (GPAQ). All sedentary people should receive brief advice and counselling to encourage physical activity.

1.2.1.4 Primary care clinicians should assess diet and nutrition for all people with IBS and provide the following general advice.

- Have regular meals and take time to eat.
- Avoid missing meals, or leaving long gaps between meals.
- Drink at least 8 cups of fluid per day, especially water or herbal teas.
- Restrict tea and coffee to not more than 3 cups per day.
- Reduce intake of alcohol and fizzy drinks.
- It may be helpful to limit high-fibre cereals (such as wholemeal or high-fibre breads and wholegrains).
- Reduce intake of 'resistant starch', which is often found in processed or re-cooked foods, as it may increase symptoms.
- Limit fruit to 3 portions per day (approx 80 g each).
- People with diarrhoea should avoid sorbitol, which is found in sugar-free sweets (including chewing gum) and drinks, and some diabetic and slimming products.
- People with wind and bloating may find it helpful to eat oats (such as oat-based breakfast cereal or porridge) and linseeds (up to one tablespoon per day).

1.2.1.5 Primary care clinicians should review the fibre intake of a person with IBS, adjusting (usually decreasing) it according to effect while monitoring symptoms. People with IBS should be actively discouraged from taking insoluble fibre (bran). If an increase in dietary fibre is advised, this should be soluble fibre (such as ispaghula powder) or foods high in soluble fibre (for example, oats).

1.2.1.6 Primary care clinicians should not discourage people with IBS from trying specific probiotic products. If people with IBS choose to do this, it should be for at least 4 weeks, and they should monitor their effect. The probiotic should be taken at the dose recommended by the manufacturer.

1.2.1.7 Primary care clinicians should discourage the use of aloe vera in the treatment of IBS.

1.2.1.8 If diet is considered to be a major factor in a person's symptoms and general lifestyle/dietary advice has been followed, they should be referred to a dietitian for advice, including single food avoidance and exclusion diet, to ensure that the diet remains well-balanced.

1.2.2 Pharmacological therapy

1.2.2.1 Primary care clinicians should consider prescribing antispasmodic agents, to be taken as required, alongside dietary and lifestyle advice.

1.2.2.2 Laxatives should be considered for the treatment of constipation in people with IBS, but they should be actively discouraged from taking lactulose.

1.2.2.3 Loperamide should be considered as first-line treatment for diarrhoea in people with IBS[1].

1.2.2.4 Primary care clinicians should advise people with IBS how to adjust laxative or antimotility agent doses according to the clinical response. The dose should be titrated according to the stool consistency with the aim of achieving a soft, well formed stool (corresponding to Bristol Stool Chart type 4).

1.2.2.5 Primary care clinicians should consider the benefit of prescribing tricyclics as second-line treatment for people with IBS. Treatment should be initiated at a low starting dose (5–10 mg equivalent of amytriptyline), once at night, which should be reviewed regularly. The dose can subsequently be increased, but does not usually need to exceed 30 mg.

1.2.2.6 Primary care clinicians should consider prescribing selective serotonin reuptake inhibitors (SSRIs) only when tricyclics have been shown to be ineffective.

1.2.2.7 Primary care clinicians should consider reported side effects when prescribing tricyclics or SSRIs. Following prescribing of any of these drugs for the first time at low doses for the treatment of pain/discomfort, the person should be followed up after 4 weeks and then at 6–12 monthly intervals thereafter.

1.2.3 Behavioural therapies

1.2.3.1 Primary care clinicians should consider referring for behavioural therapies (cognitive behavioural therapy, hypnotherapy, psychological therapy) people with IBS who do not

[1] In certain situations the daily dose of loperamide required may exceed 16 mg, and the GDG notes that this is an out of licence dose.

respond to first-line therapies after 12 months and who develop a continuing symptom profile (described as refractory IBS).

1.2.4 Complementary and alternative medicine (CAM) therapies

1.2.4.1 Primary care clinicians should not encourage the use of acupuncture in the treatment of IBS.

1.2.4.2 Primary care clinicians should not encourage the use of reflexology in the treatment of IBS.

1.2.5 Follow-up

1.2.5.1 Follow-up should be mutually agreed between primary care clinicians and people with IBS based on symptom response to interventions. This should form part of the annual patient review.

4 AIMS OF THE GUIDELINE

The aims of the guideline are:

- To evaluate and summarise the clinical and cost evidence relating to all aspects of the diagnosis and treatment of Irritable Bowel Syndrome (IBS).
- To highlight gaps in the research evidence.
- To formulate evidence-based cost effective clinical practice recommendations relating to the diagnosis and treatment of IBS.
- To formulate consensus recommendations shaped around available evidence and expert GDG opinion in those areas of diagnosis and treatment of IBS where there is no clear clinical and cost effective evidence base.

4.1 Who the guideline is for

The guideline is of relevance to all people with IBS, carers for those people with IBS, primary healthcare professionals and social care staff that are involved in the care and/or support of those people diagnosed with IBS.

4.2 Groups covered by the guideline

Adults (18 years and older) who present to primary care with symptoms suggestive of IBS are covered by the guideline.

4.3 Groups not covered by the guideline

The following groups are not covered by the guideline:

a) Patients with other gastrointestinal disorders such as non-ulcer dyspepsia or coeliac disease will not be covered, except when a co-morbidity has specific relevance to the management of IBS.
b) Children and young people under 18 years of age.

4.4 Healthcare setting

It is recognised that the NHS is rapidly developing patterns of service delivery, with primary and secondary care borders blurring. The guideline will cover the care that is provided by primary healthcare professionals and it will indicate where secondary care referral is appropriate. The guideline is sensitive to the variations in commissioning of services relating to the diagnosis and treatment of IBS. The guideline recognises that there is current variation to service availability in both primary and secondary care across England and Wales, and at times will not state where care is accessed.

4.5 Diagnosis and management interventions covered by the guideline

The following diagnostic and treatment interventions will be covered. They have been classified into logical coherent areas of the guideline, supported by clinical and cost effectiveness reviews, and are consistent with the patient algorithm which typically reflects the patient pathway.

Diagnosis

Positive Diagnosis utilises criterion based reference tools. Negative diagnosis uses exclusion diagnosis through negative test results. This is typically characterised by primary care clinicians requesting a raft of investigations to rule out other co-morbidities. Diagnosis also addresses the identification of red flags that may lead to an alternative diagnosis such as bowel cancer. This guideline is cross referenced to NICE clinical guideline 27 (Suspected Cancer Referral).

Lifestyle: diet and exercise

This section of the guideline reviews clinical and cost effectiveness evidence relating to patient lifestyle. It is focussed on shared care decision making between the primary care clinician and the person with IBS. This develops coping behaviours and modifies lifestyle relating to dietary input/changes and levels of exercise that work towards alleviating symptom based IBS profiles.

Drug therapy

This section of the guideline reviews clinical and cost effectiveness evidence relating to different pharmacological treatments options that are prescribed to alleviate symptom based IBS profiles.

Referral and follow-up

This section provides consensus based recommendations and narrative on the importance of referral and follow up once diagnosis has been made. This also incorporates clinical and cost

effective reviews and recommendations on referral for people with intractable IBS, defined as a continuing symptom profile and lack of response to first line treatment interventions.

4.6 Interventions not covered by the guideline

If during the process of diagnosis for IBS another disease is suspected, further diagnosis and treatment of this disease will not be covered. Management and diagnosis of co-morbidity will not be covered. New drugs in development are not covered as they are not licensed for use.

4.7 Guideline Development Group

The guideline recommendations were developed by a Guideline Development Group (GDG) convened by the NICE-funded National Collaborating Centre for Nursing and Supportive Care (NCC-NSC) with membership approved by NICE. Members included representatives from patient groups; nursing; general practice and gastroenterology medicine; pharmacy; dietetics; public health; technical team from the NCC-NSC.

The GDG met 13 times between May 2006 and July 2007. All members of the GDG were required to make formal declarations of interest at the outset, and these were updated at every subsequent meeting throughout the development process. This information is recorded in the meeting minutes and kept on file at the NCC-NSC.

5 METHODS USED TO DEVELOP THE GUIDELINE

5.1 Summary of development process

The methods used to develop this guideline are based on those outlined by Eccles and Mason (2001). The structure of the recommendations sections (sections 6 to 11) (i.e. recommendations; evidence statements, evidence narrative and guideline development group commentary) came from McIntosh et al. (2001).

The stages used in the development of this guideline were as follows:

- Guideline scope development following referral from the department of health
- NICE stakeholder review and feedback
- Multidisciplinary guideline development group convened with formal appointment of the clinical lead and chair of the group by competitive interview
- Establish key clinical questions
- Identify sources of evidence
- Retrieve potential evidence

- Evaluate potential evidence relating to clinical and cost effectiveness, quality of life, for eligibility, quality and relevance
- Extract relevant data from studies meeting methodological and clinical criteria
- Interpret each paper, taking into account the results (including, where reported, beneficial and adverse effects of the interventions, cost, comfort and acceptability to patients), the level of evidence, the quality of the studies, the size and precision of the effect, and the relevance and generalisability of the included studies to the scope of the guideline
- Analyse, where appropriate using statistical synthesis, the results reported in the studies
- Prepare evidence reviews and tables which summarize and grade the body of evidence
- Formulate conclusions about the body of available evidence based on the evidence reviews by taking into account the above factors
- Agree final recommendations
- Submit drafts (short version and full version) of guideline for feedback from NICE registered stakeholders
- Consider stakeholders comments (GDG)
- Submit final version of the guideline to NICE.

NCC-NSC technical team members searched bibliographic databases for evidence, examined and quality assessed the evidence. The technical team compose successive drafts of the recommendations and guideline documents (including the full version of guideline; the NICE version and the quick reference guide), based on the evidence reviews and GDG input and deliberations. The GDG having interpreted the evidence formulated the recommendations. The NICE patient and public involvement programme produced the information for the public version, using the NICE version of the guideline, in collaboration with the NCC-NSC. The general methods for the evidence reviews are reported in sections 5.2 and 5.3. This linear relationship, demonstrating the relationship between the clinical and cost effectiveness results, evidence statements and resulting recommendations, is reported for each review in sections 6 to 11.

The search strategies for the reviews are presented in Appendix B. The included studies for each review are reported in Appendix C; the methodological assessments of the included studies are in Appendix D and the studies excluded from each review are listed in Appendix E.

5.2 Clinical effectiveness review methods

This section describes the methods of systematic reviewing that are common to all clinical effectiveness reviews of intervention studies. At the start of the guideline development process, a general protocol was discussed with the GDG which resulted in the selection criteria and approaches to analysis described below. Further details specific to the reviews are given for each review.

Selection criteria

The following selection criteria were to be applied to studies to determine their suitability for inclusion in the reviews

Types of studies

For intervention studies, the randomised trial (RCT) is the primary trial design. Quasi randomised studies could also be included (e.g. allocation by alternation, date of birth, etc). Where there is insufficient evidence from RCTs or quasi RCTs, cohort studies could be considered.

Both parallel and crossover trial designs could be included in the guideline: in the former, patients are randomised to one of two (or more) interventions; in the latter, patients receive interventions in a randomised order, crossing over to the second (and third) interventions after a specified period ('washout period').

Crossover trials are common in chronic conditions: they have the advantage that the patient acts as their own control, so there are no differences in baseline patient characteristics for each intervention, unlike parallel trials in which different patient groups receive the interventions. The crossover design is only appropriate when the condition is truly chronic (i.e. no progression or regression) and when the interventions make no permanent or slow decaying changes to the patient's condition. Crossover trials have the disadvantage that effects of the second intervention may be influenced by those in the first period (carryover effects). To avoid errors of this type, better designed crossover trials have a washout period between interventions, in which the patient characteristics are allowed to return to the levels present before the first intervention. Ideally, the characteristics are measured at intervals following the first intervention period, and the second intervention is delayed until the baseline values are retrieved, but, especially in older studies, this is usually approximated by the trialists using a 'washout period' they believe to be appropriate.

Some studies do not have a washout period, and the GDG's view was that crossover trials without washout periods should not be included, unless first period data are available – although, even this should be treated with caution, unless individual patient data are reported. For each review, the GDG decided if crossover trials were allowable, and, if so, defined the washout period. Factors taken into consideration included the lifetime of the intervention (especially for drugs). The washout period for each review is given in the methods section for that review. Trials with washout periods shorter than the pre-determined value should be excluded. Studies that do not state a washout period should be assumed to have none, and therefore should be excluded.

Studies should be restricted to the English language, with the exception of studies translated for Cochrane reviews, but the date should not be restricted.

Types of participants

Participants should be adults (18 years and older). However, studies could be included if they had some participants slightly below 18 years, provided that the mean age indicated that the majority were adults.

Participants should have a diagnosis of IBS. Suitable definitions included Rome I, Rome II or Manning criteria. Studies could also be included if the authors stated the patients had IBS, or if they described patients who had a set of symptoms suggestive of IBS. Studies reporting patients with single symptoms such as chronic constipation/diarrhoea in isolation should not usually be included. Studies could be included if a proportion of the patients had IBS, provided the IBS subgroup was reported separately, but such studies should be treated with caution unless the IBS subgroup members were separately randomised to treatments.

All settings could be included, but those in secondary/tertiary care should be distinguished from those in primary care only. This decision was taken regardless of the date of the study (people who were outpatients 20 years ago would now be treated in primary care).

Indirect evidence may be considered for some reviews, where direct evidence is not available, or is insufficient (for example, the use of laxatives in the treatment of constipation in non-IBS patients). In all cases, indirect evidence should be used to provide additional information, and its quality should be downgraded accordingly. Indirect evidence should not be combined in a meta-analysis with direct evidence. The indirect evidence permitted is given in the methods section for each review.

Types of intervention

The interventions varied across reviews and are detailed at the beginning of each review.

Interventions could be given in three different ways:

- As short-term rescue medication (e.g. antimotility agents for acute diarrhoea episodes)
- As a longer-term maintenance treatment (e.g. antispasmodics)
- As a 'one-off' intervention or series of treatments at the start of the management period (e.g. psychotherapy).

For the longer-term, maintenance interventions, the GDG specified a minimum acceptable period for the intervention. This was set at four weeks, and the reason for this was partly to take into account women's menstrual cycles. Maintenance studies with intervention durations of less than four weeks should not be included.

Types of outcome measures

The GDG decided on a number of outcomes related to symptom control. These would either be measured as the number of patients with a particular feature (dichotomous outcomes) or as a mean measurement, preferably on a validated scale (continuous outcomes). The following outcomes were considered to be primary:

- Global improvement of symptoms
- Global symptom scores.

Other outcomes were also considered important:

- Abdominal pain
- Bloating
- Stool score/general improved bowel habit
- Quality of life, using a validated scale
- Adverse effects.

The time of measurement and duration of follow-up should be recorded, together with information on whether the studies reported a change in symptoms from baseline, final values following treatment, or a mean value based on diary records.

'Global' meant a measure that took into consideration a combination of the following IBS symptoms: pain, bloating and stool properties (e.g. frequency, consistency, ease of passage). Alternatively, the participants could have assessed their overall symptoms as improved/same/worse; provided this did not obviously refer to just one component of IBS, these measurements could also be included in the 'global' category. Studies in which the authors labelled their outcomes as 'global' but in fact only measured one component should be analysed as single components.

The GDG decided that different definitions of improvement should not be distinguished (e.g. 100%, 75% improvement, slight, much), and that categorical outcomes should be dichotomised, e.g. grouping together 'much improvement' with 'slight improvement'.

For the individual symptom components, studies could record the number of people with that symptom at the end of the study or during the study, or they could record changes in symptoms over time, or a final symptom score at a particular time. For a positive outcome, the number of people with fewer symptoms (e.g. less pain) or the number with no symptoms should be recorded. For a negative outcome, the number with more symptoms (e.g. increased bloating), and the number of people with that symptom should be used. These two types of outcomes (absolute and increase/decrease) could be recorded on the same forest plot, but should not be combined in a meta-analysis.

For continuous outcomes, we recorded the severity score of the symptom (negative outcome) or the improvement in the symptom score (positive outcome).

Stool scores can have various formats: sometimes the raw values are recorded (e.g. stool frequency or consistency) or the severity may be assessed on a visual analogue scale. In the former, this measurement is only meaningful when the results are given separately for the different types of IBS - whether this is a positive or negative outcome depends on what type of IBS the person had. Therefore, if a study has people with a range of types of IBS, this type of raw value measurement should be disregarded. The severity score may be included as an acceptable outcome measure, as may the patient's assessment of improved bowel habits.

We note that the majority of these outcome measures are subjective and therefore, have potential for bias.

SEARCH STRATEGY

The search strategies and the databases searched are presented in detail in Appendix B. All searches were carried out on the following core databases: Medline, Embase, Cinahl (all using the OVID interface) and The Cochrane Library. Additional databases were searched for individual reviews where appropriate.

For this guideline, a general set of terms was produced relating to IBS. The relevance of the terms diarrhoea and constipation was explored before they were included in the IBS filter. For each review, terms related to the intervention were combined with the set of IBS terms. Where appropriate, study design filters (RCT and systematic review) were applied. Results were limited to papers published in English where possible. All searches were updated to June 2007.

Hand-searching was not undertaken following NICE advice that exhaustive searching on every guideline review topic is not practical or efficient (Mason 2002). Reference lists of articles were checked for studies of potential relevance.

METHODS OF THE REVIEW

Sifting process

Once the search had been completed, the following sifting process took place:

- 1st sift: one reviewer sifted the title/abstract for articles that potentially met the eligibility criteria
- 2nd sift: full papers were ordered that appeared relevant and eligible or where relevance/eligibility was not clear from the abstract

- 3rd sift: full papers were appraised, generally by one reviewer using an inclusion criteria form, and this was checked where necessary by a second reviewer.

Quality assessment and validity

Once individual papers were retrieved, the articles were checked for methodological rigour (using quality checklists appropriate for each study design), applicability to the UK and clinical significance. Assessment of study quality concentrated on dimensions of internal validity and external validity. At this stage, some studies were excluded if the interventions were not licensed for use in the UK or they were not regularly used in the UK. Studies in which the interventions were obsolete were also excluded.

Studies for which the methodological quality indicated a high potential for bias were included in the review, but were not included in the analysis.

Data abstraction

Data from the included studies were extracted by one reviewer for each review, with random checking by a second reviewer, and entered into a Microsoft Access relational database that had been especially designed for the guideline. The use of the database provided a more structured extraction, for example, only certain choices could be made for some items, although free text fields were also used. The main advantage of using a database for this purpose is that a large amount of detail can be input, and then an overview obtained using database sorting procedures. The following data were extracted from each study:

- Review being addressed
- Study details: study design (RCT, quasi-randomised, cohort study, etc); parallel/crossover, washout period; country where trial conducted; setting; funding
- Study quality
- Participants: age (mean and range), gender (ratio male:female), co-morbidities, inclusion/exclusion criteria, IBS diagnosis method, type of IBS, presence of bloating, presence of pain, measure of severity of IBS, symptom status at trial entry, length of time since diagnosis, duration of symptoms, ethnicity, socio-economic group, weight, post-infective/non post-infective initiated IBS
- Interventions: class (e.g. insoluble fibre) and sub-class (e.g. wheat bran), total amount per day, frequency/time of consumption, means of delivery (oral capsule, taken as a food, drink, etc), duration of treatment; concurrent treatment in both arms
- Comparator: placebo (details of what it is), other control group, other intervention
- Outcome: including follow-up period, scales used, definition of success (if using “improved”, “complete response”, etc)
- Results for each outcome.

If studies were published more than once, data were extracted from the most recent report where there were differences; otherwise all papers were used for data extraction.

Masked assessment, whereby data extractors are blind to the details of journal, authors etc, was not undertaken.

Appraisal of methodological quality

The methodological quality of each trial was assessed by one reviewer and randomly checked by a second. The following quality items were assessed:

- *A priori* sample size calculation:
  - Whether or not this was carried out
- Method of generation of the randomisation sequence:
  - The means by which interventions are distributed amongst the participants
  - Whether the method was reported or unclear (i.e. no details given)
  - Whether the reported method was adequate, inadequate or partial (Table 1)
- Allocation concealment at randomisation:
  - The means of preventing the treatment assignment being known *before* the time of allocation
  - Whether the method was reported or unclear (no details)
  - Whether the reported method was adequate, inadequate or partial (Table 1)
- Baseline comparability of treatment groups:
  - For relevant risk factors
- Patients stated to be blinded, especially for comparisons with placebo:
  - Blinding involves hiding the nature of the intervention from participants, clinicians and treatment evaluators *after* allocation has taken place
  - Blinding may be not be possible depending on the nature of the interventions
  - Blinding may be more important for some outcomes than others (this is noted in the reviews)
- Outcome assessor stated to be blinded
- No loss to follow-up for each outcome:
  - Studies with at least 20% of data missing from any group were considered to be potentially biased
  - Those with moderate loss to follow up (20 to 50%) were considered in sensitivity analyses
  - Those with 50% or more patients missing from any one group were regarded as flawed and not analysed further
- Intention to treat analysis:

    - Trial participants should be analysed in the groups to which they were randomised regardless of which (or how much) treatment they actually received, and regardless of other protocol irregularities
    - All participants should be included regardless of whether their outcomes were actually collected
- For crossover trials, the washout period relative to the minimum for the review:
    - Studies in which the washout period was shorter than the minimum were not included, as were studies with no washout or none stated
    - Studies reporting first period only data as individual patient data were included
- The intervention time relative to a minimum of 4 weeks or as defined for the particular review:
    - Studies in which the intervention time was shorter than 4 weeks were usually excluded, but slightly shorter durations could be included in the absence of other data.

Table 1:

Adequate Sequence Generation

- Coin toss, throwing a dice, shuffling, drawing lots (from a container).
  Partial: drawing a card from a pack.
- Computer- or calculator- generated sequence (including minimisation and biased-coin/urn design). **Partial**: "random permuted blocks".
- Random number table or statistical tables. **Partial**: random numbers, randomisation table.
- Randomised Latin square design.

Inadequate Sequence Generation

- Randomised Latin square. For example, allocation by alternation, birthdate, day of week.

Adequate Allocation Concealment

- Central randomisation: with contacting details and/or statement that central office retained schedule; must apply to all patients. **Partial**: vague statement of central randomisation.
- Independent 3rd party: allocates interventions *and* retains schedule, or statement that *allocator* has no knowledge of patients. **Partial**: 3rd party, but unclear treatment allocation.
- 3rd party cluster randomisation: 3rd party has no knowledge of clusters.
  Partial: unclear what 3rd party knew.
- Different parties (including one of authors): should have no knowledge of the patients *and* retain the schedule.
- Secure computer assisted method, e.g. locked file. **Partial**: as adequate, but unclear access.
- Sequentially numbered, opaque, sealed envelopes - all required, else **partial**.
- Serially numbered, identical containers, allocated sequentially - all required, else **partial**.

Inadequate allocation concealment

- For example, schedule known in advance, birthdate, case record number.

Data synthesis

Meta-analysis of similar trials, where appropriate, was carried out using *The Cochrane Collaboration's* analysis software, Review Manager (Version 4.2). Trials were pooled using a fixed effects model and plotted on forest plots. Where there was significant heterogeneity, a random effects model was used as a sensitivity analysis.

For dichotomous studies, we used the analyses reported by the authors, which was usually those reporting an outcome. Where there were incomplete data reported (more than 20% missing in any one group), we carried out sensitivity analyses, excluding these studies.

Where it was possible to combine studies, outcomes were summarised for dichotomous data using odds ratios (as default), relative risks (where the event rate in either arm was greater than 20%), or Peto odds ratios (where there were studies with no events in one arm). Numbers needed to treat (with the control group rate to which they apply) were calculated from the risk difference, where appropriate. The number needed to treat (NNT) is the number of people who would have to be treated for one to have an improved outcome.

For continuous data, weighted mean differences were used and where the studies reported measurements on different scales, standardised mean differences were used. Studies reporting final values and studies reporting change scores were combined if the scales used were the same, otherwise they were reported separately. Summary statistics and their 95% confidence intervals (95% CI) were reported where sufficient detail allowed their calculation.

In some studies, the mean difference was given with a p-value for the difference; this allowed calculation of the standard error. Results from such studies could then be combined in a meta-analysis with other studies reporting means and standard deviations: the standard error and mean difference were calculated for each study and then the studies pooled using the fixed effects generic inverse variance method in RevMan to give a weighted mean difference and 95% confidence intervals. This procedure is only appropriate when the same scales are used or transformation between scales is possible.

Crossover and parallel studies were analysed separately because there were insufficient data to calculate correlation factors. Trials were analysed by the conventional approach of treating the two arms of the crossover as if they were from a parallel trial with separate groups. Alternatively, if first period data were available, these were used in the analysis and the parallel and first period (pseudo-parallel) trials combined.

Stratifications

We planned *a-priori* to separate studies by the type of IBS, into patients with constipation predominant, diarrhoea predominant and alternating types. Studies that did not say or that considered all types of IBS together were treated as a separate group. Other stratifications were planned depending on the review.

Subgroup analyses

Randomised trials generally report four different types of subgroup analysis:

- Between-trial, in which the *studies* are separated according to the particular variable considered (e.g. dose)
- Within-trial subgroup analyses, with stratification of the *participants* by the particular characteristic (e.g. post-infective or not) followed by randomisation

- *A-priori* defined within-trial subgroup analyses, in which the *participants* were not stratified, but later separated according to pre-specified characteristics – these analyses should be included cautiously, because the interventions are not randomised to the subgroups
- Post-hoc within-trial subgroup analyses, in which the *participants* were separated afterwards without pre-specification.

All subgroup analyses are non-randomised comparisons between the different subgroups, however, types 1 and 2 are more reliable. Type 3 analyses can be included in meta-analyses with caution, but post-hoc within trial subgroup analyses were considered to be data-driven and were included only under exceptional circumstances. Most commonly in the guideline, the term 'subgroup analysis' refers to between-study comparisons.

Subgroup analyses were carried out in order to investigate heterogeneity or to investigate pre-specified features. We assessed heterogeneity between trials by visual inspection of forest plots, noting where there was poor overlap of horizontal lines, and by using statistical measures: the χ^2 test for heterogeneity and the level of inconsistency, I^2 ($I^2 = [(\chi^2 - df)/ \chi^2] \times 100\%$, where df is the degrees of freedom). We considered that there was heterogeneity if the p-value (for heterogeneity) was less than 0.1 and I^2 was greater than 50%. Any heterogeneity was explored further and unexplained heterogeneous results were not used as the basis for recommendations.

The following pre-specified factors were proposed for subgroup analyses:

- Type of intervention (e.g. soluble fibre/insoluble/both)
- Dose (defined for the particular review)
- Duration of intervention
- Post-infective/Non-post-infective
- Symptom severity.

Subgroup analyses specific to each review were also carried out, as appropriate.

Sensitivity analyses

Sensitivity analyses were carried out to investigate assumptions within the analyses. These included the following:

- Methodological quality
- Setting.

For methodological quality, we paid particular attention to allocation concealment, loss to follow-up and blinding of patients. We did not include studies with more than 50% loss to follow-up for a particular outcome in the analyses. Otherwise we carried out sensitivity analyses on studies

that had between 20 and 50% withdrawals from any group (or protocol deviations that were eliminated from the study's analyses).

Sensitivity analyses were also carried out where there were quasi-randomised studies (e.g. sequence generation by alternate allocation or date of birth) or inadequate allocation concealment. If these represented the only evidence, their quality was downgraded accordingly.

Significance

Sometimes the results were statistically significant, but small in size. In this case, the GDG decided on what was a clinically important difference in the summary statistics for a particular outcome. Some meta-analyses gave pooled summary statistics close to the null value. Where the confidence interval was narrow, we considered this to be 'evidence for no significant difference' between interventions and the approach became similar to that of an equivalence trial (Alderson 2004). Where the confidence interval was wide, there was considered to be insufficient information to determine if there was a difference between interventions. For most outcomes, the GDG judged what constituted a wide confidence interval; if there was any doubt, they decided there was uncertainty.

General approach to reviewing

The clinical effectiveness reviews seek to determine answers to a number of questions, which were investigated using the following comparisons:

- Does the intervention work? (and is it harmful?):
  - Direct comparisons of intervention with placebo/none
- Is there a dose effect?
  - Direct dose comparisons
  - Subgroup analyses (across trials) of intervention versus placebo, by dose
- Is the duration of treatment important?
  - Direct comparisons of different durations
  - Subgroup analyses of intervention versus placebo, by duration
- Is the intervention better than another treatment?
  - Direct comparisons
  - Subgroup analyses of interventions versus placebo, by type of intervention
- Is the intervention useful as an adjunct to another treatment?
  - Direct comparisons (A + B versus B alone)
- Are there (pre-specified) subgroups of patients for whom the intervention is more effective?
  - E.g. type of IBS (constipation, diarrhoea, alternating); severity of IBS
  - Subgroup analyses: preferably within trials (stratification then randomisation for each subgroup) or across trials; less acceptably, within trials.

We note that the best type of information is from direct comparisons in which two values of the variable considered (e.g. dose 1 and dose 2) are randomised to different groups of people. However, some useful information can be obtained from between-study subgroup analyses.

Grading evidence

For some reviews, we used the GRADE[‡] scheme (GRADE working group 2004) to assess the quality of the evidence for each outcome using the approach described below, and evidence summaries across all outcomes were produced.

According to the GRADE scheme, evidence is classified as high, moderate, low or very low:

- High - further research is very unlikely to change our confidence in the estimate of effect
- Moderate - further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
- Low - further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
- Very low - any estimate of effect is very uncertain.

The procedure adopted when using GRADE was:

1. A quality rating was assigned, based on the study design – for example, RCTs started as high and observational studies as low.
2. This rating was up or downgraded according to specified criteria: study quality, consistency, directness, preciseness and reporting bias. These criteria are detailed below. Criteria were given a downgrade mark of -1 or -2 depending on the severity of the limitations.
3. The downgrade/upgrade marks were then summed and the quality rating revised. For example, a decrease of -2 points for an RCT would result in a rating of 'low'.
4. Wherever possible, reasoning was explained for the downgrade marks.

Study quality

Study quality is assessed against standard criteria, depending on the study design. For randomised trials, we took into account: the adequacy of allocation concealment; blinding of participants for comparisons and outcomes susceptible to bias; loss to follow-up and deviations from intention to treat. The GDG regarded blinding of participants to be important for the comparisons with placebo, but did not necessarily consider blinding of different active interventions to be critical. They did not consider blinding to be important for the behavioural therapies, mainly because this was not possible to achieve. The majority of outcomes in the IBS guideline are subjective and therefore susceptible to bias. A downgrade mark of -1 was given for inadequate allocation concealment and for a loss to follow-up of more than 20% in any one arm or overall. A loss to follow-up of 50% or more was given a downgrade of -2 (but was more usually excluded from the analysis). If the evidence was a meta-analysis of several studies, we

[‡] GRADE – Grading of Recommendations Assessment, Development and Evaluation

took into consideration the proportion and weighting of poor quality studies, and in some instances carried out sensitivity analyses disregarding these studies and giving a separate rating for the new meta-analysis.

Consistency

When several RCTs have widely differing estimates of treatment effect (heterogeneity or variability in results) the results are regarded as inconsistent. We defined this as a p-value for heterogeneity less than 0.1 and an I^2 value greater than 50%. Where this was the case, we gave a downgrade mark of -1. Where possible, we carried out pre-defined subgroup analyses to investigate heterogeneity and reported these results separately. Generally, we did not regard single trials (especially smaller ones) as having inconsistency unless there were *a-priori* defined subgroups showing widely different effects.

Directness

Directness refers to the extent to which the population, interventions, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is only relevant if there is a compelling reason to expect important differences in the size of the effect. For example, many interventions have more or less the same relative effects across patient groups, so extrapolation is possible and reasonable. There are various types of indirectness found in studies:

- When the guideline-defined drugs differ from those in the studies, but are within the same class. Similar issues arise for other types of interventions, for example, different types of psychotherapy.
- When there are no direct comparisons of interventions, investigators must make comparisons across studies. For example, we want to know the difference in effectiveness between interventions A and B, but we only have information on A versus placebo and B versus placebo.
- Specifically for IBS, the GDG decided that a difference in setting – secondary care in the studies rather than primary care in the guideline – was a relevant indirectness factor. Their reasoning was supported by differences found in surveys of IBS in primary and secondary care (Miller 2006).

Preciseness

This is a rather subjective, but nevertheless important category. Evidence is considered to be imprecise if:

- The sample size is small. This is a subjective measure and is more important in a single study. If there was a power calculation for that outcome and comparison, it was used to decide if a study was ‘small’. Otherwise we used the rule of thumb that if the study had less than 25 patients in any one arm, this was too small. The rationale for this was that below this size, assumptions about normal distributions become much less valid. However, if these

small studies were combined in a meta-analysis, we regarded their use as much more acceptable.

- There are sparse data (only a few events and they are uninformative).
- If confidence intervals are sufficiently wide that the effect estimate is consistent with both important harms and important benefits, and would lead to conflicting recommendations. This category requires the GDG to decide what are important harms and benefits for that outcome measure. Where the confidence intervals were very wide, we gave a downgrade mark of -2.

Reporting bias

Reporting bias occurs in two main ways:

- Publication bias, in which papers are more likely to be published if their results are statistically significant. The existence of publication bias in the studies in a meta-analysis can be investigated in a limited way using funnel plots, in which the standard error is plotted against the log odds ratio, the log relative risk or the mean difference. Asymmetry is indicative of reporting bias. This method is usually only useful when there are at least five studies. Industry sponsored studies are also regarded as potentially biased.
- Outcome bias, in which authors do not report some outcomes (probably because they have non-significant results), even though they say in the methods section that they have measured them.

We note that the GRADE approach, although rigorous, still requires judgements to be made, for example, what is a 'wide' confidence interval; what is a 'small' study; how important is blinding of patients for a particular outcome; how serious is it that the study population is treated in secondary care rather than primary? We have indicated how we considered these difficulties in the bullet points above, and the GDG made judgements as appropriate.

Evidence Statements

The GRADE summary (where used) was condensed into evidence statements, which are based on the quantity and quality of the evidence as shown in Table 2. Sometimes the evidence statements summarised more than one outcome measure. Where there were no GRADE summaries, evidence statements were made based on the analyses.

Table 2: Evidence statements

Description	Quality	Quantity
Strong evidence	Good quality	Large amount of data / meta analysis
Good evidence	Good quality	
Moderate evidence		Reasonable amount
Fair evidence	Acceptable quality	
Limited evidence		Not much evidence: trial < 50 people
Weak evidence	Poor quality	
Insufficient evidence		Not enough evidence to judge: trial size < 20 people or wide confidence interval

Generally, for randomised trials, a GRADE rating of 'good' equated with the wording 'good' or 'strong' evidence; a rating of 'moderate' with 'fair' evidence; a rating of 'low' was given the wording 'weak' evidence and a rating of 'very low' was described as 'insufficient' evidence.

5.3 Cost effectiveness review methods

Health economic evidence is useful in guideline development as it assesses the costs and benefits of alternative courses of action which could be recommended within the guideline. Cost-effectiveness evidence can be used to determine whether a particular recommendation would result in the efficient use of NHS resources by considering whether it achieves additional health gain at an acceptable level of cost. Whilst cost-effectiveness is an important consideration for all recommendations made within the guideline, two areas were identified as being priority areas for which cost-effectiveness evidence would have particular importance for informing recommendations. These were identified by the health economist in conjunction with the GDG

after consideration of the importance of each clinical question in terms of the number of patients likely to be affected and the impact on costs and health outcomes for those patients.

The use of tests to exclude alternative diagnoses in people with IBS-like symptoms was considered to be a high priority area for economic evaluation for the following reasons: diagnostic testing has the potential to result in earlier diagnosis of organic disease which may improve health outcomes; the widespread use of tests may have significant cost implications; the use of tests may result in unnecessary anxiety for patients, particularly if the rate of false positive results is high; invasive tests may have adverse consequences for patients in terms of complications.

The use of pharmacological and behavioural interventions in the management of IBS was also identified as a high priority area for economic evaluation. Pharmacological interventions were identified as an area of high priority because the ongoing use of these interventions in a large number of IBS patients would have significant implications for the use of NHS resources. Behavioural interventions were identified as an area of high priority because these are not widely used at present in the management of IBS and therefore significant additional resources may be required if these are recommended for widespread use.

Two approaches were employed to provide cost-effectiveness evidence for the GDG to consider when making recommendations. Firstly, a review of the health economic literature was carried out and relevant health economic evidence was presented to the GDG. Secondly, further economic analysis was carried out in the priority areas where there was insufficient evidence available from the published literature to inform recommendations and where there was sufficient evidence to demonstrate the clinical effectiveness for the intervention or diagnostic strategy. This further economic analysis was conducted in the form of a cost-effectiveness analysis where the additional benefits were measured in terms of quality-adjusted life-years (QALYs) and the additional costs were assessed from an NHS and personal social services perspective. The GDG considered the incremental cost per QALY for alternative management and diagnostic strategies alongside the clinical effectiveness evidence when formulating recommendations. Where one clinical strategy was clearly more effective and less costly than another it was considered cost-effective. Where one strategy was more effective but also more costly, the incremental cost per QALY was estimated and this was compared to a cost-effectiveness threshold of £20,000 to £30,000 per QALY in line with the principals laid out in the NICE Guidelines Manual (NICE 2007). For those clinical questions not prioritised for economic analysis, the GDG considered the likely cost-effectiveness of associated recommendations by making a qualitative judgement on the likely balance of costs, health benefits and any potential harms.

5.3.1 Economic literature review methods

Background

The diagnostic review described in chapter 6 provides evidence on several criterion based reference tools that are useful in the diagnosis of IBS in patients who do not have "red-flag" symptoms. However, some patients meeting the diagnostic criteria for IBS, following the application of a criterion based reference tool, may have another disease which has similar symptoms to IBS, such as inflammatory bowel disease (Crohn's disease and ulcerative colitis), coeliac disease or lactose intolerance. In some patients these conditions may be mistakenly diagnosed as IBS and sometimes they may be present alongside IBS. The health economic review aimed to assess whether further diagnostic testing to identify patients with alternative diagnoses is cost-effective in patients meeting the diagnostic criteria for IBS who do not have any "red-flag" symptoms.

The clinical effectiveness reviews presented in Chapters 7 to 10 assess the effectiveness of various interventions which may be useful in the management of IBS. The economic review aimed to assesses the cost-effectiveness of these interventions to manage IBS based on the published literature. Whilst pharmacological interventions and behavioural interventions were identified by the GDG as being priority areas for which cost-effectiveness evidence would have particular importance for informing recommendations, this review was not restricted to these interventions and evidence was included on any of the management interventions covered by this guideline.

OBJECTIVES

- To determine the cost-effectiveness of tests to identify alternative diagnoses in patients meeting the diagnostic criteria for IBS who do not have any "red-flag" symptoms.
- To assess the cost-effectiveness of interventions used in the management of IBS.

SELECTION CRITERIA

Types of studies

The types of studies included in the review were trial or model based economic evaluations including cost-effectiveness analyses, cost-utility analyses and cost-benefit analyses. Cost-minimisation studies were excluded except where therapeutic equivalence had been demonstrated.

Population

The population considered was patients meeting the diagnostic criteria for IBS who do not have any "red-flag" symptoms.

Types of intervention

The following interventions were considered: diagnostic tests for inflammatory bowel disease; coeliac disease; lactose intolerance; all interventions used in the management of IBS.

Outcomes

The outcomes assessed by the review were: cost per QALY; cost per LY; cost per correct diagnosis; cost per unit of clinical effect; cost-benefit ratio; net benefit.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the MEDLINE database for objective 1 using the strategy given in appendix B. Specific searches were also performed on the NHS EED database using the MeSH terms for inflammatory bowel disease (exploded to include Crohn's disease and ulcerative colitis), lactose intolerance and coeliac disease. Free-text searching on the NHS EED database was explored but did not yield any further relevant papers.

Searches were performed on the MEDLINE database for objective 2 using the strategy in Appendix B. Specific searches were also performed on the NHS EED database using the MeSH term for irritable bowel syndrome which yielded two further papers. Free-text searching on the NHS EED database was explored but did not yield any further relevant papers.

Included papers

The search results for both objectives were sifted together to allow identification of any cross-relevant information. Twenty-five papers were retrieved in full, of which 10 addressed the cost-effectiveness of management strategies (objective 2), with 4 included in the review, and 15 addressed the cost-effectiveness of tests to identify alternative diagnoses (objective 1), with 4 included in the review. Excluded papers and the reasons for exclusion are detailed in Appendix E. The most common reasons for exclusion were that the paper was not an economic evaluation or that it considered an inappropriate population. Included studies were reviewed by the health economist and the quality of each study was critically appraised using a validated check-list for economic analyses (Drummond 1997). Each study is discussed under the clinical question it addresses within chapters 6 to 10 of the guideline. The characteristics of the included studies are given in Appendix C and the details of the quality assessment are provided in of Appendix D.

5.3.2 Cost-effectiveness modelling methods

Having considered the published clinical and cost-effectiveness evidence on the use of diagnostic tests in people with IBS, the GDG decided that further economic analysis was needed to determine the cost-effectiveness of serological tests for coeliac disease in people meeting the IBS diagnostic criteria compared to initiating IBS management without testing for

coeliac disease. This was done by adapting one of the published economic analyses to make it more applicable to the NHS in England and Wales. Further details on the cost-effectiveness analysis carried out for this area of the guideline is provided in Chapter 6.

There was insufficient cost-effectiveness data identified from the published literature to allow the GDG to determine whether each of the various management interventions were cost-effective. An economic analysis was carried out to estimate the cost-effectiveness of pharmacological interventions and behavioural interventions in the management of IBS as these had been identified by the GDG as areas where cost-effectiveness evidence would be particularly important in informing recommendations. The remainder of this chapter describes the methods used in this economic analysis. The results are presented in the relevant chapter subsection for each pharmacological intervention and behavioural therapy.

The general approach

- Two models were developed to estimate the cost-effectiveness for different types of IBS management interventions:
  - A long-term maintenance therapy model for pharmacological interventions which are taken on a regular basis such as laxatives, anti-motility agents, antispasmodics, tricyclics and SSRIs.
  - A "one-off" intervention model for behavioural interventions (CBT, psychotherapy and hypnotherapy) which are given over a defined period with the expectation that benefit continues beyond the intervention period.
- Modelling was carried out using the best available evidence
- Assumptions made in the model have been described explicitly. The validity of these assumptions was discussed with the GDG during the development of the model and the interpretation of the cost-effectiveness results
- The importance of model assumptions was examined through univariate sensitivity analysis
- Parameter uncertainty was explored by carrying out a probabilistic sensitivity analysis (PSA)
- Limitations of the analysis are explicitly discussed alongside the cost-effectiveness results.

Identifying evidence on prognosis, resource use and quality of life

A rapid literature review was carried out to identify data which could be used to inform the health economic modelling. This review had three objectives:

- To identify cohort studies providing prognostic data which could be used to inform the health economic model by determining health states which could be used to describe the natural history of IBS
- To identify quality of life data measured in people with IBS and determine what factors influence quality of life in IBS and how estimates of quality of life could be incorporated to reflect the natural history of IBS or the impact of interventions on quality of life in the economic model

- To identify estimates of health care resource use and costs for people with IBS and determine what factors influence resource use in IBS and how estimates of resource use could be incorporated to reflect the natural history of IBS or the impact of interventions on resource use in the economic model.

The methods and results of this review are described in Appendix F. Where the data from this review has been used to inform the economic model it has been discussed in the relevant methods section below.

Key assumptions

- The model used estimates of clinical effectiveness that were obtained from the systematic reviews of RCTs. These clinical effectiveness reviews combined the results from studies across the whole class (e.g. all antispasmodics), but also examined subgroups of that class (e.g. antimuscarics and direct-action smooth muscle relaxants). The model used a combined estimate of clinical effectiveness across the whole class unless there was evidence to demonstrate a significant difference in effectiveness between sub-groups or between interventions (e.g. individual drugs).

- Clinical effectiveness was estimated in the model by considering the proportion of patients who experienced a global improvement of symptoms. This was the primary outcome of the clinical effectiveness review and was also considered by the GDG to be closely related to an improvement in quality of life across the many different interventions considered by the economic model. Where evidence on global improvement of symptoms was unavailable, a symptom specific response rate was used after discussion with the GDG as to which of the available outcomes was most relevant. The efficacy data used for each individual class of interventions is discussed within the relevant chapter sub-section for that intervention.

- Cost-effectiveness was estimated for each IBS subtype (e.g. IBS-D/C/A) for which there was evidence of clinical effectiveness or for the population as a whole if trials did not show a difference in effectiveness between subgroups or did not provide effectiveness evidence by subgroup. The GDG considered whether the estimated cost-effectiveness was likely to apply equally to all IBS subtypes when formulating recommendations.

- Interventions which did not have sufficient evidence to demonstrate clinical effectiveness were excluded from the cost-effectiveness analysis. This judgement was made by the GDG after considering the clinical effectiveness evidence for each intervention.

- The model for long-term maintenance therapies estimated the cost-effectiveness of initiating therapy with interventions from within a particular class using a defined patient pathway. This management strategy was compared to a “no treatment” alternative in which patients

were not given any specific intervention and were not advised to return for follow-up. The "no treatment" alternative provided a common baseline, against which the costs and benefits of interventions from different classes could be assessed.

- The model for "one-off" interventions considered the addition of behavioural therapies to usual care compared to usual care alone in patients with refractory IBS. The population and comparator were selected to reflect the available RCT evidence on the clinical effectiveness of behavioural therapies. The RCTs for these behavioural interventions were considered by the GDG to be representative of patients with refractory IBS. In the majority of these trials ongoing IBS drug therapy was continued in both arms of the trial. The GDG interpreted these RCTs as reflecting the clinical effectiveness of adding behavioural therapy to usual care rather than replacing usual care with behavioural therapy.

- The cost-effectiveness of initiating therapy with *either* interventions from class A *or* interventions from class B was assessed where these represented mutually exclusive alternatives. Direct evidence comparing interventions from different classes was used where available. Otherwise, an indirect comparison was made using "no treatment" as the common comparator. However, these indirect comparisons should be treated with caution as they were not based on randomised comparisons and may be subject to bias.

- The majority of the pharmacological interventions are used to treat a specific aspect of the individual's symptom profile and can therefore be used in combination if more than one symptom is problematic. In this case they are not mutually exclusive alternatives and the incremental cost-effectiveness of one compared to the other has not been estimated.

Cost-effectiveness of intermittent use of maintenance treatments

- The intermittent use of maintenance treatments was considered by scaling drug costs and benefits by the proportion of days on which the treatment is used.
- If two interventions are used intermittently but not concurrently, for example laxatives and anti-motility agents in patients with IBS-A, the costs and benefits of each intervention were scaled according to the proportion of days on which they were used and the total costs and benefits have been summed across both interventions. The assumption here was that the benefit gained from treating a particular IBS symptom which was present on some days was independent of the benefit gained from treating another IBS symptom which was present on other days.

Cost-effectiveness of combined use of maintenance treatments

- The combined use of two interventions from different classes concurrently was not explicitly modelled as there was no direct evidence on the use of combined versus single interventions in the management of IBS. The cost-effectiveness of using maintenance

treatments in combination was considered qualitatively by the GDG based on the cost-effectiveness evidence for individual treatments and the likely additive effects of the interventions on costs and benefits.

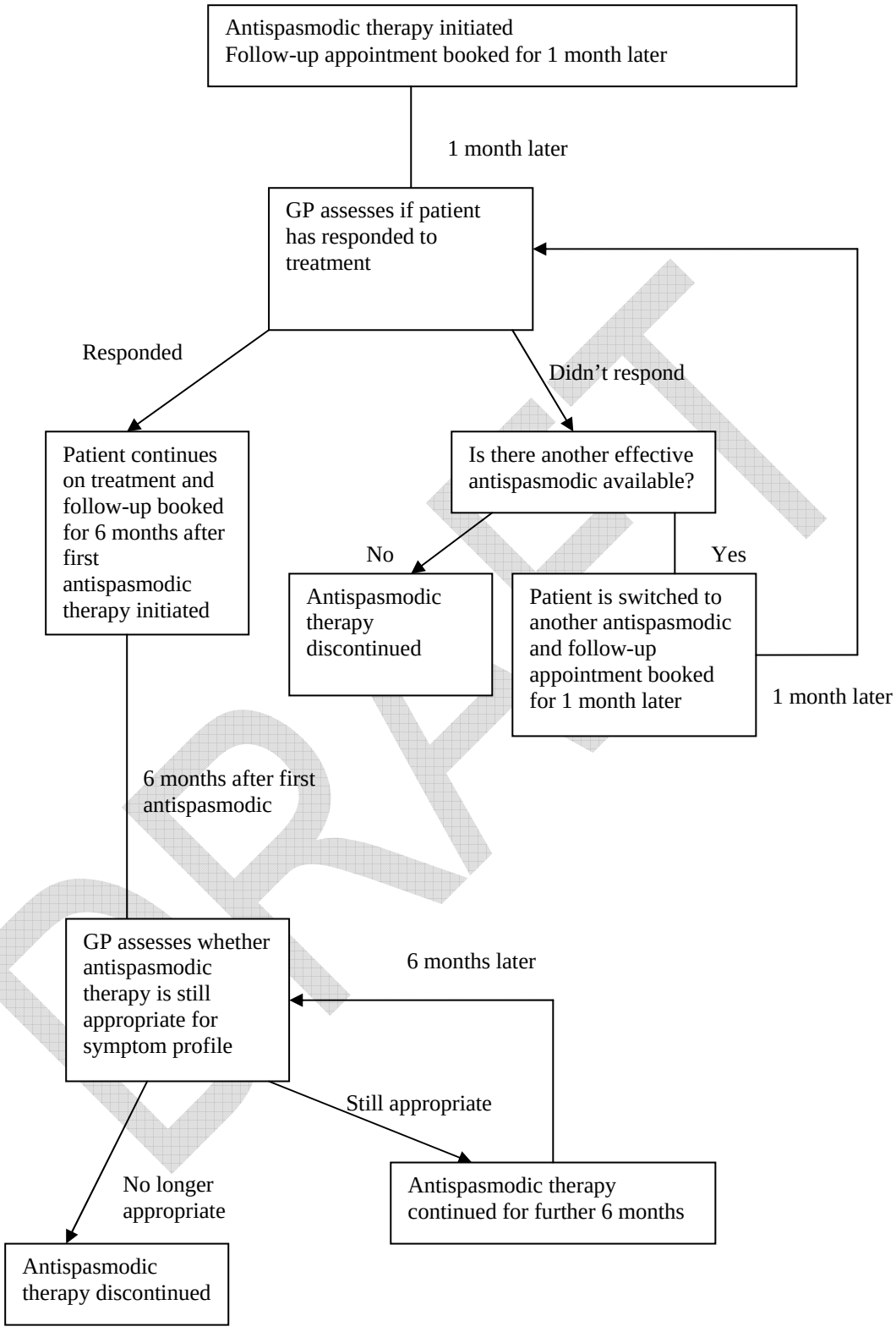
Determining the clinical pathway for maintenance interventions

In order to estimate the cost-effectiveness of maintenance interventions it was necessary to quantify the costs associated with prescribing and monitoring interventions and an appropriate time-frame for the analysis in terms of the duration over which costs and benefits were expected to differ as a result of a decision by a health care professional to initiate a particular a intervention.

There was evidence from the prognostic data reviewed in Appendix G that a patient's predominant symptom may change over medium term intervals (1-3 months) resulting in them switching between IBS subtype classifications. Evidence from Drossman (2005) showed that only 24.2% of patients remained in their baseline subtype over the study duration of 15 months. This suggests that any long-term maintenance therapy should be regularly reviewed to assess its continued relevance to the patient's evolving symptom profile. On the basis of this evidence the model was developed to consider periods of 6 months. In the first 6 months we estimated the cost-effectiveness of initiating a long-term maintenance therapy. We then estimated the cost-effectiveness of continuing the intervention for another 6 months in individuals who continue to experience a therapeutic benefit from the intervention.

The clinical pathway modelled is described in detail below and summarised in Figure 1 using antispasmodic therapy as an example. A slightly modified patient pathway has been used for tricyclics and SSRIs as these interventions require more frequent follow-up. This is described in detail in the tricyclics and SSRI section of Chapter 8.

Figure 1. Patient pathway for maintenance therapies illustrated for antispasmodics



Clinical pathway for maintenance model (See Figure 1 above)

- Patients initially receive the lowest cost intervention from within a class if there is no difference in effectiveness within the class (if there is a difference, each of the alternative interventions has been considered to estimate which is the most cost-effective to use first).
- Patients who demonstrate a successful response after 1 month continue on therapy until 6 months after treatment was initiated.
- Patients who do not respond switch to the next lowest cost therapy and response is assessed again after 1 month.
- The number of switches is limited by the number of effective interventions available.
- All patients receiving pharmacological maintenance interventions are reviewed after 6 months to assess whether the class of intervention is still relevant to the symptom profile.
- The above treatment pathway was compared to a "no treatment" alternative in which patients are not given any specific intervention and are not advised to return for follow-up.
- An analysis was undertaken to assess the maximum number of switches that are cost-effective by considering the additional cost and benefit of each additional switch of therapy.
- Probability of response to each subsequent intervention within a class was assumed to be independent of the response to previous interventions. A sensitivity analysis using lower response rates of 50% and 0% was carried out to test the impact of this assumption on cost-effectiveness.
- It was assumed that there is no fall off in treatment effect during the six month period for patients who have responded during the first month. This is an approximation, as some patients may experience a reduction in efficacy over time and may withdraw from treatment but the impact of this on cost-effectiveness is likely to be small given that treatment is reviewed every 6 months and patients are likely to discontinue therapy if it is no longer effective.
- It was assumed that the treatment effects do not persist after an intervention has been discontinued. This means that patients who stop therapy are assumed to return to their previous health state and patients who switch therapy do not experience the combined effects of both therapies in the cross-over period.

Clinical pathway for one-off interventions

- One-off interventions are given over a defined period with the expectation that benefit continues beyond that period.
- Follow-up data from trials were used to estimate the rate of fall-off in effectiveness and the time until no further benefit is expected. This determined the duration of the cost-effectiveness analysis.
- The number of patients responding over the duration of intervention and follow-up was fitted to the data available from the RCTs. Between the time points for which data is available we have assumed that the rate of change in effect is constant.

- Where the evidence was equivocal, such that alternative assumptions on the rate of fall-off in effectiveness could be justified, these alternative assumptions were considered in sensitivity analysis to assess how they alter the cost-effectiveness.
- Where the duration of continued effectiveness is over 1 year, discounting at 3.5% was applied to estimate the net present value of future costs and benefits.

Estimating the benefits associated with response to treatment

- In order to estimate cost-effectiveness it was necessary to estimate the benefits associated with treatment. In general these may be a gain in duration or quality of life, or a reduction in NHS resource use (such as fewer GP consultations).
- There was evidence from the literature review detailed in Appendix G to show that HRQoL is lower in patients with IBS than in matched controls (Akehurst 2002) and that HRQoL varies significantly by symptom frequency and severity but not by IBS subtype (El-Serag 2002). Akehurst (2002) found that resource use was significantly higher in patients with IBS than matched controls, but the evidence on resource use by symptom frequency, severity or IBS subtype was inconsistent (see Appendix G). We assumed in the model that patients responding to treatment experience a gain in health related quality of life but no reduction in resource use unless there was direct evidence from RCTs to demonstrate reduced resource use. We did not consider survival gains as IBS management interventions are not expected to affect survival.
- Utility is a measure of health related quality of life where a score of 1 represents full health and a score of 0 is a health state equivalent to death. Using the data presented in Mearin (2004) we estimated health state utility scores for high and low severity symptoms by aggregating scores across the IBS subtypes for patients with high frequency symptoms (present >50% of the time). This gave an estimated mean health state utility of 0.704 for patients with high severity symptoms and 0.775 for patients with low severity symptoms. We assumed that the utility gain associated with response to treatment was equivalent to an improvement in symptom severity (high to low severity). This was equivalent to an additional 0.071 QALYs per year of continued response (Mearin 2005). For comparison, an additional 0.135 QALYs would represent a complete resolution of IBS symptoms (Akehurst 2002).
- A threshold analysis was carried out to estimate the minimum treatment associated QALY gain for which treatment is still cost-effective.
- Adverse effects were not explicitly included in the model. Many of the adverse outcomes of interest considered in the adverse effects review (see section 8.5) were very similar to the symptoms of the IBS itself and were also considered within the effectiveness outcomes. It is likely that these adverse effects would have been captured by the clinical effectiveness estimate as this was based on global symptom score improvement. Therefore, patients who experienced a worsening of their IBS symptoms as a result of a specific intervention would be considered to have not responded to that intervention in the model and would discontinue that treatment. No other adverse effects were identified by the GDG as having

the potential to significantly impact on costs and quality of life for the interventions considered by the economic model.

Estimating the costs of the patient pathway

- Costs were considered from an NHS and PSS (Personal Social Services) perspective and included: drug costs for prescribed medications, consultation costs for the behavioural therapies and consultation costs for initiating and monitoring pharmacological interventions.
- Drug costs were based on the doses used in clinical trials and it was assumed that the lowest cost preparation would be prescribed regardless of whether this is proprietary or generic. Drug costs were based on the published costs given in the British National Formulary (Joint Formulary Committee 2007).
- Sensitivity analysis was carried out to consider whether the cost-effectiveness would be significantly different if the most costly preparation were to be used.
- Sensitivity analysis was carried out on alternative doses to those used in the trials where the GDG advised that these alternative doses were likely to be equally efficacious and more relevant to clinical practice.
- The cost of non-pharmacological interventions was estimated using the duration of clinical contact time required to deliver the intervention and the reference costs (Netten 2006) for face-to-face time with the relevant healthcare professional.

Estimating the probability of an improvement in global symptoms

- The probability of response was taken from the clinical effectiveness review using the probability of an improvement in global symptoms, unless this was unavailable. In that case an alternative symptom specific response rate was used after discussion with the GDG as to which of the available outcomes was most relevant. The efficacy data used for each individual class of interventions is discussed within the relevant chapter sub-section for that intervention.
- In the management intervention model, the cost-effectiveness was dependent on (i) the number of additional patients who respond in the treatment arm compared to the control arm, and (ii) the number failing to respond to treatment as these patients incur one month of treatment cost without benefit. In the one-off intervention model, the cost-effectiveness was also dependent on the probability of response in the comparator arm as this determines the absolute difference in response rates and therefore the clinical benefit.
- There was evidence from cohort studies that some patients experience an initial improvement in symptoms without any specific intervention. This may be a non-specific treatment effect following diagnosis and reassurance or it may be that symptoms fluctuate naturally and patients consult when their symptoms are particularly bad but symptoms then improve without any intervention. There was also evidence from randomised controlled trials that some patients in the placebo arms of controlled trials experienced an improvement in symptoms.

- Therefore we assumed a non-zero response rate in the no treatment arm of the model.
- The probability of moving from a high to low symptom severity state estimated from the Mearin (2004) cohort study (45%) was used to estimate the response rate in the no treatment arm in the base case analysis, except where the population was deemed to be refractory.
- The RCTs for behavioural interventions (CBT, psychotherapy and hypnotherapy) were considered to be representative of patients with refractory IBS. In the majority of these trials ongoing IBS drug therapy was continued in both arms of the trial. The mean response rate from the comparator arms of these trials (25%) was used to estimate the proportion of patients with refractory IBS that experienced an improvement in global symptoms under usual care which included the continuation of any ongoing drug therapy.
- A sensitivity analysis was carried out using the average response rate in the placebo arm of the RCTs. The response rate in the comparator arm of the RCTs varied from 0% to 71% over the studies used to estimate efficacy for the economic model with a mean value of 47.5%. The studies from the laxative review could not be used to estimate the placebo arm response rate as a different outcome was used to determine response for this intervention. However, the response rate using the alternative outcome was similar to that found in the other studies for the standard outcome.
- For refractory patients, the mean response rate from the control arms of the CBT trials (9%) was used in a sensitivity analysis to examine the impact of assuming a lower response rate in refractory patients continuing usual care.
- A sensitivity analysis was carried out assuming zero response in the no treatment arm but maintaining the absolute difference in response between treatment and no treatment from the basecase analysis.

Probabilistic sensitivity analysis (PSA) is used to provide an estimate of the uncertainty in the cost per QALY estimate due to uncertainty in the model parameters used to estimate the cost-effectiveness. The most obvious example of parameter uncertainty in the model was the confidence intervals surrounding the clinical effectiveness estimates, but other parameters used in the model which were based on empirical measurement also had some uncertainty associated with them. We carried out a PSA which considered the parameter uncertainty around the clinical effectiveness estimates, the response rate in the comparator arm, the utility gain associated with a response to treatment and the costs of behavioural therapies due to variation in the number and duration of sessions used in the RCTs. Where direct evidence from the RCTs on resource use reduction was applied in the model, the parameter uncertainty around this was also estimated in the PSA. The reference costs for pharmaceutical interventions and clinical contact time with health care professionals were assumed to be fixed in the model, as was the discounting rate which was fixed by the NICE “reference-case” for economic evaluations (NICE 2007). In the PSA we characterised the parameter uncertainty by using a probability distribution to describe each of the parameters, details of which can be found in Appendix H. We then

sampled from each distribution independently under the assumption that there was no correlation between the different input parameters. However, the same random number set was used to sample common parameters across the different cost-effectiveness comparisons to prevent sample bias being introduced when comparing the incremental cost-effectiveness of two interventions. We then calculated the model outcomes (incremental costs, incremental QALY gains) for each set of sampled parameters and used these to estimate the uncertainty surrounding the cost per QALY estimate.

We based our PSA on 1000 samples of the parameter distributions. The results are presented as cost-effectiveness acceptability curves which show the proportion of samples that resulted in a cost per QALY value below various thresholds. It should be noted that the PSA did not account for uncertainty around the model assumptions and these were explored separately using univariate sensitivity analysis. Table 1 gives the basecase parameters that were used in estimating the cost-effectiveness of all of the pharmacological and behavioural interventions. Parameters that were specific to each intervention, such as efficacy estimates and intervention costs, are tabled in the relevant section of Chapters 8 and 9.

Table 1: Base case parameters applied in the economic model for all interventions

Description	Mean (95%CI)	Evidence
Utility gain associated with a response to treatment	0.071 (0.02 -0.147)	Mearin (2004), difference between high and low intensity symptoms
Response rate for no treatment arm	45% (33% - 57%)	Mearin (2004), 1 month probability of transition from high to low intensity symptoms
Response rate for usual care in people with refractory IBS	25% (19% - 32%)	Comparator arms of RCTs in behavioural therapies*
Discounting rate for costs and benefits	3.5%	NICE (2007), NICE reference case value
Cost for GP appointment to initiate intervention / review medication	£18	Netten (2006), GP cost per surgery consultation (excluding qualification and direct care staff costs)

* Behavioural therapies includes CBT, psychotherapy and hypnotherapy

5.4 Submission of evidence

No formal request was made for submission of evidence.

5.5 Formulating recommendations and determining key recommendations

EVIDENCE TO RECOMMENDATIONS

Each review summarises the evidence, and the GDG are asked to interpret the evidence before drafting recommendations. In each case, this includes a consideration of the clinical and cost effectiveness evidence; an indication of the factors the GDG took into account, including the balance between benefits and harms; the GDG's reasoning and conclusions, and, where relevant, the level of agreement amongst the group.

This is reported in each individual review section, illustrating the linear relationship between published clinical and cost effective evidence and recommendation for clinical practice.

KEY RECOMMENDATIONS

Methodology

There are generally three main methods reported for developing consensus. These are Delphi, consensus development panels and nominal group processes (Bowling 2002). The nominal group technique (NGT) was originally developed by Delbecq et al (1971) as an organisational planning tool. The methodology allows individuals to work in the presence of others, but verbally interaction is prevented, enabling consensus to be developed without the social pressures normally exerted through open dialogue (Zastrow and Navarre 1977). Individual ideas are shared within the group, with facilitated discussion enabling the group to see how individuals are expressing their ideas. Normal practice is for the facilitator to then ask the group to prioritise, with aggregated rankings recorded. This methodology works extremely well towards the end of guideline development, particularly in relation to developing consensus agreement.

The GDG having worked together for the previous 12 meetings had become a mature working group; individuals within the group were able to express their views relating to key recommendations within a social setting (the last GDG meeting). This was important for the group, who were able to use this experience and the content of discussion to then go into a round of voting to move agreed recommendation into a potential top 10 list, which reflected the key priorities for the guideline. Iteration is usual within consensus methodology, and a second round of voting was necessary in order to gain full consensus within the group.

Process

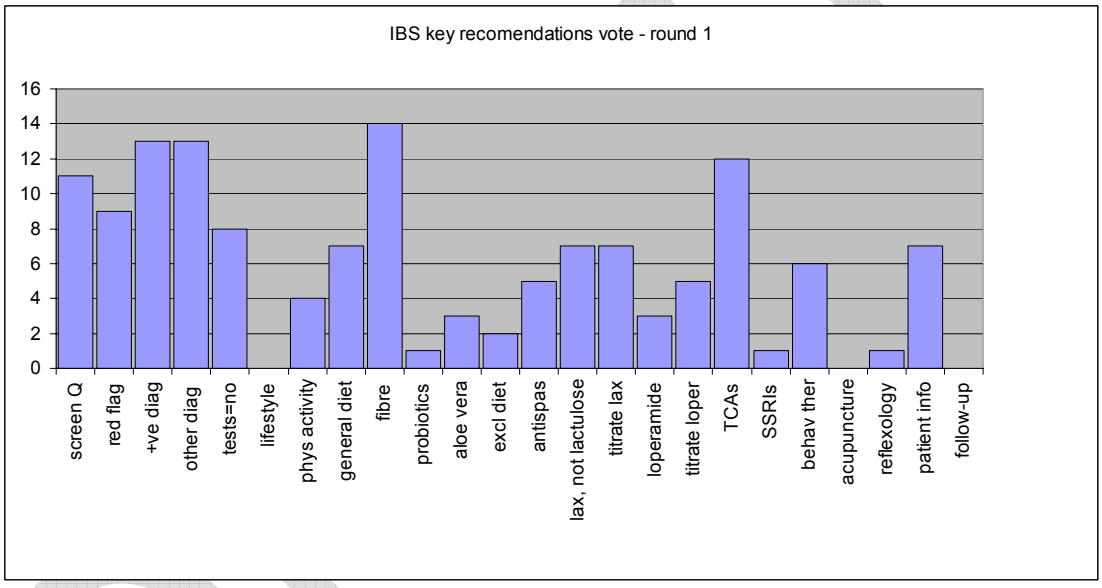
The GDG was asked to vote on key recommendations by secret email ballot using an Excel spreadsheet. This incorporated the full list of recommendations and votes were allocated to the

group, in order to try and determine the key priorities for the guideline. Developing consensus through validated instruments is key to ensure that the final list of up to ten key recommendations fully reflect the group as a whole. This enables all constituent members of the group to have equal weighting of opinion as their opinion moves towards a consensus group position. Typically, NGT works well for small groups, with 12 to 15 people widely acknowledged in the literature as the maximum number of people involved in this process.

Results in round 1: 15 GDG members voted (100%), but one voting paper was spoiled and we were unable to obtain clarification from this member. Therefore results were based on a 93% representive opinion of the GDG relating to Round 1 voting.

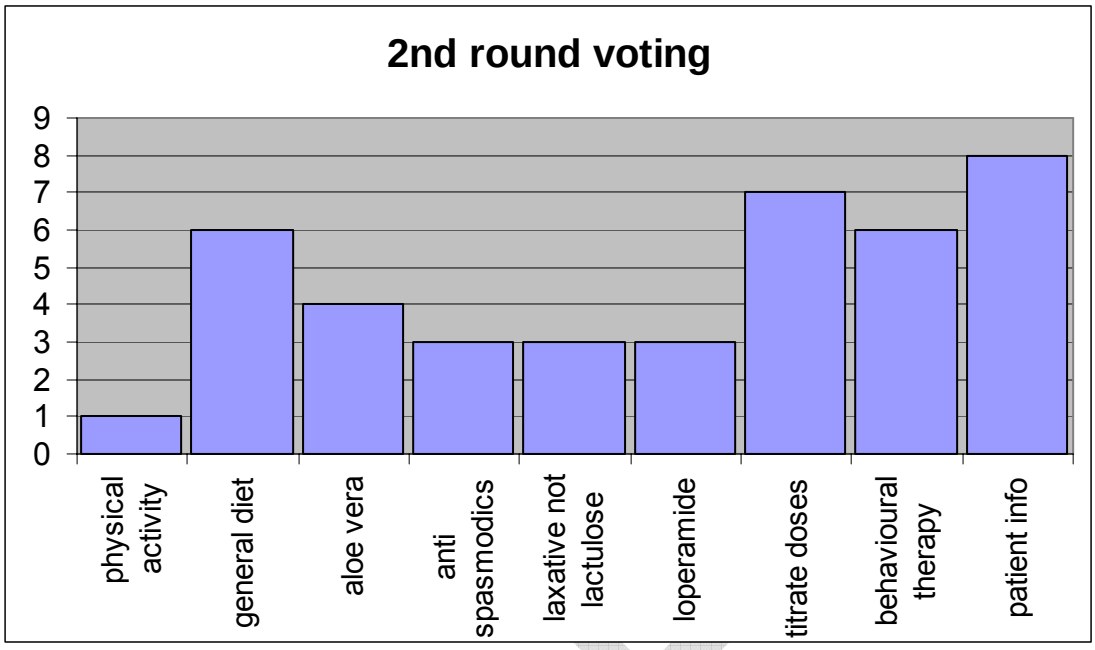
The results for this round of voting are seen below in table 1.

Table 1.



All recommendations with more than 50% of the vote were selected automatically as key recommendations; i.e. screening question, red flags, positive diagnosis, other diagnostic tests, tests that should not be done, fibre, and tricyclics. This gave seven recommendations, but the next highest results gave four recommendations with 7 votes. This determined the need for a second round of voting. Those recommendations with 2 or fewer votes were excluded, and the GDG were asked to choose three of nine recommendations. Between the two rounds, two recommendations were combined (the two relating to titration of medication doses) and the patient information recommendation was revised.

Results in round 2: 14 of 15 GDG members voted and one member only voted for two recommendations. Results are based on 93% group representive opinion of the GDG.



In analysing the voting for round 2, two further recommendations were selected: patient information and titrating doses of laxatives and antimotility agents. Two other recommendations had six votes each, general diet and behavioural therapies and it was decided to exclude both of them, leaving the following nine key recommendations.

Summary

The NGT worked well in developing consensus opinion, reflected by the key recommendations emergent from the process. The nine key recommendations represent the heart of the full guideline and full guideline recommendations. They articulate the evidence supporting the key areas of healthcare practice that will be shaped by the guideline, providing the possibility with effective implementation for people with IBS symptoms being properly diagnosed and managed within primary care.

6 DIAGNOSIS

INTRODUCTION AND BACKGROUND

Irritable bowel syndrome (IBS) is a chronic, relapsing and often life-long disorder. It is characterised by the presence of abdominal pain associated with defaecation, or a change in bowel habit together with disordered defaecation (constipation or diarrhoea or both), and the sensation of abdominal distension. Symptoms sometimes overlap with other gastrointestinal disorders such as non-ulcer dyspepsia, or with coeliac disease. Diagnosis of IBS has proven difficult historically for many reasons, not least that traditionally an exclusion diagnostic approach has been selected by clinicians. Each year, typically approximately 10% of the population will experience IBS symptoms, with up to half of these presenting to primary care clinicians. In reviewing the literature, it is clear that in the absence of gold standard diagnostic criteria, several criterion referenced diagnostic tools have emerged over the last two decades. These have been used in both prevalence and incidence studies, and have proven to be useful for clinicians in enabling them to provide a diagnosis for those patients presenting with IBS symptoms. These criteria have also allowed for standardisation of IBS diagnosis in research.

Definition

For the purpose of this guideline, IBS is defined using the Rome II criteria, used mainly in the context of research. The Rome group is a pan-European clinician group that have met for the last decade, seeking to provide both clarity and direction for clinicians and patients alike.

The Rome II criteria characterises IBS as:

- At least 12 weeks (which need not be consecutive), in the preceding 12 months, of abdominal discomfort or pain with two of the following three features:
  - Relief by defaecation
  - Onset associated with a change in frequency of stool
  - Onset associated with a change in stool appearance.

The IBS population

IBS most commonly affects people between the ages of 20 and 30 years and is twice as common in women as in men. The prevalence of the condition in the general population in the UK is estimated to lie somewhere between 10 and 20%. Recent trends indicate that there is also a significant prevalence of IBS in older people; therefore, IBS diagnosis should be a consideration when an older person presents with unexplained abdominal symptoms. Because incidences of other conditions with similar symptoms are higher in the elderly population, use of certain diagnostic tests is warranted. Co-morbid conditions and poly-pharmacy are common in this patient population. The true prevalence of IBS in the whole population may be higher than estimated, because it is thought that many people with IBS symptoms do not seek medical

advice; NHS Direct online data suggest that 75% of people using this service rely on self-care. In England and Wales, the number of people consulting for IBS is extrapolated to between 1.6 and 3.9 million. Evidence suggests that age and race have no consistent effect on the incidence of symptoms.

Investigations commonly requested by clinicians

Primary care investigations are likely to include: routine blood tests such as full blood count, urea and electrolytes, liver function tests; tests for thyroid function, tissue transglutaminase anti-endomysial antibodies (test for coeliac disease); inflammatory markers, stool microscopy; urinary screen for laxatives; and lactose tolerance testing. Other investigations such as gut transit studies (radiological tests to measure the time required for food to move through the digestive tract) and sigmoidoscopy (endoscopy of the lower part of the bowel) are routinely performed in secondary care. Determining the criteria for such requests and appropriate referral into secondary care will be addressed in the guideline.

The need for effective diagnosis – clarifying concepts

IBS is associated with a disproportionately high prevalence of abdominal and pelvic surgery, although the cause of this has not been established. Diagnostic test methodology has traditionally been applied when comparing a new or alternative test with the acknowledged gold standard reference.

Gold standard reference points aim to represent the 'truth', and when a test is carried out there are four possible outcomes. These are:

1. True positive (detects disease when present)
2. False positive (detects disease when it is absent)
3. True negative (can identify absence of disease)
4. False negative (can identify someone as being disease free when they have it).

It is widely acknowledged within the literature that there is no gold standard reference for the diagnosis of IBS, which means that comparison of definitive diagnostic tests remains difficult. Diagnostic criteria in themselves can be seen as having enormous value, and these could be directly compared in other disease areas to the gold standard reference. In this narrative review of 170 studies/papers, comparisons of criteria are made against a definitive diagnosis of IBS through clinician expertise, augmented by a whole battery in many cases of diagnostic investigations.

In measuring accuracy of diagnostic test/criteria, two measures are used. These are sensitivity and specificity.

Sensitivity

This is a measure (usually expressed as a % of the total population that the test is applied to) that indicates how good the test is in identifying people with the disease.

Specificity

This is a measure (usually expressed as a % of the total population that the test is applied to) that indicates how good the test is in identifying people without the disease.

Problems associated with using these as single measures are acknowledged, as they are difficult to interpret for individual patients. For example, if a test has a sensitivity of 85%, what if I am one of the 15% that the test has failed to identify. In real world situations, what patients and clinicians generally want to know is *'If this test is positive, does it mean that I have a positive diagnosis?'* or; *'If this test is negative, does this mean that I do not have the disease?'*

What may be more useful is for these single measures to be expressed as a probability; a likelihood of accuracy. Again this is expressed as a %, with positive tests measured against a whole study population who had the test. For example, 37 positive results out of 100 would be a 37% prediction, expressed as positive predictive value (PPV). This can be viewed from the reverse perspective; how many negative tests were recorded out of the total study population who had the test. For example 63 negative results out of 100 would be a 63% prediction, expressed as negative predictive value (NPV). Using real data (Steurer 2002), of 1000 women who received a positive mammogram result, 90 actually had breast cancer, meaning that the PPV for mammography is 9%. Converted to probability, this means that women have a 1 in 11 chance of having breast cancer if they have a positive mammogram result. Of the 12,102 negative mammogram results, 12,090 did not have breast cancer (meaning that 12 did have breast cancer). This means that the NPV for mammography is 99.9%. Converted to a probability, this means that women have a 1 in 1000 chance of having breast cancer if they receive a negative mammogram result.

Odds ratios

This is another way of measuring test accuracy (see Appendix 2 of this chapter). Its real value is in estimating test accuracy. This is calculated using test Likelihood Ratio's (LR) by taking the positive LR and dividing this by the negative LR. Likelihood ratios are useful in estimating the value of diagnostic tests, and as a general principle, the higher the likelihood ratio the more useful that test will be. A high odds ratio is an indicator of a good diagnostic test.

A main aim of the guideline

One of the main aims of this guideline is to identify diagnostic criteria for people presenting with symptoms suggestive of IBS and to ensure that primary care clinicians and people who may have IBS have a reference tool that is both sensitive and specific, with high predictive value of

the syndrome. This is an area of healthcare practice which is currently absent, and creates great uncertainty for both clinicians and people who may have IBS.

OBJECTIVES

1. To determine the effectiveness of diagnostic criteria for people with IBS.
2. To determine the clinical utility of diagnostic tests to exclude alternative diagnoses in people meeting the diagnostic criteria for IBS.

SELECTION CRITERIA

The selection criteria for this systematic narrative review was to analyse all relevant literature related to diagnosis of IBS. Due to the absence of a gold standard reference for this disease, diagnostic review methodology was not applicable. On this basis, the GDG accepted that a systematic narrative review was the best way of measuring current practice against peer reviewed literature. This review formed the basis for GDG consensus discussions and recommendations for diagnosis of IBS. Studies identified were then quality assessed. Studies included in the review importantly had to have used a criterion referenced diagnostic tool, studies that failed to do so were excluded from the review. This ensured that all relevant studies provided the evidence base in validating a diagnostic tool, enabling primary care clinicians to make a positive IBS diagnosis around symptom recognition.

Types of studies

All published literature on IBS diagnosis was included. This resulted in a large search and, post-sifting, a large number papers being reviewed for potential inclusion in the review.

SEARCH STRATEGY FOR IDENTIFICATION OF THE LITERATURE

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additional databases were not searched for this review. A sensitive search strategy was employed, as recommended by Haynes and Wilczynski (2004) in determining optimal search strategies for retrieving scientifically strong studies of diagnosis. The search strategies are listed in Appendix B.

METHODOLOGY

The benefits of a systematic narrative review of the clinical evidence in the absence of diagnostic test studies are highlighted by Oxman and colleagues. Applying the quality assurance principles advocated by Oxman (1994), a valid review article can provide the best possible source of information that can lay a foundation for clinical decisions to be made. There is argument that focused narrative reviews for individual outcomes, in this case, IBS diagnosis, are more likely to provide valid results that are useful for clinicians.

STRUCTURE OF THE REVIEW

Having provided the background and context for this review, diagnostic criteria are presented that emerge from the systematic narrative review of the literature. Data is presented in three main sections of the review.

In the first section, the evidence relating to the use of criterion based tools in the diagnosis of IBS is presented and discussed. The effectiveness data for each tool is summarized in Table 1 and specificity, sensitivity and positive predictive value of the criteria are reported where available. The excluded studies are listed in Appendix E and are excluded on the basis that no criterion reference tool was used. This systematic narrative review is followed by a description of an interactive exercise used by the GDG to assess the strengths and weaknesses of each of the identified tools.

In the second section, the evidence relating to the utility of tests to exclude alternative diagnoses is presented and discussed. This is followed by a review of the economic literature for diagnostic testing and an adaptation of one of the cost-effectiveness models identified in the economic literature review.

IBS DIAGNOSTIC CRITERIA

The use of diagnostic criteria has merged over the last three decades, with leading GI specialists such as Manning and Kruis leading the way. Such diagnostic criteria were forerunners to a consensus process amongst leading clinicians which became known as the Rome process. Rome III is the latest iteration and builds on the validated work from authors, in particular the Manning criteria.

Pre-Rome

The first paper to address diagnostic criteria for IBS was a working team report published in 1989 for the 1988 International Congress of Gastroenterology in Rome, Italy. This is acknowledged as the Rome criteria.

Establishment of Rome Committee Process

Following the 1989 publication, a committee was set up the same year to develop for the first time a classification system for all the 21 functional gastro intestinal disorders (FGID). This report was published in 1990 heralding the beginning of the Rome Criteria process. The criteria for IBS in 1989 did not feature pain as a symptom, which is now a current ROME criteria requirement for the diagnosis of IBS.

Rome I

From 1990-1995, seven committees formed to elaborate on the 1990 classification system. Knowledge of this classification system was quite limited, since the journal had a small

circulation and was not listed in MEDLINE. The committee however were able to publish a book which featured the updated Rome I criteria in 1992 and it was the first time that pain was required for the diagnosis.

Rome II

By 1995, interest had grown from both clinicians and the pharma industry. Funding was secured from industry to support the development of Rome II. The number of committees was expanded, with wider international contributions forming the basis of this updated set of criteria. Emerging from this process, the criteria were available from 1999 and first published in 2000.

Rome III

Because of the success of the Rome II process, funding support from industry was forthcoming to maintain this consensus process. A co-ordinating committee was formed in 2001 for Rome III (Drossman, Corazziari, Delvaux, Spiller, Talley, Thompson and Whitehead). Work began in May 2003, leading to publication of new criteria in 2007.

Kruis criteria

The aim of the original study was to create a scoring system for IBS diagnosis incorporating history, physical examination and some basic investigations (ESR and blood count).

Validated criterion reference tool reviewed and acknowledged as used within practice over the last 3 decades.

Kruis patient questionnaire

1. Did you come because of abdominal pain? No Yes
   Do you suffer from flatulence? No Yes
   Do you suffer irregular bowel movements? No Yes

2. Have you experienced this for > 2years? No Yes

3. How can your abdominal pain be described: burning, cutting, very strong, terrible, feeling of pressure, dull, boring, not so bad?

4. Have you alternating diarrhoea/constipation? No Yes

5. Have your stools any of the following properties? Pencil-like; rabbit pellets; hard in the first portion and looser in the second portion; mucus?

If the patient answers yes in any of sections of each question, a scoring system is allocated as follows:

Question 1	34 points
Question 2	16
Question 3	23
Question 4	14

Question 5 carries no score.

Total score possible 87 points.

The patient questionnaire is then validated by the clinician who can subtract from the original total if they identify markers or indicators of disease, potential red flags.

Kruis clinician questionnaire

1. Abnormal physical findings, and/or history for alternative diagnosis of IBS	No	Yes
2. ESR > 20mm/2hr	No	Yes
3. Leucocytosis > 10.000/ccm	No	Yes
4. Haemoglobin F < 12g/ M < 14g	No	Yes
5. History of blood in stool	No	Yes
6. Fever (> 38.5) in the last week	No	Yes
7. Underweight	No	Yes
8. Loss of weight > 5kg in last 6 months	No	Yes

If the clinician answers yes to questions 1 – 5, a scoring system is allocated as follows:

Question 1	- 47 points
Question 2	- 13
Question 3	- 50
Question 4	- 98
Question 5	- 98

Questions 6 – 8 carry no score. This is then subtracted from the original patient score.

Manning criteria

The patient should present with at least 2 of the following symptoms for an IBS diagnosis to be made:

- Onset of pain associated with more frequent bowel movements
- Onset of pain associated with more loose bowel movements
- Relief of pain with defaecation
- Abdominal distension
- Sensation of incomplete evacuation with defaecation
- Passage of mucus.

Rome Criteria

At the 13th International Congress of Gastroenterology in Rome in 1988 a group of physicians defined criteria to more accurately diagnose IBS. The Rome criteria are:

The patient should present with 3 months of continuous or recurring symptoms of abdominal pain or irritation that:

- May be relieved with a bowel movement
- May be coupled with a change in frequency, or
- May be related to a change in the consistency of stools.

Two or more of the following are present at least 25 percent (one quarter) of the time:

- A change in stool frequency (more than 3 bowel movement per day or fewer than 3 bowel movements per week)
- Noticeable difference in stool form (hard, loose and watery stools or poorly formed stools)
- Passage of mucous in stools
- Bloating or feeling of abdominal distention
- Altered stool passage (e.g. sensations of incomplete evacuation, straining, or urgency).

Rome I criteria (1992)

The patient should present with at least 3 months of continuous or recurrent symptoms for an IBS diagnosis to be made:

Abdominal pain or discomfort, which is:

- Relieved with defaecation
- and/or associated with altered bowel frequency
- and/or associated with altered stool consistency
- and/or two or more of the following, on at least 1/4 of days:
- Altered stool frequency
- Altered stool form
- Altered stool passage (straining, urgency or tenesmus)
- Passage of mucus
- Usually with bloating or a feeling of abdominal distension.

Rome II criteria

The Rome II Criteria, published in 2000, were developed by 10 multinational working teams that collaborated over 4 years to arrive at a consensus for symptom-based diagnostic standards.

Twelve weeks* or more in the past 12 months of abdominal discomfort or pain that has 2 out of 3 features:

- Relieved with defaecation

- Associated with a change in frequency of stool
- Associated with a change in consistency of stool.

*The twelve weeks need not be consecutive

The following are supportive, but not essential to the diagnosis. One or more are usually present. They add to the clinician's confidence that the intestine is the origin of the abdominal pain. The more of these symptoms that are present, the greater the confidence with an IBS diagnosis:

- Abnormal stool frequency (> 3/day or < 3/week)
- Abnormal stool form (lumpy/hard or loose/watery) > 1/4 of defaecations
- Abnormal stool passage (straining, urgency or feeling of
- incomplete evacuation) > 1/4 of defaecations
- Passage of mucus > 1/4 of defaecations
- Bloating or feeling of abdominal distension > 1/4 of days.

ROME III Diagnostic Criteria*

Recurrent abdominal pain or discomfort** at least 3 days/month in last 3 months associated with two or more of criteria #1 - #3 below:

Pain or discomfort at least 2-3 days/month (question 1>2)

For women, does pain occur only during menstrual bleeding? (question 2=0 or 2)

1. Improvement with defaecation

Pain or discomfort gets better after BM at least sometimes (question 4>0)

2. Onset associated with a change in frequency of stool

Onset of pain or discomfort associated with more stools at least sometimes (question 5>0), OR Onset of pain or discomfort associated with fewer stools at least sometimes (question 6>0)

3. Onset associated with a change in form (appearance) of stool

Onset of pain or discomfort associated with looser stools at least sometimes (question 7>0), OR Onset of pain or discomfort associated wit harder stools at least sometimes (question 8>0)

* Criteria fulfilled for the last 3 months with symptom onset at least 6 months prior to diagnosis

Yes. (question 3=1)

**"*Discomfort" means an uncomfortable sensation not described as pain*.

In pathophysiology research and clinical trials, a pain/discomfort frequency of at least two days a week is recommended for subject eligibility.

Pain or discomfort more than one day per week (question 1>4)

How to use the questionnaire?

Criteria for IBS-C

(question 9>0) and (question 10=0)

Criteria for IBS-D

(question 9=0) and (question 10>0)

Criteria for IBS-M

(question 9>0) and (question 10>0)

Criteria for IBS-U

(question 9=0) and (question 10=0)

The timely publication of ROME III is beneficial to this guideline, it brings together many studies that have incorporated ROME criteria, and this latest iteration closely aligns the thinking of the GDG, in particular in relation to the implementation of diagnostic criteria for primary care clinicians to use.

ROME III Criteria - Questionnaire

1. In the last 3 months, how often did you have discomfort or pain anywhere in your abdomen?	0 Never → 1 Less than one day a month 2 One day a month 3 Two to three days a month 4 One day a week 5 More than one day a week 6 Every day	***Skip remaining questions***
2. For women: Did this discomfort or pain occur only during your menstrual bleeding and not at other times?	0 No 1 Yes 2 Does not apply because I have had the change in life (menopause) or I am a male	
3. Have you had this discomfort or pain 6 months or longer?	0 No 1 Yes	
4. How often did this discomfort or pain get better or stop after you had a bowel movement?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	
5. When this discomfort or pain started, did you have more frequent bowel movements?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	
6. When this discomfort or pain started, did you have less frequent bowel movements?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	
7. When this discomfort or pain started, were your stools (bowel movements) looser?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	
8. When this discomfort or pain started, how often did you have harder stools?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	
9. In the last 3 months, how often did you have hard or lumpy stools?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	Alternative scale: 0 Never or rarely 1 About 25% of the time 2 About 50% of the time 3 About 75% of the time 4 Always, 100% time
10. In the last 3 months, how often did you have loose, mushy or watery stools?	0 Never or rarely 1 Sometimes 2 Often 3 Most of the time 4 Always	Alternative scale: 0 Never or rarely 1 About 25% of the time 2 About 50% of the time 3 About 75% of the time 4 Always 100% time

Table 1: Summary Table of Diagnostic Papers with diagnostic data provided

Criteria used in study	Study authors/Name of study	N= (if appropriate)	Sensitivity/Specificity Se= Sp=	Predictive value (PPV)
Kruis	Kruis et al 1984	N=108	Se= 83% Sp = 97% Accuracy if score ≥ 44 is 99%	Based on IBS prevalence if score is ≥ 44 10% 87.1% 30% 96.4% 50% 98.4%
	Dogan and Unal 1996a,Turkey	N=347	Se= 81% Sp= 91% if score of 44 points was positive	PPV= 90%
	Frigerio et al 1992 Italy	N=1257	Se= 47% in men, 60% in women Sp= 94% men, 95% women	PPV= 54% men, 82% women Negative Predictive value 91.6% men - 87.3% women
	Osset et al 1991 Italy	Quoting from Kruis 1984	Se= 83% Sp= 97% 99% accurate if score is > 44 points	
	Dogan and Unal 1996b,Turkey : Manning discriminated IBS from OGD	N=347	Se= 90% Sp= 87% if > 3 positive.	PPV=87%
	Rao et al 1993	N=123	Se=67% Sp=93%	PPV=93.4%
	Talley et al 1990	N=361	Se= 42% Sp= 85%	
Manning + Kruis	Dogan and Unal 1996c,Turkey: Correlation significant in IBS r=0.714 p=<0.05 but not in OGD r =	N=347	Se= 80% Sp= 97%	PPV=96%

	0.190 p=>0.05			
Manning (3/6)	Jeong et al 1990 Smith 1992 Manning ≥ ¾	N=172 N=109	Se= 67% Sp= 70% Se= 63% Sp= 85%	
Manning (>3/6)	Talley et al 1990 Kruis et al 1984	N= ?? N=479	Se= 84% Sp= 76% Se= 64% Sp= 99%	 94%
Rome	Saito et al 2003a, USA Vanner et al 1999	Prevalence Cohort study 1st survey 1987n=1121 2nd survey 1989 3rd survey 1992 n=892 response n=643 (72%) N=384 (retrospective) N=95	Prevalence rates by criteria: Rome (1989) 27.6 per 100 (95%CI:23.6-31.5) Rome (1990) 5.1 per 100 (95%CI:3.2-7.1) Se= 63% Sp= 100%	 98%
Rome I	Saito et al 2003b, USA Chey et al 2002a, USA Mearin et al 2001a Spain Patients diagnosed with Manning, Rome I and Rome II. > 2/3 of subjects fulfilling Manning or Rome I would not be diagnosed as having IBS if using RomeII.	 N= 1014 women N=281	Rome I (1992) 6.8 per100 (95%CI 4.7-8.9) Se=83% Sp= not given Se/Sp = not given	Good agreement between Rome I and II (>95% Kappa >0.68)
Rome II	Saito et al 2003c, USA		Rome II (1999) 5.1 per 100 (95% CI:3.1-7.0)	Rome II and Rome (79% kappa 0.29)

	Chey et al 2002b, USA Difference in sensitivity seemed to be attributable to more restrictive time requirement for pain with Rome II Boyce et al 2000, Australia (prevalence study)	N=1014 women N=2910	Se= 47% Sp=not given See table 2 in paper	Rome II more restrictive. Results similar for other studies Mearin *et al*, Thompson *et al* Chey *et al* If different thresholds are used subjects identified are not the same. Manning identified less severe symptoms. Treatment would be no different using any criteria.
BDQ (Talley et al) Validated q'aire for identifying IBS	Bijkerk et al 2003, Netherlands	N= 99	All patients had diagnosis of IBS but only18% (n=14) met Rome II GP diagnosis based on Bloating (87%) and absence of alarm features (87%) rather than diagnostic criteria. GP diagnosis correlated most closely with Manning. GP's reported tests to exclude organic disease in pts over 50.	

RESULTS OF THE REVIEW AND DISCUSSION

Of the 170 papers reviewed, 45 were excluded as no diagnostic criteria were used in the study population, or they were discussion/professional papers highlighting aspects of care relating to IBS, of which diagnosis was mentioned. Of the remaining 125 papers, 18 studies provided useful data which has been extracted and presented in Table 1. The remaining papers all provided useful background information on the use of diagnostic criteria; many studies used the reference criteria as a way of measuring prevalence and incidence of IBS in general populations. Literature was drawn from a wide international base, with Europe, North America and South East Asia providing the main source of data.

Issues to consider

From this extensive review of the literature, a number of key observations have emerged which the GDG will need to consider in moving towards consensus opinion as to how IBS is diagnosed; which criteria to use in diagnosing IBS and how this potentially will move clinicians away from an exclusion diagnostic approach towards positive diagnosis of IBS, management and patient follow-up. Of significance is the potential cost saving to the NHS of tests that are routinely requested but prove to be of little added value.

Diagnostic thresholds

Clinicians need to be able to determine whether a person has IBS (or not). Balance between missing the diagnosis and over diagnosing is a possibility because criteria may be too vague. Thresholds can be set across different parameters. These include:

- Severity of bowel symptoms
- Symptom count – either all symptoms given equal weight (e.g. Manning – 2 or more symptoms being given equal weight) or identified symptoms being considered as essential with others considered as accessory symptoms (Rome I)
- Duration threshold in combination with symptom count (Rome II) and symptom frequency appears to be highly relevant (Mearin 2003)
- Rome II requires the presence of both abdominal pain and changes in bowel habit and duration of symptoms (at least 12 weeks in last 12 months)
- Rome considers abdominal pain and changes in bowel habit independently and no minimum duration of symptoms
- Kruis developed according to symptoms and evaluation of lab tests
- Following a positive diagnosis of IBS based on clinical criteria clinicians can be reassured that the diagnosis is durable. (Adeniji 2004, follow-up of 196 patients with a diagnosis of IBS between 1989-92 35/86 pts had had further diagnostic investigations without diagnosis changing)
- Prevalence of IBS measured across a New Zealand population cohort (N=980) using 2 of the manning criteria emerged as 18.1%. This decreased to 10.3% if 3 or more of the

Manning criteria were used to identify an IBS diagnosis. This then fell to 3.3% when 4 or more of the criteria were used (Barbezat 2002).

A key question for the GDG was "Should a positive diagnostic approach include severity threshold or disabling threshold?"

What happens to patients who do not meet Rome II?

- How important is diagnostic precision for clinicians (this is a different priority for research)? Is the priority to exclude structural cause and/or red flags for symptoms?
- Would treatment be different if patients were diagnosed using Manning or Rome? Does this matter?
- Do they have alternative diagnosis (eg FBD) and/or go for lots of investigations and then turn out to be IBS?
- Manning: discriminates upper GI disease from IBS but not IBD (3 or more Manning criteria were frequent in patients with ulcerative colitis in remission (Isgar 1983).
- Kruis: was not able to discriminate IBS from organic GI Disease (Frigerio 1992).

Clinician ignorance of IBS diagnostic criteria

- An Italian study (Bellini 2005) of 28 generalist GPs – 17 judged knowledge of IBS to be insufficient but only three thought that further education might be useful. Ten GPs were familiar with Rome II prior to the study; 19 agreed with Rome II criteria when they used them as part of the study. They reported satisfactory management in approximately 60% of patients.
- Important symptoms for this group – Primary symptoms: changes in bowel habit, abdominal pain relieved by evacuation, bloating. Secondary symptoms: difficult or incomplete evacuation, passing mucous with BM.
- Following clinical evaluation GPs ordered further investigations in large numbers of patients. GPs with more than 20yrs experience requested less investigations than younger colleagues (p= 0.001). 168 patients had routine bloods, 30 – abdominal ultrasound, 87 FOB, 83 faecal analyses, 81 – Thyroid Function, 70 – lower GI endoscopy.
- All agreed that counselling, reassurance, information about natural course of condition and suggestions for coping strategies were first step in management. Patients with diarrhoea were prescribed wider ranging therapeutic interventions and perceived to be in more need of further investigations than those with constipation.
- In 145/229 cases referral to at least one specialist was made (GI specialist most common but gynaecological referral for 19% of women). Referral did not vary with clinical presentation and the most common reasons for referral were diagnostic confirmation, patient need for reassurance and patient request.

- Patients frequently attribute food intolerance (37%) and stress (43%) as the causes of IBS. GPs consider stress (71%), fibre deficiency (83%) and disturbed motility (62%) as most important factors (Bijkerk 2003).
- Helpful clues to aid diagnosis: symptoms chronic or recurrent; pain is variable in location and timing; D and C may alternate; onset sometimes follows GI tract infection; findings on physical examination are usually normal except for abdominal tenderness (Paterson 1999, Canadian IBS position statement).
- USA Primary Care practice based diagnostic evaluations differ significantly from speciality expert opinion based guidelines. Implementation of speciality guidelines in Primary Care would increase utilisation but with limited improvement of diagnostic outcomes (Yawn 2001).
- Patients under 50 yrs of age who meet Manning and have no red flags require no investigations (Paterson et al 1999 Canadian IBS position statement).
- Patient expectations: reassurance, counselling, pharmacotherapy, diagnostic tests and referral to specialist. Dietary interventions were considered less important. Most people with IBS in this study stated that improvement in worst symptom should be target of treatment. Global improvement and QoL were considered much less important as treatment goal.
- A British study in general practice (n=400) Gladman and Gorard (2003) Sent a questionnaire to a random selection of 200 GPs and 200 clinician members of British Society of Gastroenterologists asking about their knowledge of functional GI disorders and their knowledge and use of Manning and Rome criteria for diagnosis. 68/137 GPs believed functional GI disorders were psychosomatic compared to 36/167 of consultants (p=<0.001). Consultants believed that understanding had increased over last 20 years; 50%GPs believed it had not changed. Both believed diagnosis and management had not improved in past 20 years. Only 29/137 GPs had heard of Manning; 16/137 of Rome compared to 134/166 and 139/167 Consultants respectively (p=0.0001). Only 18 GPs used either Manning or Rome in practice and despite increased awareness only 40% consultants used one or other diagnostic criteria in their practice.

Many studies of IBS and developed guidelines to date have been produced by specialists who had seen patients with particularly severe or intractable symptoms. This clinical guideline was developed from a different starting point, with the essential focus being in primary care. All development and implementation must be framed with questions of applicability to General Practice.

The GDG noted that GPs consider Rome II too complicated and more suited for use in secondary care or for research purposes (Thompson 1997; Bellini 2005). The need for a pragmatic useful diagnostic tool was felt to be the most important aspect to the review. As the majority of IBS patients are treated by GPs, any recommendations for the use of diagnostic criteria should ensure that their ease of use by GPs in their practice is established, rather than

expecting GPs to use criteria that do not apply in the reality of day-to-day practice (Corsetti and Tack 2004).

What is the role of red flag symptoms alongside diagnostic criteria IBS diagnosis?

- The Manning criteria do not consider red flag symptoms. The addition of red flag symptoms seems to enhance diagnostic accuracy (Paterson 1999, Canadian IBS position statement). The GDG considered this aspect of the review at length, recognising the need for recommendations supporting the IBS algorithm to ensure that red flag symptoms take the patient out of this guideline and into other related NICE guidance.
- The addition of red flags to the Manning criteria increases the PPV of Manning and Rome I and II (Vanner 1999; Hammer 2004).
- Red flag symptoms – these seem to enhance original criteria and importantly relate this guidance to other relevant NICE guidelines, in particular NICE Clinical Guideline 27 'Suspected Cancer Referral' published in 2005.

Discussion

The need for clinicians to be guided in the diagnosis of IBS has emerged as a strong recurrent theme in this systematic narrative review. The seminal work of Manning laid a foundation to enable clinicians to be guided by such criteria in attempting to provide direct answers to patients presenting with a range of symptoms. This work has undoubtedly influenced the development of thinking within the Rome group, and the Rome criteria recently published as Rome III reflects the benefit of validation of the key aspects of the criteria and pragmatic decisions relating to the length of presenting symptom such as pain (6 months). The use of available diagnostic criteria summarised in this review have typically been augmented by further diagnostic investigations that have limited or no benefit and these add considerable costs to the NHS.

The use of consensus agreement regarding the recommendation of single diagnostic criteria serves three main purposes:

- Increased patient confidence through positive diagnosis
- Increased clinician confidence
- Potential for considerable NHS disinvestment in avoiding unnecessary investigations and referrals to multiple specialities.

When looking at combinations of possible criteria used in the available diagnostic tools reviewed, the emergence of Rome III has proven to be timely in relation to guideline development. It features strengths of previous diagnostic criteria, while minimising weaknesses of reviewed tools.

Key questions that emerge from the literature aim at identifying a gold standard or best reference tool. The challenge is that when thresholds differ, results are inconsistent. For

clinicians, diagnostic precision of IBS is often of low priority when compared to excluding other structural cause. This is a conceptual misinterpretation which can be explained as under-confident application of clinical examination and clinical history interpretation. Perhaps of significant note, regardless of which criteria were used in included studies in the review, treatment rarely differed against symptom profiles.

Over a decade ago, Jeong and colleagues having identified that Manning had reasonable specificity, called for better diagnostic criteria with improved accuracy to be developed. The road to Rome and the subsequent development of international consensus over the last 15 years has been useful in predetermining consistency in research application. It however, may have distracted from refinement of a tool that is easy and straightforward to use for primary care clinicians.

What clearly emerges from the literature is that with careful history and physical examination, positive diagnosis of IBS is possible. This, augmented by simple laboratory investigations to rule out more serious underlying pathology in the absence of red flag symptoms, would seem to be a step forward for both clinicians in diagnostic practice and patients in receiving timely interventions.

Perhaps it is fitting to highlight within this review that clinicians have been seeking to change the way that they think about diagnostic approaches in relation to this chronic syndrome. The Manning (1978) paper titled “Towards a positive diagnosis of Irritable Bowel Syndrome’ clearly indicates a diagnostic aim, this over the last three decades has been lost, with negative diagnosis by ruling out other conditions being the predominant clinician approach. Returning to the original aim of Manning and colleagues by seeking to influence the behaviour of primary care clinicians in the way that they think and approach diagnosis of people presenting with IBS symptoms is an important objective of this guideline.

GDG interpretation of the review and application of the guideline

General practitioner (GP) training has focussed on the importance of what happens within a typical patient consultation. This is usually recorded and analysed to enable new GPs to reflect on the detail within the consultation, in particular the quality of verbal and non-verbal behaviour, the sequencing of questions and information gathered to enable diagnosis. This is based around simulation and objective structured clinical examination methodology and has effectively enabled GP trainees to experience and develop understanding related to the importance of clinical history prior to physical examination. Using this approach, the NCC-NSC planned an interactive session for the GDG to fully engage with relevant issues. This was felt to be important in demonstrating that in guideline development, the GDG had explored the utility of different criteria that would then inform any consensus recommendations and lay the foundation for a positive implementation experience.

In order to test the utility of different criteria, NCC-NSC staff ran an interactive diagnostic simulation with members of the GDG. A number of typical IBS patient profiles were written by the technical team, which were then shared with four subgroups of the GDG. Details of the patient profiles are listed in Appendix 3 to this chapter.

Sub group constituency:

- Primary or secondary care doctor and/or
- Primary or secondary care nurse and/or
- Allied health professional (eg. Dietician or pharmacist) and/or
- Patient representation allocated to comment and input across the four groups.

One patient profile randomly selected from the total number of prepared patient profiles were randomly allocated to each of the GDG subgroups. Each group then had to discuss the use of allocated diagnostic criteria and elect one member of the group to role-play a GP consultation, responding to the simulated patient profile. Group A were asked to use Kruis Criteria; Group B were asked to use Manning criteria; Group C were asked to use Rome criteria;Group D were asked to use Rome II criteria.

Each group selected a physician to role-play the consultation, timed at a typical 8 minute general practitioner consultation. Two groups selected their GP member, with the other group selecting the GDG clinical lead who is a Gastroenterologist. Members of the NCC-NSC team role-played the four different patients. ROME III at the time of the patient simulation was unpublished, and therefore was unable to be used in this exercise. During each of the four role-plays, GDG members were asked to observe the consultation and record their observations. These typically related to the ease and logical progression of the consultation, shaped by the diagnostic criteria. This simulation enabled the GDG to both interpret the evidence and evaluate how easily the criterion reference based tool could work within a busy primary care environment.

The NCC-NSC technical team transcribed the detail within each of the GP consultations and recorded the information gathered from the patient using each of the three criteria referenced tools. The content was analysed and grouped in emerging themes (see Table 1) to enable the group to fully understand what was possible in recreating the primary care consultation. Typically patients are reticent to come and see a primary care clinician with issues relating to bowel habit/function, and this reticence was simulated with behaviours that demonstrated both hesitancy and embarrassment.

Observations were recorded from three main areas for feedback:

1. How the GDG clinician felt using the diagnostic criteria allocated to them
- 2. How GDG members felt each of the diagnostic criteria worked in this simulated patient consultation

- 3. How the NCC team member felt when role playing the IBS patient, in relation to the sequencing of ideas and extracting of important patient information, facilitating an effective diagnosis.

This was a powerful exercise in embedding the evidence review into a simulated patient-clinician exchange. The importance of ensuring that the guideline recommendations are able to be effectively implemented into routine primary care is clearly important in ensuring that current variations in diagnosing IBS are addressed.

Outcomes from the evidence review and diagnostic criteria simulated exercise were:

- A strong evidence base for the use of diagnostic criteria with good predictive value.
- Expert (GDG) evaluation of how potential tools could enable primary care clinicians to make a positive diagnosis of IBS, supported by a limited number of investigations that may augment an IBS positive diagnosis.
- Agreement of evidence based positive diagnostic criteria for use in primary care which reflects current evidence.
- A contemporised Manning criteria which are consistent with ROME III criteria.
- The decision to refer to agreed criteria as 'Positive Diagnostic Criteria'.

KRUIS (A SCORE OF >44 = IBS)

STRENGTHS	WEAKNESSES
VERY SPECIFIC	**NEEDS OTHER TESTS – TIME AND COST IMPLICATIONS**
	TOO MANY QUESTIONS
	SCORING CONFUSING
	COUNTER INTUITIVE – CLOSED QUESTIONS
	OMISSION – RELIEF OF PAIN BY DEFAECATION
	PATIENT NOT REASSURED – NO DIAGNOSIS
	NO WAY TO EXPLORE EXTRA COLONIC SYMPTOMS

MANNING (> 3 CRITERIA = IBS)

STRENGTHS	WEAKNESSES
SIMPLICITY	**VERY 'PAIN' FOCUSSED – NO MENTION OF DISCOMFORT**
CLEAR QUESTIONS	**NO RED FLAGS**
INSPIRED CONFIDENCE	**DOESN'T MENTION CONSTIPATION SPECIFICALLY**
EASIEST TO USE IN WORKING PRACTICE	**LANGUAGE OLD FASHIONED**

TIMESCALES	**OMISSION – FLATUS**
	NOT NATURAL FLOW
	? VALIDITY OF DIAGNOSIS FROM 2 SYMPTOMS

ROME I

STRENGTHS	WEAKNESSES
SIMILAR TO MANNING	**APPLICATION – TIME TAKEN**
ENABLED DIAGNOSIS	**COMPLEXITY**
	COUNTER INTUITIVE – CLOSED QUESTIONS

ROME II

STRENGTHS	WEAKNESSES
THOROUGH CONSULTATION	**APPLICATION – TIME TAKEN**
REASSURING FOR PATIENT	**COMPLEXITY**
	MISSED DIAGNOSIS
	DID NOT INFLUENCE FINAL DIAGNOSIS – USED CLINICAL JUDGEMENT
	COUNTER INTUITIVE – CLOSED QUESTIONS

GDG DISCUSSION

Following the simulated consultation, members of the GDG discussed the high importance of good communication in establishing the clinician-patient relationship. Typically, the evidence demonstrates that diagnostic tendency within primary care is for an exclusion diagnostic approach which is experienced as a negative diagnosis by people with IBS. This can be time consuming, sub-optimal and cost inefficient in relation to unnecessary investigations that are likely to add little or no benefit to predictive value, and can mean that patients are subject to inappropriate referrals to other specialities such as gynaecology. The GDG reflected the importance of language used in the first meeting between people with IBS symptoms and primary care clinicians, supporting the positive diagnostic approach of *'you have IBS'* as opposed to '*all investigations are negative and you have nothing else wrong with you, it must be IBS'*. This exclusion diagnosis approach is widely reported as typical in the patient population, supported by both patient representatives on the GDG.

Themes emerging from the exercise and focussed GDG discussion

- IBS is a lifelong condition that needs to be managed effectively.
- Symptoms that are most crucial in diagnosis are pain/discomfort relieved by bowel movement, bloating (more common in women; men describe it as abdominal tension) and disordered bowel habit. It was noted that language in the Manning criteria needed to be

contemporised; this was discussed and agreed, e.g. pain was contemporised to pain/discomfort.

- Pain/discomfort induced by eating is also common symptom.
- Extra-colonic symptoms were commonly reported in secondary care – with good discussion around their prevalence in primary care.
- The severity of the condition may or may not be useful as a threshold. Whilst high sensitivity maybe attractive, it is important not to miss patients by having too high an exclusion criteria, as reported in the evidence for Rome II, and supported by the simulated consultation feedback and analysis.
- IBS co-exists with other conditions. The possibility of missing inflammatory bowel disease initially would not be perceived by the GDG as problematic
- There is clear evidence supporting diagnostic criterion based reference tools, but their use in practical clinical settings has been reported to be difficult, this was noted by the group and it was felt that published criteria in this guideline should reflect the validated tools, but ordered in such a way that ensures that the tool is intuitive for clinicians to use. It should also facilitate the type of discussion that enables a full history to emerge.
- The individual patient story is very important, emphasising the need for the primary care clinician to focus on the most severe symptom while also establishing other related symptoms.

Published evidence from the diagnostic tools has shaped recent diagnostic approaches for IBS. Whorwell (2006) refers to this as a diagnostic triad, seen below:

- **Pain/discomfort – quality and quantity**
  Site of pain: in IBS it can be anywhere in the gut. If the site of pain varies it is unlikely to be cancer (tumour fixed). Need to distinguish this IBS pain discomfort from that caused by gall bladder disease. IBS patients do not tolerate abdominal surgery well.

- **Bowel habit – quality and quantity**
  Giving patients' descriptive examples (e.g. like porridge, rabbit pellets) and using the Bristol Stool Chart helps. Incomplete evacuation is reported, creating rectal hypersensitivity. Urgency is increased in Diarrhoea; prevalence for those incontinent is 20% (patients often do not disclose unless asked directly).

- **Bloating in women**
  (Absence of bloating in women = red flag) Less common in men, although they may report that the abdomen is tight/hard.

The diagnostic triad clearly reflects the valuable work published by Manning and the Rome group. It also highlights the importance of the extra-colonic features that maybe reported by

people presenting with IBS symptoms, typically these include nausea, low back pain, bladder symptoms and thigh pain.

The GDG agreed that primary care clinicians should ask open questions to establish the multiple features of the syndrome, recognising that a potential conflict may exist within Primary Care in terms of the time available to the clinician in exploring the whole range of presenting symptoms.

Diagnostic certainty

In establishing the sensitivity and specificity of different diagnostic criteria, looking at a pragmatic diagnostic reference tool appears to be of great value to the primary care clinician. The advent of ROME III during the development of this guideline, was both timely and beneficial in shaping the and further strengthening the diagnostic criteria agreed within final recommendations. Of equal value, is the provision of clear economic evidence relating to supplementary diagnostic tests.

UTILITY OF TESTS TO EXCLUDE ALTERNATIVE DIAGNOSES

In order to determine the utility of diagnostic tests used to exclude alternative diagnoses in people meeting symptom based criteria for IBS, we needed evidence on the pre-test probability of organic GI disease in people meeting IBS diagnostic criteria and the accuracy of diagnostic tests in identifying organic GI disease. A published systematic review by Cash (2002) was identified which considered the utility of diagnostic tests by evaluating the evidence in these two areas. The selection criteria for the review were:

- Use of a cohort of IBS patients explicitly diagnosed via symptom based criteria (a priori).
- Performance of common diagnostic tests with either blinded comparison with gold standard.
- Results quantified as normal or abnormal with abnormal test resulting in alternative diagnosis of organic disease.

Six studies were included in the Cash (2002) systematic review and these were quality assessed using eight quality criteria (Hamm 1999; Tolliver 1994; Pimental 2000; Sanders 2001; Francis 1996; MacIntosh 1992). All were prospective cohorts of consecutive patients. All were in secondary care, except Hamm (1999) which did not state whether the participants were in primary or secondary care. The patients in Hamm (1999) were all enrolled in a treatment trial. One study had a control group of healthy volunteers (Sanders 2001). In addition to these six studies our search identified two further studies which had been published since the systematic review and which met the inclusion criteria (Sanders 2003; Pimentel 2003). Each study used different criteria for recruiting patients. These are summarised as follows:

- Referral for abdominal pain not previously evaluated (Tolliver 1994)
- Diagnosis of IBS made at first attendance and evaluated within 6 months (Francis 1996)
- Enrolment in treatment trial i.e. not all recent diagnosis (Hamm 1999)

- Referral for altered bowel habit or requesting investigation to reassure following clinical diagnosis of IBS (Sanders 2001)
- Referral for breath testing (Pimentel 2000)
- All patients attending gastroenterology practice (MacIntosh 1992)
- Primary care attendees, including people entering GP surgery for any reason (Sanders 2003)
- Advertisement within community and IBS support groups (Pimentel 2003).

The study characteristics and results are summarised in Tables 3 to 10 below for each class of diagnostic test. The number of abnormal test results is reported alongside the alternative diagnoses resulting from these tests. Where the tests were not given to the whole study cohort this has been noted. For lactose intolerance and bacterial overgrowth we have noted in Table 4 10 whether the diagnosis was confirmed by an improvement in symptoms following treatment, as an abnormal hydrogen breath test result does not provide a definitive diagnosis of either condition.

Table 2 is reproduced from Cash (2002) and summarises the evidence on the pre-test probability of organic GI disease from the 6 studies included in the systematic review. This is compared to general population data presented by Cash (2002), although it was not clear how the general population sample was defined. In addition to the data presented by Cash, Sanders (2003) reported a general population prevalence of 1% for coeliac disease in people recruited from a UK primary care setting and a prevalence of 3.3% in people meeting IBS diagnostic criteria.

Table 2: Pre-test probability of organic GI disease in people meeting symptom based criteria for IBS and in the general population

Organic GI disease	IBS patients (%)	General population (%)
Colitis/IBD	0.51-0.98	0.3-1.2
Colorectal cancer	0 – 0.51	4-6
Celiac disease	4.67	0.25-0.5
Gastrointestinal infection	0-1.7	N/A
Thyroid dysfunction	6	5-9
Lactose malabsorption	22-26	25

The evidence on the clinical utility of tests for alternative diagnoses in patients meeting IBS diagnostic criteria can be summarised as follows:

- The pre-test probability of organic disorders, including colon cancer, inflammatory bowel disease, thyroid disease and lactose malabsorption was no different in IBS populations when compared to the general population.
- One exception was coeliac disease which did appear to higher incidence amongst the IBS population.

- In the IBS population, common investigations including endoscopy of the colon, ultrasound, stool ova and parasite testing, faecal occult blood, thyroid function testing and hydrogen breath testing for lactose intolerance and bacterial overgrowth were unlikely to lead to the diagnosis of organic disease. Rectal biopsy was also demonstrated to be ineffective.

"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." Jeong et al (1993). Repeated testing can also undermine patient confidence in a positive IBS diagnosis.

Table 3: Colonic evaluation

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Hamm (1999)	Rome criteria met for at least 6 months, and no colonic endoscopic exam in previous 2 years. i.e. not all recent diagnosis	Age <50: Flexible sigmoidoscopy Age >50: Colonoscopy or flexible sigmoidoscopy plus barium enema	None	7/306 (2%) 1146 patients not tested	3 IBD 1 colonic obstruction 3 colonic polyps without malignancy
Tolliver (1994)	International Congress of Gastroenterology Symptom Criteria for IBS. Referred to secondary care without prior diagnosis	Air contrast barium enema, flexible sigmoidoscopy and / or colonoscopy.	None	43 abnormalities in 23 patients (all 196 tested)	2 which could be cause of IBS symptoms 1 IBD 1 cancer
MacIntosh (1992)	IBS patients referred to secondary care, (89% fulfilled Manning 3 or more and 84% fulfilled Rome criteria)	Sigmoidoscopy, colonoscopy, phosphate enema, rectal biopsy	None	0/89 (all patients tested)	None
Francis (1996)	Patients evaluated within 6 months of diagnosis, met Rome criteria and normal stool exam, haematological and biochemical indices including ESR.	Sigmoidoscopy in all, plus barium enema or colonoscopy in over 45 year olds .	None	0/125 (all patients tested)	None except diverticular disease

Table 4: Lactose intolerance

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Hamm (1999)	Rome criteria met for at least 6 months. Not all recent diagnosis	Hydrogen breath test	None – ideally should report response to lactose restricted diet	23% of 1122 patients 330 not tested	Unconfirmed lactose intolerance as no response to treatment recorded
Tolliver (1994)	International Congress of Gastroenterology Symptom Criteria for IBS. Referred to secondary care without prior diagnosis	Hydrogen breath test	3 year follow-up to assess symptoms	48/186 (10 not tested, doesn't state why)	Possible lactose malabsorption but no difference in symptoms at 3 years compared to those without diagnosis

Table 5: Thyroid function

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Hamm (1999)	Rome criteria met for at least 6 months and without test in previous 12 months. Not all recent diagnosis	TSH and thyroxine	None – ideally should report resolution of symptoms following treatment	67/1209 (6%) 3% hypo and 3% hyper	Hypo or hyperthyroidism
Tolliver (1994)	International Congress of Gastroenterology Symptom Criteria for IBS. Referred to secondary care without prior diagnosis	T3 T4 TSH	None – ideally should report resolution of symptoms following treatment	1/171, author states this provided no useful clinical information 25 not tested	Not clear

Table 6: Stool tests

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Hamm (1999)	Rome criteria met for at least 6 months and without test in previous 3 months. Not all recent diagnosis	Faecal ova and parasite test	None – ideally should report resolution of symptoms following treatment	19/1154 (2%) 298 not tested	Enteric infection of unconfirmed clinical significance
Tolliver (1994)	International Congress of Gastroenterology Symptom Criteria for IBS. Referred to secondary care without prior diagnosis	Occult blood and parasites	Occult blood - structural evaluation Parasites – none, should report resolution of symptoms following treatment	Occult blood 15/183 (13 not tested) Parasites 0 /170 (26 not tested)	1 Hemorrhoids, 2 annal fissures, 1 melanosis coli

Table 7: Other laboratory tests

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Tolliver (1994)	International Congress of Gastroenterology Symptom Criteria for IBS. Referred to secondary care without prior diagnosis	FBC, HgB, ESR, Chemistry panel, urine analysis	None	FBCand HgB; 0/196 Chemistry: 2/196 Urine: 4/157 (39 not tested_	No useful clinical information
Sanders (2001)	Rome II without “sinister symptoms” of weight loss, rectal bleeding, nocturnal diarrhoea or anaemia (2ndary care)	FBC, ESR, blood urea nitrogen, serum electrolyte conc, thyroid function, CRP, blood glucose.		CRP: 2/300 ESR: 1/300 Liver function: 2/300 Anaemia: 1/300 All patients tested	3 IBD (abnormal CRP / ESR) 2 excess alcohol (IBS symptom response to reduced intake not reported) Anaemia was secondary to coeliac disease

Table 8: Coeliac screening

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Sanders (2001)	Rome II without "sinister symptoms" of weight loss, rectal bleeding, nocturnal diarrhoea or anaemia (2ndary care)	IgA and IgG antiglandin, endomysial antibody	Duodenal biopsy	66/300 All patients tested	14 Coeliac disease confirmed by biopsy, 1 positive serology but refused biopsy Response to diet not reported
Sanders (2003)	Primary care cross-sectional study, IBS diagnosis from Rom II (subgroup of whole cross-sectional cohort)	IgG/IgA antiglandin and EMA	Small bowel biopsy, and follow-up after diet	Positive tests not reported for IBS subgroup All patients tested	4/123 IBS patients had coeliac disease, all responded to diet

Table 9: Ultrasound

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Francis (1996)	Patients evaluated within 6 months of diagnosis, met Rome criteria and normal stool exam, haematological and biochemical indices including ESR.	Ultrasound of abdomen and pelvis	None	22/125 (18%) All patients tested	No change to IBS diagnosis

Table 10: Bacterial overgrowth

Study	Population tested	Tests used	Gold standard	Abnormal tests	Alternative diagnosis
Pimentel (2000)	Referred for lactulose hydrogen breath test Rome I criteria. Excluded if evidence of rapid transit	Hydrogen breath test	Reported symptom resolution and repeat test result but only in minority of treated patients	157 of 202 (78%)	Only 47 had repeat test to confirm response to therapy 25 achieved eradiation and 45% of these no longer met Rome criteria
Pimentel (2003)	Community and IBS support group advertisement, Rome criteria	Hydrogen breath test	Reported symptom response and repeat test results	84% of 111 had positive first test	20% of those with positive test and antibiotic treatment achieved normal second test, symptom improvement associated with treatment and normal second test

ECONOMIC EVIDENCE ON DIAGNOSTIC TESTS TO IDENTIFY ALTERNATIVE DIAGNOSES

Of the four included studies, two consider the cost-effectiveness of screening for coeliac disease in the IBS population, one considers the cost-effectiveness of endoscopy in the IBS population and one considers the cost-effectiveness of diagnostic strategies for IBD in patients who do not meet the Rome criteria for IBS. The characteristics of the included studies are given in Appendix C. All four were model based economic evaluations with two considering the short-term diagnostic period (Suleiman 2001; Dubinsky 2002) and two considering patient outcomes over longer time-frames of 10 years or more (Spiegel 2004; Mein 2004). The quality of each study has been critically appraised using a validated check-list for economic analyses and details are provided in Appendix D. Due to variation in the interventions and populations considered, each study will be discussed separately.

Mein (2004)

The primary aim of this study was to assess the cost-effectiveness of coeliac disease testing in patients with suspected irritable bowel syndrome in the US health care system. This was done by using a decision tree to estimate the number of coeliac disease cases detected, QALYs gained and costs resulting from three testing strategies and comparing these to no testing. The three strategies were; tissue transglutaminase antibody (TTG), antibody panel, or upfront endoscopy with biopsy. All positive serological tests were followed by endoscopy with small bowel biopsy and the potential complications of this procedure were accounted for. All positive upfront endoscopies were assumed to be confirmed by an antibody panel. Long-term treatment costs were assumed similar between patients with IBS and those diagnosed with coeliac disease. The increase in health-related quality of life associated with correctly diagnosing coeliac disease in patients with suspected IBS and initiating a gluten-free diet was estimated indirectly by comparing utility estimates for treated and untreated coeliac disease and IBS measured in different populations. This is less reliable than direct utility measurement as it combines estimates from different populations. However, the uncertainty surrounding this parameter has been adequately examined in a sensitivity analysis. The authors stated that they used conservative assumptions to deliberately bias the model against testing for coeliac disease. These included assuming no reduction in resource use or increase in life-expectancy following correct diagnosis of coeliac disease and initiation of treatment. The base case prevalence for coeliac disease in patients with suspected IBS was 3% and was varied from 1 to 5% in a sensitivity analysis.

The probabilistic model estimated that testing with TTG would detect 28 out of 30 cases present in a population of 1000 individuals but testing with a full antibody panel would only detect one further case. The median incremental cost per case detected was $6,700 (interquartile range $4,800 - $9,700) for TTG vs no testing and $12,300 ($8,900 - $17,700) for antibody panel vs no testing. The incremental cost per case detected for antibody panel vs TTG was $167,000

($110,000 - $279,000). The incremental cost per QALY was $11,200 ($7,200 - $17,900) for TTG vs no testing and $20,900 ($13,500 – $34,300) for antibody panel vs no testing. The incremental cost per QALY for antibody panel vs TTG was $287,000 ($99,400 - $675,000). The upfront biopsy strategy resulted in a lower QALY gain and higher costs than TTG testing and was therefore dominated by TTG testing. In the one-way deterministic sensitivity analysis, reducing the prevalence to 1% increased the cost per QALY of TTG vs no testing from $7,400 to $19,900 and decreasing the utility gain associated with treatment from 0.024 to 0.01 increased the cost per QALY to $17,900. This demonstrates that whilst the cost-effectiveness results are sensitive to changes in these parameters, the TTG testing strategy is still cost-effective compared to no testing at the thresholds considered ($50,000 to $100,000 per QALY).

This study provided evidence that TTG testing followed by confirmatory endoscopy with biopsy would be cost-effective in patients with suspected IBS in the US health-care system. We converted the cost per QALY directly from 2003 US$ to 2006 UK£ using Health Care Purchasing Power Parity rates (2003 PPP rates UK/US = 2317/5711, OECD 2006) and Hospital and Community Health Services Pay and Pricing Index (2006/2003 = 241.3/224.8 (Netten 2006) and this gave a cost per QALY for TTG vs no testing of £4,900. This is a crude estimate as it assumes that each component of the total cost has an equal weighting in both counties, which may not be true due to differences in the health care systems between the US and UK. However, the relatively low value of this estimate compared to typical UK thresholds of £20,000 to £30,000 per QALY, suggests that this intervention may also be cost-effective from a UK NHS and PSS perspective.

Spiegel (2004)

This aim of this study was to assess the cost-effectiveness of screening for coeliac disease in patients fulfilling the Rome II criteria for diarrhoea predominant IBS (IBS-D) in the US health care system. A strategy of screening for coeliac disease using serological tests followed by confirmatory endoscopy with biopsy was compared with a strategy of initiating IBS therapy without screening for coeliac disease. This was done using a decision tree to estimate the number of patients receiving appropriate therapy for either IBS or coeliac disease, the number of missed coeliac disease diagnoses and the number of patients for whom IBS treatment was delayed due to coeliac disease testing. It was assumed that 1 in 4 clinicians eventually test for coeliac disease in patients who do not respond to empiric IBS treatment, resulting in an average diagnostic delay of 6 months. A Markov model was then used to estimate transitions between states of symptomatic improvement and remission once patients have begun treatment for either IBS or coeliac disease. The analysis was based on a generic serological test using data which reflected the range of serological tests available (anti-EMA and anti-TTG IgA antibodies). The results are presented in terms of the cost per additional patient with symptomatic improvement after 10 years. The authors state that the model was deliberately biased in favour of IBS treatment without testing for coeliac disease in order to place the burden of proof for cost-

effectiveness on coeliac testing. This was done by using estimates from the unfavourable end of the range presented in the literature for the following parameters; coeliac disease prevalence, sensitivity and specificity of tests for coeliac disease, rate of coeliac disease testing in patients not responding to empiric IBS therapy, IBS treatment effectiveness and cost. For example, IBS treatment was assumed to be effective in 75% of patients based on the effectiveness of alosetron but the cost of therapy was assumed to be $45 per month which is similar to the cost of loperamide. The model also assumed that 30% of the population with coeliac disease had "latent" or "potential" coeliac disease which would not be detected by small bowel biopsy but would have the potential to benefit from a gluten free diet. It also assumed that 5% of the population with coeliac disease would have concurrent IgA deficiency which would render serological screening for IgA antibodies ineffective. These assumptions were based on limited data but were included to bias the model against testing for coeliac disease and their impact was explored through sensitivity analyses.

The deterministic base case model estimated that testing for coeliac disease resulted in 51.6% of the cohort achieving symptomatic improvement at 10 years, whilst initiating IBS therapy without testing for coeliac disease resulted in 50.9% of the population achieving symptomatic improvement. The incremental cost was $77 per patient resulting in a cost per additional symptomatic improvement after 10 years of $11,000. The probabilistic model resulted in a median cost per symptomatic improvement of $12,983 (95% CI: Dominating to $41,031). The results were sensitive to the prevalence of coeliac disease in the population considered. The cost-effectiveness ratio was under $50,000 when the prevalence was >1% and screening dominated no screening (resulted in more health gain at reduced cost) when the prevalence was over 8%.

These results were difficult to interpret as they were presented for the US health care system and did not provide benefits measured in QALYs. The aim of a Markov model is usually to determine the proportion of time a patient spends in each health state over the duration of the model and to use this to estimate their aggregate health gain over the time-horizon considered. This analysis did not present results in terms of the time spent in the symptom remission state, but instead presented the results in terms of the number of patients in this state at the end of the model, which may not accurately reflect the amount of health benefit accrued over the duration of the model. It was therefore less useful in determining whether testing for coeliac disease is cost-effective compared to no testing than the evidence provided by Mein (2004).

Dubinsky (2002)

This study examined the cost-effectiveness of initial serodiagnostic screening followed by standard invasive testing, compared to standard testing alone in patients presenting with symptoms suggestive of IBD from a third-party payer perspective in the US health care system. The authors state that the population considered by this analysis was patients presenting with

symptoms which did not meet the Rome I criteria for IBS. As the aim of this review is to consider the cost-effectiveness of testing for alternative diagnoses in patients meeting the diagnostic criteria for IBS, following the application of a criterion based reference tool, this study was not directly relevant to the target population. We would expect the patient population meeting the diagnostic criterion for IBS to have a lower prevalence of IBD. As the analysis considered a wide range of prevalence values (5% to 75%) in a sensitivity analysis, the results for the lower end of this prevalence range were considered to have some relevance to the target population.

The decision analytic model considered six alternative diagnostic strategies. Two levels of serodiagnostic screening were evaluated. In the primary screening strategy (PR 1) patients received a primary assay followed by a gold standard invasive diagnostic test if the primary screen was positive or if a negative primary screen was followed by persistent symptoms. In the sequential screening strategy (SS 1) a positive primary assay was followed by a confirmatory assay and if this was positive it was followed by a gold standard test. Negative results followed by persistent symptoms were investigated using the gold standard test as in the primary screen. These were compared to gold standard testing upfront (GS 1) which consisted of colonoscopy with biopsies and histological examination as well as a barium upper GI series and small bowel follow-through. Three additional strategies were also considered in which the first three strategies were extended to include a second gold standard test in patients with persistent symptoms following the first gold standard test (PS 2, SS 2 and GS 2). The proportion of patients returning with persistent symptoms due to IBS not meeting the Rome I criteria or other causes of symptoms was assumed to be 50% based on expert opinion and varied from 0 to 100% in a sensitivity analysis. A decision tree model was used to estimate the accuracy and cost of each of the six strategies. No costs or health benefits following diagnosis were estimated. An incremental cost-effectiveness analysis was also presented which compared the relative cost-effectiveness of the six competing strategies.

In the basecase model all of the serodiagnostic strategies had lower costs and higher diagnostic accuracy than the gold standard strategies. The SS 1 strategy had the lowest cost and a diagnostic accuracy of 96.95%. The SS 2 strategy cost $20.30 more per patient but had a slightly higher diagnostic accuracy of 97.90% resulting in a cost per % increase in accuracy of $2,137. The SS 1 strategy dominated all other strategies by having higher accuracy and lower cost and also resulted in the lowest number of invasive procedures out of all six strategies (610 for SS 1 vs 1000 for GS 1 and 1010 for GS 2). In the cost sensitivity analysis, standard invasive testing was more cost-effective when the costs of testing were varied outside of the plausible range considered by the sensitivity analysis.

Standard invasive testing was more cost-effective when the prevalence of IBD was varied to >76%, or when the proportion of patients with persistent symptoms was varied to over 89%.

These results suggest that serodiagnostic screening for IBD in patients with “atypical” IBS symptoms would be less costly and more effective than immediate gold standard invasive testing. These results are based on cost data from the US and the conclusions may be different in a UK analysis if the relative costs of invasive and non-invasive testing are significantly different. Whilst these results apply to patients with “atypical” IBS who do not meet the Rome I criteria, the sensitivity analyses carried out demonstrate that they will apply equally to groups with lower prevalence rates of IBD, provided that less than 89% of patients are given the gold standard test after returning with persistent symptoms following a negative serodiagnostic test. This study does not address whether further testing in patients returning with persistent symptoms is beneficial but assumes that this occurs in practice regardless. This number may be higher or lower in the group meeting the diagnostic criteria for IBS depending on the confidence placed on the positive diagnosis. This study did not address whether these strategies for diagnosing IBD are cost-effective compared to a strategy of initiating IBS treatment, following a positive IBS diagnosis, without excluding IBD. It therefore did not demonstrate the cost-effectiveness of serological testing for IBD in patients meeting diagnostic criteria for IBS.

Suleiman (2001)

The aim of this study was to assess the incremental cost-effectiveness of endoscopic procedures in the work-up for IBS. It did not consider the incremental cost-effectiveness of a specific test for an alternative diagnosis in patients with IBS, but considered the increase in diagnostic probability achieved by using various sequences of tests to exclude alternative diagnoses. These tests included; hydrogen breath test to exclude lactase deficiency or bacterial overgrowth, flexible sigmoidoscopy and colonoscopy to exclude inflammatory colitis, diverticular disease and colon cancer, and small bowel follow-through to exclude small bowel cancer. A decision tree was used to estimate the probability of IBS in the remaining population following each diagnostic test. Various sequences of tests were considered but each began with a clinical history, physical examination and laboratory tests. The costs of further testing following false positive tests and the costs and health impact of delayed diagnosis of IBS or alternative diagnoses were not considered. The authors presented incremental cost-effectiveness ratios (ICERs) for flexible sigmoidoscopy and colonoscopy in terms of the incremental cost per 1% increase in IBS probability, but these figures considered the cost in the individual and did not take into account the number of patients who would be tested at each stage of the diagnostic sequence. The authors also presented average cost-effectiveness ratios, “ACERs” which gave the cost of the whole diagnostic sequence in a cohort of patients divided by the number of correct diagnoses.

The model demonstrated that lower ACERs are achieved by using flexible sigmoidoscopy after rather than before hydrogen breath testing and small bowel follow-through. The same was found for colonoscopy in the absence of a previous flexible sigmoidoscopy. The results demonstrated that carrying out colonoscopoy without flexible sigmoidoscopy at the end of the diagnostic

sequence would result in a lower ACER than carrying out colonosocopy following flexible sigmoidoscopy at the end of the diagnostic sequence.

The relevance of these results to this review was limited as the study did not consider the incremental cost-effectiveness of testing for a specific alternative diagnosis in patients meeting IBS diagnostic criterion. However, it did demonstrate that the cost of diagnostic testing can be reduced by using more costly interventions at the end of a diagnostic sequence without changing the number of correct diagnoses. The clinical outcomes did not vary when the ordering of tests was varied due to the assumption that each test is independent of the next. In practice this may not be strictly true and there may be some dependence resulting in slightly different clinical outcomes depending on the test sequencing. However, it is still likely that lower costs would be achieved in practice by using more costly invasive investigations at the end of the diagnostic sequence as this would reduce the number of people who require these invasive tests. This would also minimise adverse health outcomes due to complications.

Summary

There was some relevant published literature concerning the cost-effectiveness of screening for coeliac disease in patients with suspected IBS. The study by Mein (2004) provided a cost per QALY for coeliac testing vs no testing from a US perspective. Whilst this could not be applied directly to the population under consideration, due to differences in the health care systems between the US and the UK, the low cost per QALY suggested that this intervention may also be cost-effective from a NHS and PSS perspective. The studies by Dubinsky (2002) and Suleiman (2001) did not consider directly whether further diagnostic testing would be cost-effective in patients meeting diagnostic criteria for IBS compared to no further diagnostic testing. They did provide some evidence that where diagnostic testing does take place, it is cost-effective to use less costly and less invasive tests first in the diagnostic sequence with positive results confirmed by standard invasive testing compared to invasive testing early in the diagnostic sequence.

Having considered the evidence on the clinical utility of diagnostic tests in patients meeting IBS diagnostic criteria, the GDG decided that there was insufficient evidence of clinical utility to warrant further economic analysis on the cost-effectiveness of diagnostic testing, except for serological testing for coeliac disease.

Cost-effectiveness of screening for coeliac disease in patients meeting IBS diagnostic criteria – adaptation of a published economic evaluation

Further analysis was carried out to adapt the cost-effectiveness estimate provided by Mein (2004) to make it more applicable to the NHS in England and Wales. UK specific data was obtained for the prevalence of undiagnosed coeliac disease, diagnosis costs, and HRQoL and ongoing resource use for individuals with IBS. A discounting rate of 3.5% was applied to both

costs and QALYs in line with the NICE reference case for cost-effectiveness analysis (NICE 2007). Mein (2004) did not consider the additional cost of gluten-free foods in their analysis, but as gluten-free foods can be prescribed through the NHS this cost was also considered in our analysis. Mein (2004) did not allow for any increased life-expectancy that may result from adherence to a gluten-free diet in patients diagnosed with coeliac disease. This was considered to be overly conservative as one of the main aims of adherence to a gluten-free diet in coeliac disease is to reduce the risk of malignant diseases associated with coeliac disease such as Non-Hodgkin Lymphoma (West 2004). The model was therefore adapted to include an estimated survival difference between patients with diagnosed and undiagnosed coeliac disease. The economic model reported by Mein (2004) compared serological testing for IgA tissue transglutaminase (TTG) antibodies against a strategy of no testing. However IgA EMA testing is more commonly used in the UK than IgA TTG so this was used in the UK adaptation and TTG was considered in a sensitivity analysis.

The prevalence of undiagnosed coeliac disease in patients meeting IBS diagnostic criteria was taken from a cross-sectional study conducted in a UK primary care setting (Saunders 2003). The study population was randomly sampled from all adults entering the GP premises on study days. A subgroup of individuals meeting the ROME II diagnostic criteria for IBS was identified. The prevalence of undiagnosed coeliac disease was 3.3% (4/123) in individuals who fulfilled the ROME II criteria for IBS. This estimate was used in the model as the prevalence of undiagnosed coeliac disease in patients meeting IBS diagnostic criteria. The prevalence from the primary care sample as a whole (1%) was used in a sensitivity analysis as the expected lower limit for the prevalence in the IBS population.

Sensitivity and specificity values for IgA EMA were taken from a published health technology assessment which included a systematic review of autoantibody testing in children with type I diabetes (Dretzke 2004). This systematic review included studies carried out in symptomatic populations or populations at a higher risk of developing coeliac disease but not exclusively type I diabetes. The sensitivity and specificity estimates used in the model were the Q values (overall best test performance with equal sensitivity and specificity) from the well-described studies as given in Table 17 of Dretzke (2004).

The NHS cost of an IgA EMA antibody test was also taken from Dretzke (2004). The cost of esophagogastroduodenoscopy (EGD) with biopsy to confirm coeliac disease was taken from the NHS references costs (2005-06) for day case endoscopic procedures on the stomach or duodenum (Department of Health 2006). The cost for an EGD with complications was assumed to be equal to the NHS reference cost for the same procedure as a non-elective in patient (average length of stay of 1 day). The cost of care for IBS was taken from a study by Akehurst (2002) which estimated the NHS costs for IBS patients and matched controls. As in the cost-effectiveness analysis by Mein (2004), it was assumed that the NHS costs of managing coeliac

disease are equal to the costs of managing IBS except that there is the additional cost of gluten free foods on prescriptions. This may be an overestimate if IBS-like symptoms are reduced when patients with coeliac disease are established on a gluten-free diet.

The NHS cost of supporting a gluten-free diet by providing foods on prescription was calculated by estimating the total cost of gluten-free foods prescribed by the NHS in England in 2005 (£21.2million) from the Prescription Cost Analysis (NHS Health and Social Care Information Centre 2006) and the number of people diagnosed with coeliac disease based on a population for England of 50.4million and a prevalence for diagnosed coeliac disease of 0.26%.(Fowell 2006). This gave an annual cost of £162 per diagnosed case of coeliac disease.

The health utility of IBS was taken from the study by Akehurst (2002) which estimated health utility for IBS patients using the EQ-5D. Mein (2004) attempted to estimate the utility gain associated with diagnosing coeliac disease in patients with IBS-like symptoms. However, this estimate was considered to be unreliable as it was calculated by comparing utility values for health states estimated in different populations. No direct evidence was available on the utility gain achieved by diagnosing and treating coeliac disease in patients with IBS-like symptoms. O'Leary (2002) found that coeliac patients with IBS-like symptoms had a lower HRQoL than those without symptoms, but these symptoms were equally common in coeliac patients who did and didn't adhere to a gluten-free diet. Casellas (2005) found that recently diagnosed patients who had not started a gluten-free diet had a lower quality of life and a higher prevalence of IBS-like symptoms compared to patients who had been established on a gluten-free diet, but the study design was cross-sectional, so it was not possible to say from this whether the diet itself provided an improvement in quality of life. In the basecase analysis it was assumed that the gluten-free diet did not provide any gain in health utility, so the only benefit was from improved survival. A threshold analysis was carried out to assess the size of health utility gain that would need to be achieved by adherence to a gluten-free diet, in order to give a cost per QALY under £20,000, when assuming that the gluten-free diet does not provide a survival gain.

There is evidence that patients with coeliac disease have a significantly higher than expected mortality (SMR = 2.0, $p<0.0001$) (Corrao 2001) and that mortality risk is significantly increased for patients with a diagnostic delay of over 12 months but is not significantly increased when it is less than this. Mortality is also significantly higher than expected (SMR 2.5, $p<0.0001$) in patients with severe symptoms of malabsorption such as diarrhoea or weight loss but not significantly increased in patients with milder symptoms which may be seemingly unrelated to coeliac disease. There is evidence that survival is significantly reduced in the first 3 years after diagnosis but not beyond. The timing of the observed excess mortality may be due to excess deaths in patients who had extended diagnostic delay and who were not diagnosed until after symptoms had become severe. It may be possible to prevent the excess mortality in patients with IBS–like symptoms by prompt diagnosis through serological screening. In order to estimate

the survival gain associated with prompt diagnosis we have assumed that undiagnosed cases have a reduced survival compared to diagnosed cases. We have taken the survival ratio for coeliac patients compared to the general public and used this to estimate the reduction in mortality avoided by prompt diagnosis. This is equivalent to a relative reduction in cumulative survival of around 2% over the first 3 years of the model for patients with coeliac disease presenting with IBS-like symptoms whose coeliac disease remains undiagnosed. This may have underestimated the survival gain associated with diagnostic testing, as the SMR in the whole coeliac population is lower than in those patients with extended diagnostic delay, but it may also have overestimated the survival gain as prompt diagnosis may not result in a complete reduction of mortality to general population levels. These survival ratios were applied to UK life-tables (Office for National Statistics 2006), assuming a male to female ratio of 1:2, and gave an estimated difference in expected life-years of 1.4 LYs for patients with diagnosed and undiagnosed coeliac disease. Once discounting was applied to the expected survival this difference was reduced to 0.54 discounted LYs. A sensitivity analysis was also carried out using the upper 95% CI of the survival ratio which resulted in a lower estimated survival gain of 0.31 discounted LYs.

The parameter values used in the UK adaptation are summarised in Table 11 alongside those used by Mein (2004) in the US basecase. Univariate sensitivity analysis was used to determine whether the deterministic estimate of cost-effectiveness was sensitive to changes in the UK specific parameters. This included a threshold analysis on the utility gain associated with establishing a gluten-free diet and the cost of prescribing gluten-free foods on the NHS. The parameter values used in the sensitivity analysis are also given in Table 11. All costs were uplifted to 2005/06 values where applicable and the uplifted values are included in Table 11 in italics.

Table 11: Parameters used in Mein (2004) analysis and in the UK adaptation

Parameter	Mein (2004) basecase, (range)	UK basecase	UK range for sensitivity analysis
Age	35 (20-60)	35	20-60
Life-expectancy (IBS or diagnosed coeliac disease)	42.8LYs	45.7LYs	N/A
Prevalence of coeliac disease	3% (1-5%)	3.3%, Saunders (2003)	1%, general population prevalence, Saunders (2003)
IBS utility	0.689 (0.6-0.9)	0.675, Akehurst (2002)	0.636, lower CI from Akehurst (2002)
Utility gain resulting from correct diagnosis of coeliac disease	0.024 (0.01-0.04)	None (conservative)	Threshold analysis at £20,000 per QALY assuming no survival gain
Sensitivity of antibody test	94% (87-97%) IgA TTG	98% IgA EMA, Dretzke (2004)	92%, lower 95% CI, Dretzke (2004)

Specificity of antibody test	95% (87-98%) IgA TTG	98% IgA EMA, Dretzke (2004)	92%, lower 95%CI, Dretzke (2004)
Probability of EGD biopsy complication	0.2% (0.05 – 0.5%)	As for US model	N/A
Probability of death if complication	5% (2-10%)	As for US model	N/A
Cost of IBS care and coeliac care excluding GFD	$450 ($225-$675) *(£196)**	£123, Akehurst (2002) *(£172)**	£221, upper 95%CI, Akehurst (2002) *(£307)**
Cost of antibody test	$68 ($22-$136) IgA TTG *(£30)***	£10 IgA EMA, Dretzke (2004) *(£12)**	£11, upper 95% CI, Dretzke (2004) *(£14)**
Cost EGD with biopsy	$800 ($300-$1800) *(£348)***	£463, Department of Health (2006)	£767 upper limit, Department of Health (2006)
Cost of EGD with complication	$5,700 ($2850-$11400), *(£2482)***	£597, Department of Health (2006)	£1010 upper limit, Department of Health (2006)
Discount rate for costs and QALYs	3%	3.5%, NICE (2007)	0% (undiscounted)
Survival difference between diagnosed and undiagnosed	N/A	0.54 (discounted) Calculated using lifetables and survival ratio from Corrao (2001)	0.31 (discounted) Using upper limit of survival ratio from Corrao (2001)
Cost per annum of gluten-free foods on prescription	N/A	£162 Calculated from prescription data and lower estimate of pop prevalence	Threshold analysis at £20K per QALY

* Uplifted to 05/06 prices using Hospital and Community Health Services Pay and Prices Index, Netten (2006)

** Equivalent cost in UK£ converted from US$ using Health Care Purchasing Power Parity rates and uplifted to 05/06, OECD (2006)

A probabilistic sensitivity analysis (PSA) was carried out to estimate the uncertainty in the cost-effectiveness estimate due to the uncertainty in the model input parameters. We characterised the parameter uncertainty by using a probability distribution to describe each of the parameters, details of which can be found in Appendix H. We sampled randomly from these distributions 1000 times and estimated the model outputs (incremental costs and incremental QALYs) for each set of sampled parameters and used these to estimate the uncertainty surrounding the cost per QALY estimate. We based our PSA on 1000 samples of the parameter distributions. The results are presented as a cost-effectiveness acceptability curve (CEAC) which shows the proportion of samples that resulted in a cost per QALY value below various thresholds. It should be noted that the PSA does not account for uncertainty around the model assumptions and these have been explored separately using univariate sensitivity analysis.

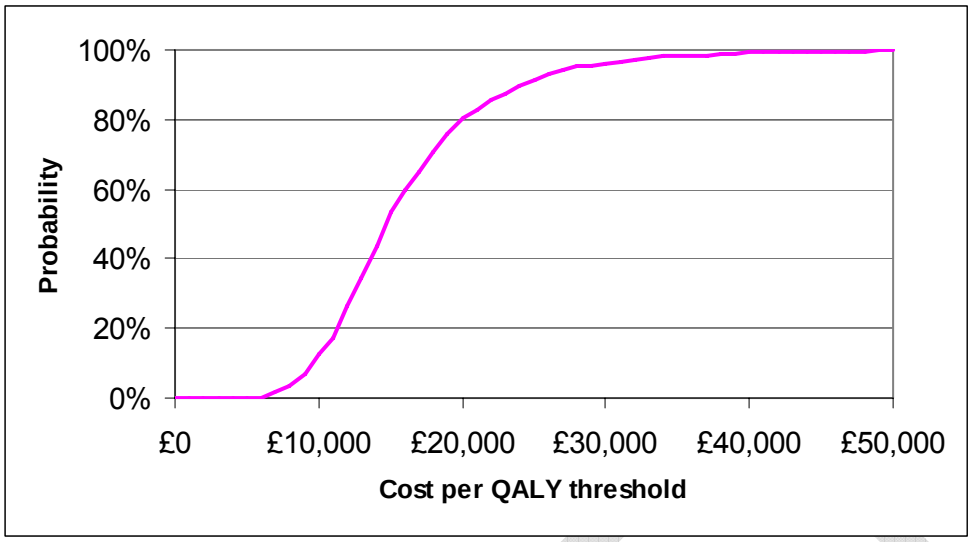
Table 12: Deterministic basecase results for 1000 patients meeting IBS diagnostic criteria. Costs, LYs and QALYs are discounted at 3.5% per annum

Outcome	No testing	Serological testing	Incremental
Number of diagnosed cases (out of 33 prevalent)	0	32	32
Total LYs	22,800	22,817	17.32
Total QALYs	15,390	15,401	11.69
Diagnostic costs	0	£36,300	£36,300
Life-time costs	£3,910,700	£4,069,300	£158,600
Cost per QALY			£13,560

The deterministic results for the UK basecase estimate are given in Table 12. The serological testing strategy identified 32 out of 33 prevalent cases of coeliac disease in the cohort of 1000 patients with IBS symptoms for a diagnosis cost of £36,300, giving a cost per correctly diagnosed case of coeliac disease of £1,122. These diagnoses resulted in an additional 43.4 LYs (undiscounted) over the lifetime of the cohort which is equivalent to 11.69 QALYs (discounted). This was associated with a further £122,300 (discounted) of treatment costs, including gluten-free products for patients diagnosed with coeliac disease, over the lifetime of the cohort. The overall cost per QALY for serological testing compared to no testing was £13,560 for a life-time horizon.

The mean cost per QALY over the 1000 samples carried out for the probabilistic analysis was £14,300. The CEAC in Figure 1 shows the probability that the cost per QALY is under various cost per QALY thresholds given the uncertainty in the parameters used to estimate cost-effectiveness. It shows that the cost per QALY had an 80% probability of being under £20K per QALY and a 96% probability of being under £30K per QALY under the basecase assumptions.

Figure 1: Cost-effectiveness acceptability curve (CEAC) for coeliac disease testing in patients with IBS-like symptoms compared to no testing



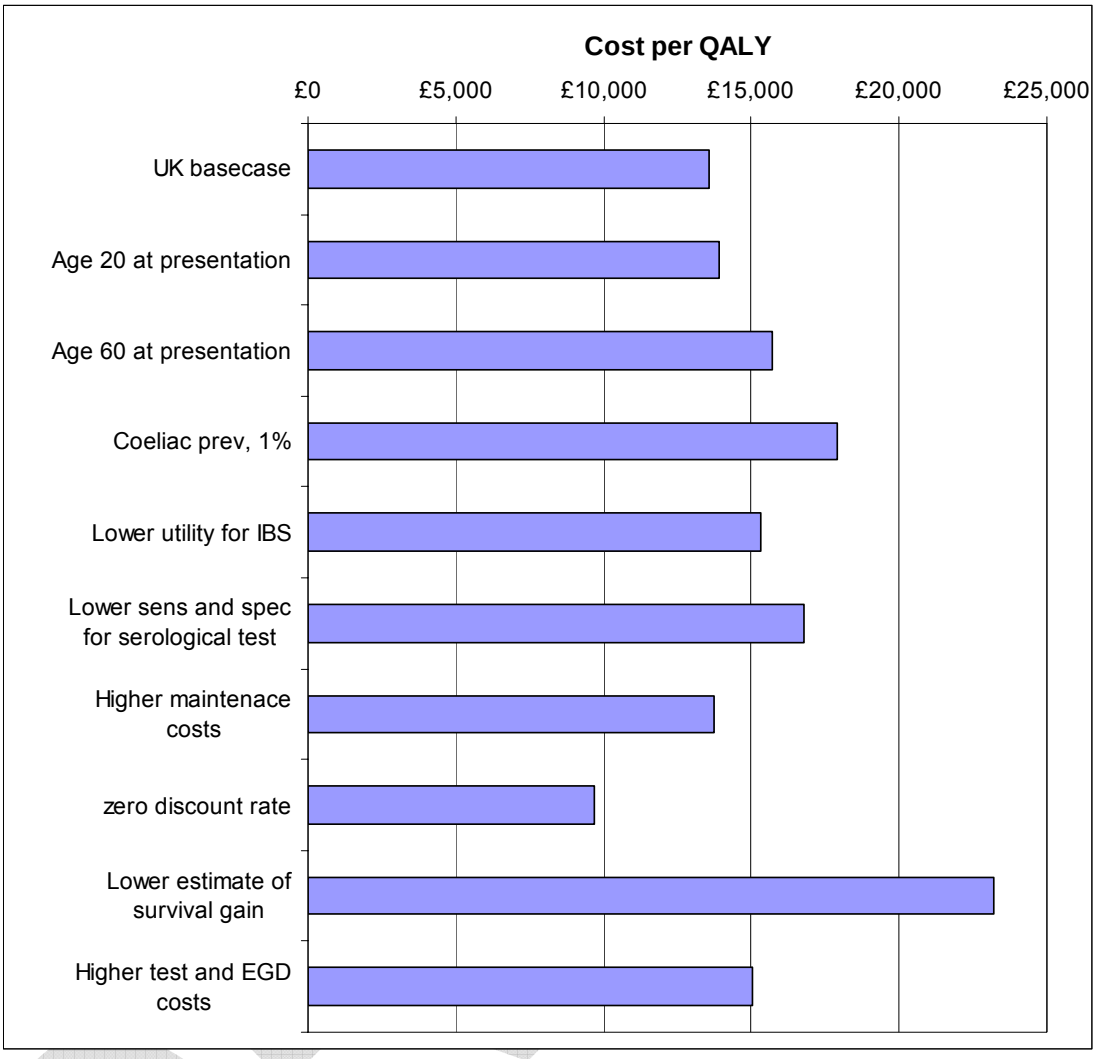
The univariate sensitivity analysis in Figure 2, shows that the cost per QALY estimate was not particularly sensitive to the age of the patient at presentation. This may be because younger patients have a longer life-expectancy, but this increases both their lifetime cost of care and their survival gain from preventing excess mortality. Cost-effectiveness was not significantly impacted by higher testing costs, higher costs for ongoing IBS / coeliac disease management, lower health state utility values for patients with IBS / coeliac disease or lower sensitivity and specificity values for serological testing. Using a zero discounting rate lowered the cost per QALY as the majority of the survival benefit was gained over the long-term whilst the upfront diagnosis costs occurred early in the model.

The cost-effectiveness estimate was sensitive to the survival gain attributed to identifying patients with coeliac disease and establishing them on a gluten-free diet, as this was the only benefit included in the basecase model. Using the lower estimate of survival benefit increased the cost per QALY to £23,000. The threshold analysis on utility gain demonstrated that establishing patients with coeliac disease on a gluten-free diet would need to produce a utility gain of 0.011 in order for the cost per QALY to remain under £20,000, when assuming that there is no survival gain. This is a small utility gain compared to the difference in health utility between IBS patients and matched controls (0.135, Akehurst 2002). These two sensitivity analyses on the survival and QALY gain demonstrated that whilst there is some uncertainty surrounding the expected benefits of identifying individuals with coeliac disease and initiating a gluten-free diet, testing is likely to be cost-effective in patients with IBS-like symptoms, so long as there is a small improvement in quality of life or a small reduction in mortality risk as a result of a correct diagnosis of coeliac disease.

The threshold analysis on the cost of prescribing gluten-free foods shows that up to £263 per patient per annum could be spent on gluten-free foods before the cost per QALY reached the threshold of £20,000. The estimated cost for providing gluten-free foods on prescription is based on the total costs of prescriptions for gluten-free foods in 2005 and an estimate of the prevalence of diagnosed coeliac disease. Using the lower estimate of prevalence from Fowell (2006) gave a higher cost of £234 per patient per annum, which based on our threshold analysis, would still provide a cost per QALY under £20,000.

We have estimated the cost-effectiveness of testing with EMA compared to no testing as this is the test most commonly available to primary care clinicians in the UK. However, TTG is also available in some areas of the NHS. Sensitivity and specificity values were available for TTG (96%, 95% CI of 92%-98%) from the Dretzke (2004) HTA, but a direct cost estimate for TTG was not available. In the economic analysis conducted as part of the HTA, the cost of TTG was assumed equal to the cost of testing for anti-gliadin antibodies (AGA) as these tests use similar techniques. The cost for these tests was estimated to be slightly higher than the cost of EMA. We carried out a sensitivity analysis to see whether testing using TTG would also be cost-effective when using the evidence on test cost and accuracy from the HTA (Dretzke 2004). The slightly lower accuracy for TTG resulted in a slightly lower QALY gain of 11.53 per 1000 people tested, for testing compared to no testing. The slightly higher test cost (£14 compared to £12) resulted in a slightly higher total cost £164,683 per 1000 people tested. The overall cost per QALY for TTG compared to no testing was therefore, £14,283. This suggests that testing with TTG would also be cost-effective compared to no testing. We have not carried out an analysis to consider which antibody test is the more cost-effective test to use as we did not feel that the cost data was sufficiently robust to allow a reliable comparison. In addition, TTG is a relatively new technology and the evidence base may have improved since the searches carried out by Dretzke (2004). There are also other factors that must be taken into account when deciding which tests should be available to primary care clinicians in the NHS, which we have not considered. Therefore we did not feel that our cost-effectiveness analysis was sufficiently robust to recommend the use of either test in preference to the other. However, there is good evidence that using either of these tests is cost-effective compared to no testing in people with IBS.

Figure 2: Univariate sensitivity analysis results for coeliac disease testing in patients with IBS-like symptoms compared to no testing



SUMMARY OF THE EVIDENCE

There is a good evidence base for the application of diagnostic criteria in the diagnosis of people presenting with IBS symptoms, allowing primary care clinicians to make a positive diagnosis with confidence. This has potential to change the current approach to diagnosis, avoiding unnecessary diagnostic tests of limited or in many cases no value. Economic analysis supported by GDG interpretation demonstrates that only four investigations from the included studies for this review are of use to the clinician, in either augmenting their positive diagnosis of IBS or related co morbidity such as Coeliac disease. The cost-effectiveness of two different antibody tests for coeliac disease (EMA and TTG) was considered as both are available within the NHS but access to these tests varies across the NHS. We did not consider which of the two antibody tests is most cost-effective as there was insufficient evidence on the relative cost to make a fair assessment of the incremental cost-effectiveness. The GDG recognised the potential need to

clarify which is the better diagnostic test and determine which test was more cost effective, but it was agreed that this was not a clinical priority for this guideline.

The clinical significance of this review is two fold. Patient experience is often determined by the first exposure to healthcare, and the use of diagnostic criteria offers people who may have IBS the potential for symptom based condition to be diagnosed and managed confidently from the first consultation. The potential for cost saving is a real possibility, by determining the small number of investigations which offer primary care clinicians added benefit to confirm their clinical diagnosis. Identifying tests which are routinely requested but have little or no diagnostic value has real potential for disinvestment within the NHS. The validation of the 'Positive Diagnostic Criteria' is a clear step towards addressing the current variations in diagnostic practice within primary care for people presenting with IBS symptoms.

GDG COMMENTARY

Duration of symptom profile

Having reviewed the evidence and analysed application of the criterion referenced diagnostic tools, duration of symptom profile was recognised to be an important aspect to consider in making recommendations for practice. Three, six and twelve month durations were all discussed and the consensus of the group was that a duration of 6 months was the most appropriate.

Entry filter for use of the diagnostic tool

Primary care clinicians should consider assessment for IBS if the patent reports any of the following symptoms for at least 6 months:

- Change in bowel habit
- Abdominal pain/discomfort and or bloating.

Positive Diagnostic Criteria

For a positive diagnosis to be made, the patient must present with at least 3 of the agreed diagnostic criteria. Language used in the tool was contemporised as follows:

- Pain was modified to include discomfort
- Pain/discomfort was changed to recurrent/episodic, experienced for at least 6 months duration
- Abdominal distension/bloating/abdominal tension was added.

Supportive investigations

Appropriate investigations identified were:

- FBC (full blood count)
- ESR (erythrocyte sedimentation rate)
- CRP (inflammatory marker)
- Antibody testing for Coeliac disease (EMA or TTG)

Follow-up

Once a positive diagnosis has been made, patient follow up is a key aspect of longer term management in managing and evaluating the response to first line therapy interventions. Giving patients the opportunity either to re-attend as required or possibly making regular appointments was discussed. It was agreed by the GDG that follow up should be explicitly stated within the recommendations, and in the absence of any evidence supporting this, consensus agreement would be used.

EVIDENCE STATEMENTS

- 1. There is good evidence to support the use of positive diagnostic criteria in making a diagnosis of Irritable Bowel Syndrome.
-
- 2. There is limited evidence demonstrating that patients who meet symptom based criteria for IBS, are unlikely to have organic gastrointestinal disease. The majority of diagnostic testing in this population adds little or no clinical value, with the exception of serological testing for celiac disease.

3. There are two published studies providing evidence on the cost-effectiveness of screening for Coeliac disease in patients with suspected IBS although only one presented the results in terms of the cost per QALY gained. This study provided a cost per QALY for celiac testing vs no testing from a US perspective. This published decision analytic model was adapted to consider the cost-effectiveness of serological screening for coeliac disease from a UK perspective. This showed that antibody testing (EMA or TTG) is likely to be cost-effective in patients with IBS-like symptoms when taking into account the potential for improved survival or a modest gain in quality of life following diagnosis.

4. There is evidence from published literature that where diagnostic testing does take place, it is cost-effective to use less costly and less invasive tests first in the diagnostic sequence with positive results confirmed by standard invasive testing compared to invasive testing early in the diagnostic sequence.

1

RECOMMENDATION

Primary care clinicians should consider assessment for IBS if the patient reports having had any of the following symptoms for at least 6 months:

- change in bowel habit
- abdominal pain/discomfort
- bloating.

RECOMMENDATION

Patients should be asked if they have any of the following 'red flag' symptoms:

- unintentional and unexplained weight loss
- rectal bleeding
- familial history of bowel cancer.

Patients should be assessed for:

- anaemia
- abdominal masses
- rectal masses.

Identification of any of the above should result in referral into secondary care for further investigation (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27; www.nice.org.uk/CG027).

RECOMMENDATION

For a positive diagnosis of IBS to be made, the person must complain of abdominal pain or discomfort which is either relieved by defaecation, or associated with altered bowel frequency, or altered stool form. This must be accompanied by at least two of the following four symptoms:

- altered stool passage (straining, urgency, incomplete evacuation)
- abdominal bloating, distension, tension or hardness
- symptoms made worse by eating
- passage of mucus.

It should be noted that other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and can be used to support the diagnosis.

RECOMMENDATION

In people who meet the IBS diagnostic criteria, it is recommended that the following tests should be undertaken to exclude other diagnostic possibilities:

- 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]).

RECOMMENDATION

The following tests should not be done to confirm diagnosis in people who meet the IBS diagnostic criteria:

- ultrasound
- rigid/flexible sigmoidoscopy
- colonoscopy; barium enema
- thyroid function test
- faecal ova and parasite test
- faecal occult blood
- hydrogen breath test (for lactose intolerance and bacterial overgrowth).

Appendix 1: Sensitivity and specificity values offered as Odds Ratios

Table 3. Summary of Studies Validating Standard Diagnostic Criteria*

Criteria	Source, y	Scoring Method	Gold Standard	Type of Control	Controls, No.	IBS Patients, No.	Sensitivity, %	Specificity, %	Diagnostic Odds Ratio	Score† Raw	Score† Weighted
Manning	Manning et al,[7] 1978	≥2 of 4 criteria	Clinical	Organic GI	33	32	91	70	22.2	6	10.9
		≥2 of 6 criteria	Clinical	Organic GI	33	32	94	55	12.5		
	Talley et al,[15] 1989	No. of warnings	Clinical	Healthy	154	82	65	86	11.4		
				Organic GI	101	82	58	74	3.9	6	7
				NUD-OGD	134	82	42	85	4.1		
	Talley et al,[16] 1991; Talley,[17] 1992	≥2 of 4 criteria	Clinical	Organic GI	33	32	91	70	NC		
		≥2 of 6 criteria					94	55	NC	6	9.1
		≥3 of 4 criteria					63	85	NC		
		≥3 of 6 criteria					84	76	NC		
	Jeong et al,[18] 1993	≥2 of 6 criteria	Clinical	Non-IBS referrals	114	58	84	54	6.5	7	12.3
		≥3 of 6 criteria					67	70	4.8		
	Rao et al,[19] 1993	≥3 of 6 criteria	Clinical	Healthy	45	65	66	93	27.4		
				Organic GI	23	65	66	61	3.0	8	13.2
				NUD	35	65	66	91	20.9		
	Dogan and Unal,[20] 1996	≥3 of 6 criteria	Clinical	Organic GI	182	165	90	86	57.4	7	12.3
	Thompson,[21] 1997	≥2 of 6 criteria	Clinical	None	NC	156	NC	NC	NC	1	3
Kruis	Kruis et al,[10] 1984	Kruis ≥44	Clinical	Organic GI	209	108	64	99	183.1	8	13.2
	Frigerio et al,[22] 1992	Kruis (male) ≥44			93	15	47	94	12.7		
		Kruis (female) ≥44			108	37	59	95	30.2		
		Kruis (male + female) ≥44			201	52	56	95	21.8		
		Modified Kruis (male) ≥44	Clinical	Organic GI	93	15	67	90	18.7	8	13.2
		Modified Kruis (female) ≥44			108	37	68	92	23.8		
		Modified Kruis (male + female) ≥44			201	52	68	91	21.5		
	Dogan and Unal,[20] 1996	Kruis ≥44	Clinical	Organic GI	182	165	81	91	44.8	7	12.3
Rome	Thompson,[21] 1997	Rome guidelines	Clinical	None	NC	156	NC	NC	NC	1	3
	Vanner et al,[23] 1999	Rome guidelines	Clinical	Non-IBS referrals	52	46	65	100	18.8	6	9

*IBS indicates irritable bowel syndrome; GI, gastrointestinal; NUD-OGD, nonulcer dyspepsia–oesophagogastroduodenoscopy; and NC, not calculable.
†Sums of scores from Table 1.

From Fass et al (2001) Evidence and consensus based practice guidelines for the diagnosis of Irritable Bowel Syndrome.

Appendix 2: Comparison table for Rome Criteria

Table 1. Results of Studies Comparing Gastrointestinal Symptoms in Irritable Bowel Syndrome (IBS) and Organic Diseases

	Rome Criteria (%)							
Study	Abdominal Pain/Discomfort Relieved by Defecation	Abdominal Pain/Discomfort Associated with Change in Stool Frequency	Abdominal Pain/Discomfort Associated with Change in Stool Consistency	Altered Stool Frequency	Altered Stool Form	Altered Stool Passage	Passage of Mucus	Bloating or Feeling of Distension
Manning et al, 1978[4]								
IBS	81	81	81	—	—	59	47	53
Organic	30	30	27			33	21	21
Thompson, 1984[11]								
IBS	82	61	64	—	37	88	60	84
Organic	78	52	48		35	73	71	59
Talley et al, 1990[6]								
IBS	43	38	40	32	57	56	49	43
Organic	33	13	20	15	28	32	20	23
Poynard et al, 1992[7]								
IBS	69	23	42	—	—	78	23	66
Organic	52	20	33			54	28	45
Jeong et al, 1993[9]								
IBS	59	57	59	—	—	78	19	40
Organic	32	26	33			53	10	26
Rao et al, 1993[8]								
IBS	66	35	48	—	—	85	79	22
Organic	44	9	13			65	65	13

Table 2. Number of Patients with Specified Organic Diseases in Studies Comparing Gastrointestinal Symptoms in Irritable Bowel Syndrome and Organic Diseases

			Lower GI Disorders				
Study	No. of Subjects with Organic Disease	Upper GI Disorders	IBD	Other Colitis	Colonic Cancer	Diverticulitis	Other GI Disorders
Manning et al, 1978[4]	33	19	5	—	2	Excluded	7
Thompson, 1984[11]	98	49	49	—	—	—	—
Talley et al, 1990[6]	101	65	6	—	4	—	26
Poynard et al, 1992[7]	94	23	18	4	?*	6	43
Jeong et al, 1993[9†]	87	51			16†		20
Rao et al, 1993[8]	23	—	16	—	7	—	—
Dogan and Unal, 1996[10]	182	140	9	—	16	—	17

* Specific number of upper versus lower gastrointestinal carcinoma not specified in this investigation.
† Type of colonic disease was not specified in this study.

Appendix 3: PATIENT PROFILES FOR SIMULATED GP CONSULTATION

IBS GDG MEETING 30th November- 1st December 2006

Patient 1

Female, age 37yrs, married 2 children

Recurrent abdominal discomfort approximately a week out of every month – sometimes worse pre- menstrual, worse if eats cauliflower or spicy foods
Tired and lethargic – has tried different diets to help energy levels but nothing works
Bowel – change in bowel habits, had diarrhoea after holiday abroad now seems to have constipation followed by diarrhoea.

Patient 2

Male 44 yrs, Divorced, 4 children, 2 ex wives!

Frequent abdominal pain, most days of the week
Bloating worse by end of day with increased flatulence
Constipation, thinks there has been a little rectal bleeding but not sure. Worse since new job – v stressful over last six months. .
Social life diminished because embarrassed to go out, becoming increasingly depressed, worried he may have something serious.

Patient 3

Female 51yrs single,

Diarrhoea on and off for last 2 years
Abdominal pain
Back ache
Nausea
Weight loss
2 x visit to Doctors with urinary symptoms – no UTI but symptoms recur intermittently
describes herself as fed up – not depressed.

Patient 4

Female 24yrs

Altered bowel habit – diarrhoea and constipation – changes all the time feels she never empties her bowels, passes mucus in diarrhoea, pale bulky stools when constipated. Has had 'sensitive tummy' since she was a child

Mother has long term problems with constipation. Abdominal pain better after bowel movement. Some foods make it worse she wonders if she has food allergy – sometimes gets a rash and frequently has mouth ulcers

Drinks a lot of milk when 'off' food