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

February 2008

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   Page

Guideline development group membership and acknowledgements

Mr Jonathan 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
Consultant Dietitian in Gastroenterology, King's College 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.

Behavioural therapy: A generic term to describe forms of psychological therapy based on the concept that the way we think about things affects how we feel and act. Behavioural therapy focuses on thinkingand behaviour and it aims to help people in the ways they act (behaviour).

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 Form Scale: 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.

Cognitive Behavioural Therapy, also called **Cognitive therapy:** A relatively short-term form of psychotherapy based on the concept that the way we think about things affects how we feel emotionally. Cognitive therapy focuses on present thinking, behaviour, and communication rather than on past experiences and is oriented toward problem solving. It aims to help people in the ways they think (cognition) and in the ways they act (behaviour).

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'.

Concordance: An approach first used in the prescribing and taking of medicines but can be applied to many treatments in healthcare. It is an agreement reached after negotiation between a patient and a health care professional that respects the beliefs and wishes of the patient in determining whether, when and how medicines/treatments are to be taken. Although reciprocal, this is an alliance in which the health care professionals recognise the primacy of the patient's decisions about taking the recommended medications.

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).

Counselling: The skilled use of the relationship (between the counsellor and patient) to help the patient develop self-knowledge, self-esteem and the ability to take control of his or her own life. (The British Association of Counselling and Psychotherapy).

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: 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: 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.

Psychological interventions: The treatment of any condition by psychological means. This may utilise insight, persuasion, suggestion, reassurance, and instruction so that patients may see themselves and their problems more realistically and have the desire to cope effectively with them. There are many different psychological interventions, these include psychotherapy, biofeedback, cognitive behavioural therapy, family therapy, hypnotherapy, interpersonal therapy and psychodynamic therapy.

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, KEY RECOMMENDATIONS and IBS ALGORITHM

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 psychological 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

The key recommendations were identified by the GDG as the recommendations that are the priorities for implementation. (The numbering corresponds to the abbreviated (NICE) version of the guideline).

1.1.1.1 Healthcare professionals should consider assessment for IBS if the person reports having had any of the following symptoms for at least 6 months:
- Abdominal pain or discomfort
- Bloating
- Change in bowel habit.

1.1.1.2 All people presenting with possible IBS symptoms should be asked if they have any of the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected):
- Unintentional and unexplained weight loss
- Rectal bleeding
- A family history of bowel or ovarian cancer
- A change in bowel habit to looser and / or more frequent stools persisting for more than 6 weeks in a person aged over 60 years.

1.1.1.3 All people presenting with possible IBS symptoms should be assessed and clinically examined for the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected).
- Anaemia
- Abdominal masses
- Rectal masses
- Inflammatory markers for inflammatory bowel disease.
  ~~If there is significant concern that symptoms may suggest ovarian cancer, a pelvic examination should also be considered.~~

> This recommendation was partially updated in September 2012 in line with more recent guidance. The deleted text now reads:
>
> Measure serum CA125 in primary care in women with symptoms that suggest ovarian cancer in line with 'Ovarian cancer' (NICE clinical guideline 122).

1.1.1.4 A diagnosis of IBS should be considered only if the person has abdominal pain or discomfort that is either relieved by defaecation or associated with altered bowel frequency or stool form. This should be accompanied by at least two of the following four symptoms:

- Altered stool passage (straining, urgency, incomplete evacuation)
- Abdominal bloating (more common in women than men), distension, tension or hardness
- Symptoms made worse by eating
- Passage of mucus.

Other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and may be used to support the diagnosis.

1.1.2.1 In people who meet the IBS diagnostic criteria, the following tests should be undertaken to exclude other diagnoses:

- Full blood count (FBC)
- Erythrocyte sedimentation rate (ESR) or plasma viscosity
- C-reactive protein (CRP)
- Antibody testing for coeliac disease (endomysial antibodies [EMA] or tissue transglutaminase [TTG]).

1.1.2.2 The following tests are not necessary 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 Healthcare professionals should review the fibre intake of people with IBS, adjusting (usually reducing) it while monitoring the effect on symptoms. People with IBS should be discouraged from eating insoluble fibre (for example, bran). If an increase in dietary fibre is advised, it should be soluble fibre such as ispaghula powder or foods high in soluble fibre (for example, oats).

1.2.2.4 People with IBS should be advised how to adjust their doses of laxative or antimotility agent according to the clinical response. The dose should be titrated according to stool consistency, with the aim of achieving a soft, well-formed stool (corresponding to Bristol Stool Form Scale type 4).

1.2.2.5 Healthcare professionals should consider tricyclic antidepressants (TCAs)[**] as second-line treatment for people with IBS if laxatives, loperamide or antispasmodics have not helped. TCAs are primarily used for treatment of depression but are only recommended here for their analgesic effect. Treatment should be started at a low dose (5–10 mg equivalent of amitriptyline), which should be taken once at night and reviewed regularly. The dose may 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. The emergence of any of the 'red flags'during management and follow up should prompt referral for further investigation and/or referral into secondary care.

[**] At the time of publication (February 2008) TCAs did not have UK marketing authorisation for the indications described. Informed consent should be obtained and documented.

IBS Algorithm

Person with any of these symptoms for at least 6 months
Abdominal pain/discomfort, **B**loating, **C**hange in bowel habit.

Patient History and Clinical Examination by GP/Primary Care Clinician

Investigations in PC
- FBC (Anaemia)
- ESR
- CRP (Inflammatory Bowel Disease)
- EMA or TTG (Coeliac Disease)

IBS Positive Diagnosis Criteria

Red Flag Symptoms
- Rectal bleeding
- Unexplained unintentional weight loss
- Family history of bowel/ ovarian cancer
- Late onset (age over 60)

Assess for anaemia, abdominal, pelvic (if appropriate) and rectal masses and inflammatory bowel disease
Immediate referral to secondary care

IBS Management

Should be based on the nature and severity of symptoms and individual or combinations of medication, with lifestyle advice, directed at the predominant symptom/s.

Lifestyle: Diet & Physical Activity

Assess diet; reduce fibre intake; take soluble fibre and consider dietitian referral.
Assess level of physical activity, encourage increased levels of activity.
Patient information resource, with dietary, lifestyle and self help advice

Drug Therapy

Consider single or combination therapies:
Antispasmodics
Antimotility agents (titrate dose)
Laxatives (titrate dose)
2nd line Tricyclics (or SSRIs)

Symptom Control

Effective

Follow up to evaluate response (timescale negotiated between clinician and patient)

Not effective

Continuing symptom profile

More than 12 months duration, consider psychological interventions:
Hypnotherapy, Psychological therapy, CBT

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. Healthcare professionals need to be sensitive to and take into consideration cultural, ethnic and communication needs of people for whom English is not a first language or who may have cognitive and/or behavioural disabilities. Appropriate action should be taken to facilitate effective consultation.

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 (to identify possible inflammatory bowel disease): 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 or

ovarian cancer is suspected on clinical examination, or if there is a family history of GI or ovarian 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

Confirming a diagnosis of IBS is a crucial part of this guideline. The primary aim should be to establish the person's symptom profile, with abdominal pain or discomfort being a key symptom. It is also necessary to establish the quantity and quality of the pain or discomfort, identify its site (which can be anywhere in the abdomen) and whether this varies. This distinguishes IBS from cancer-related pain, which typically has a fixed site.

When establishing bowel habit, showing people the Bristol Stool Form Scale (see Appendix D of the NICE version and Appendix I of this document) may help them with description, particularly when determining quality and quantity of stool. People presenting with IBS symptoms commonly report incomplete evacuation/rectal hypersensitivity, as well as urgency, which is increased in diarrhoea-predominant IBS. About 20% of people experiencing faecal incontinence disclose this only if asked. People who present with symptoms of IBS should be asked open questions to establish the presence of such symptoms (for example, 'tell me about how your symptoms affect aspects of your daily life, such as leaving the house'). Healthcare professionals should be sensitive to the cultural, ethnic and communication needs of people for whom English is not a first language or who may have cognitive and/or behavioural problems or disabilities. These factors should be taken into consideration to facilitate effective consultation.

1.1.1 Initial assessment

1.1.1.1 Healthcare professionals should consider assessment for IBS if the person reports having had any of the following symptoms for at least 6 months:

- Abdominal pain or discomfort
- Bloating
- Change in bowel habit.

1.1.1.2 All people presenting with possible IBS symptoms should be asked if they have any of the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected).

- Unintentional and unexplained weight loss
- Rectal bleeding
- A family history of bowel or ovarian cancer
- A change in bowel habit to looser and / or more frequent stools persisting for more than 6 weeks in a person aged over 60 years.

1.1.1.3 All people presenting with possible IBS symptoms should be assessed and clinically examined for the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected).

- Anaemia
- Abdominal masses
- Rectal masses
- Inflammatory markers for inflammatory bowel disease.

~~If there is significant concern that symptoms may suggest ovarian cancer, a pelvic examination should also be considered~~.

1.1.1.4 A diagnosis of IBS should be considered only if the person has abdominal pain or discomfort that is either relieved by defaecation or associated with altered bowel frequency or stool form. This should be accompanied by at least two of the following four symptoms:

- Altered stool passage (straining, urgency, incomplete evacuation)
- Abdominal bloating (more common in women than men), distension, tension or hardness
- Symptoms made worse by eating
- Passage of mucus.

Other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and may be used to support the diagnosis.

1.1.2 Diagnostic tests

1.1.2.1 In people who meet the IBS diagnostic criteria, the following tests should be undertaken to exclude other diagnoses:

- Full blood count (FBC)
- Erythrocyte sedimentation rate (ESR) or plasma viscosity
- C-reactive protein (CRP)
- Antibody testing for coeliac disease (endomysial antibodies [EMA] or tissue transglutaminase [TTG]).

1.1.2.2 The following tests are not necessary 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 Healthcare professionals should encourage people with IBS to identify and make the most of their available leisure time and to create relaxation time.

1.2.1.3 Healthcare professionals should assess the physical activity levels of people with IBS, ideally using the General Practice Physical Activity Questionnaire (GPPAQ; see Appendix J). People with low activity levels should be given brief advice and counselling to encourage them to increase their activity levels.

1.2.1.4 Diet and nutrition should be assessed for people with IBS and the following general advice given.

- Have regular meals and take time to eat.
- Avoid missing meals or leaving long gaps between eating.
- Drink at least eight cups of fluid per day, especially water or other non-caffeinated drinks, for example herbal teas.
- Restrict tea and coffee to three cups per day.
- Reduce intake of alcohol and fizzy drinks.
- It may be helpful to limit intake of high-fibre food (such as wholemeal or high-fibre flour and breads, cereals high in bran, and whole grains such as brown rice).
- Reduce intake of 'resistant starch' (starch that resists digestion in the small intestine and reaches the colon intact), which is often found in processed or re-cooked foods.
- Limit fresh fruit to three portions per day (a portion should be approximately 80g).
- People with diarrhoea should avoid sorbitol, an artificial sweetener found in sugar-free sweets (including chewing gum) and drinks, and in 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 Healthcare professionals should review the fibre intake of people with IBS, adjusting (usually reducing) it while monitoring the effect on symptoms. People with IBS should be discouraged from eating insoluble fibre (for example, bran). If an increase in dietary fibre is advised, it should be soluble fibre such as ispaghula powder or foods high in soluble fibre (for example, oats).

1.2.1.6 People with IBS who choose to try probiotics should be advised to take the product for at least 4 weeks while monitoring the effect. Probiotics should be taken at the dose recommended by the manufacturer.

1.2.1.7 Healthcare professionals should discourage the use of aloe vera in the treatment of IBS.

1.2.1.8 If diet continues to be considered a major factor in a person's symptoms and they are following general lifestyle/dietary advice, they should be referred to a dietitian for advice and treatment, including single food avoidance and exclusion diets. Such advice should only be given by a dietitian.

1.2.2 Pharmacological therapy

Decisions about pharmacological management should be based on the nature and severity of symptoms. The recommendations made below assume that the choice of single or combination medication is determined by the predominant symptom/s.

1.2.2.1 Healthcare professionals should consider prescribing antispasmodic agents for people with IBS. These should 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 people should be discouraged from taking lactulose.

1.2.2.3 Loperamide should be the first choice of antimotility agent for diarrhoea in people with IBS.

1.2.2.4 People with IBS should be advised how to adjust their doses of laxative or antimotility agent according to the clinical response. The dose should be titrated according to stool consistency, with the aim of achieving a soft, well-formed stool (corresponding to Bristol Stool Form Scale type 4).

1.2.2.5 Healthcare professionals should consider tricyclic antidepressants (TCAs)[**] as second-line treatment for people with IBS if laxatives, loperamide or antispasmodics have not helped. TCAs are primarily used for treatment of depression but are only recommended here for their analgesic effect. Treatment should be started at a low dose (5–10 mg equivalent of amitriptyline), which should be taken once at night and reviewed regularly. The dose may be increased, but does not usually need to exceed 30 mg.

[**] At the time of publication (February 2008) TCAs did not have UK marketing authorisation for the indications described. Informed consent should be obtained and documented.

1.2.2.6 Selective serotonin reuptake inhibitors (SSRIs)*** should be considered for people with IBS only if TCAs have been shown to be ineffective.

1.2.2.7 Healthcare professionals should take into account the possible side effects when prescribing TCAs or SSRIs. After prescribing either of these drugs for the first time at low doses for the treatment of pain or discomfort in IBS, the person should be followed up after 4 weeks and then at 6–12 monthly intervals thereafter.

1.2.3 Psychological interventions

1.2.3.1 Referral for psychological interventions (cognitive behavioural therapy [CBT], hypnotherapy and/or psychological therapy) should be considered for people with IBS who do not respond to pharmacological treatments after 12 months and who develop a continuing symptom profile (described as refractory IBS).

1.2.4 Complementary and alternative medicine (CAM)

1.2.4.1 The use of acupuncture should not be encouraged for the treatment of IBS.

1.2.4.2 The use of reflexology should not be encouraged for the treatment of IBS.

1.2.5 Follow-up

1.2.5.1 Follow-up should be agreed between the healthcare professional and the person with IBS, based on the response of the person's symptoms to interventions. This should form part of the annual patient review. The emergence of any 'red flag' symptoms during management and follow-up should prompt further investigation and/or referral to secondary care.

*** At the time of publication (February 2008) SSRIs did not have UK marketing authorisation for the indication described. Informed consent should be obtained and documented.

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. The GDG declarations are recorded in Appendix K.

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 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

- o Blinding may be not be possible depending on the nature of the interventions
- o 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:
  - o Studies with at least 20% of data missing from any group were considered to be potentially biased
  - o Those with moderate loss to follow up (20 to 50%) were considered in sensitivity analyses
  - o Those with 50% or more patients missing from any one group were regarded as flawed and not analysed further
- Intention to treat analysis:
  - o 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
  - o 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:
  - o Studies in which the washout period was shorter than the minimum were not included, as were studies with no washout or none stated
  - o 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:
  - o 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?):
  o Direct comparisons of intervention with placebo/none
- Is there a dose effect?
  o Direct dose comparisons
  o Subgroup analyses (across trials) of intervention versus placebo, by dose
- Is the duration of treatment important?
  o Direct comparisons of different durations
  o Subgroup analyses of intervention versus placebo, by duration
- Is the intervention better than another treatment?
  o Direct comparisons
  o Subgroup analyses of interventions versus placebo, by type of intervention
- Is the intervention useful as an adjunct to another treatment?
  o Direct comparisons (A + B versus B alone)
- Are there (pre-specified) subgroups of patients for whom the intervention is more effective?
  o E.g. type of IBS (constipation, diarrhoea, alternating); severity of IBS
  o 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 psychological interventions, 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

---

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

studies, we 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 psychological interventions 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 psychological interventions. 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

Antispasmodic therapy initiated
Follow-up appointment booked for 1 month later

1 month later

GP assesses if patient has responded to treatment

Responded

Didn't respond

Patient continues on treatment and follow-up booked for 6 months after first antispasmodic therapy initiated

Is there another effective antispasmodic available?

No

Yes

Antispasmodic therapy discontinued

Patient is switched to another antispasmodic and follow-up appointment booked for 1 month later

1 month later

6 months after first antispasmodic

GP assesses whether antispasmodic therapy is still appropriate for symptom profile

6 months later

Still appropriate

No longer appropriate

Antispasmodic therapy continued for further 6 months

Antispasmodic therapy discontinued

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). Our method for estimating QALY gain is quite crude as it assumes that all patients who experience a therapeutic response have the same increase in HRQoL and it does not distinguish between varying degrees of improvement in HRQoL. Where possible, we have used an improvement in global symptoms to determine whether there has been a therapeutic response to treatment in order to prevent bias being introduced by the use of different outcomes for different interventions.
- Given the limitations of the approach used to estimate QALY gains, a threshold analysis was also 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 psychological interventions 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 psychological interventions*
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)

* Psychological interventions 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 relationship between published clinical and cost effective evidence and recommendations 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% representative opinion of the GDG relating to Round 1 voting.

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

Table 1.

IBS key recomendations vote - round 1
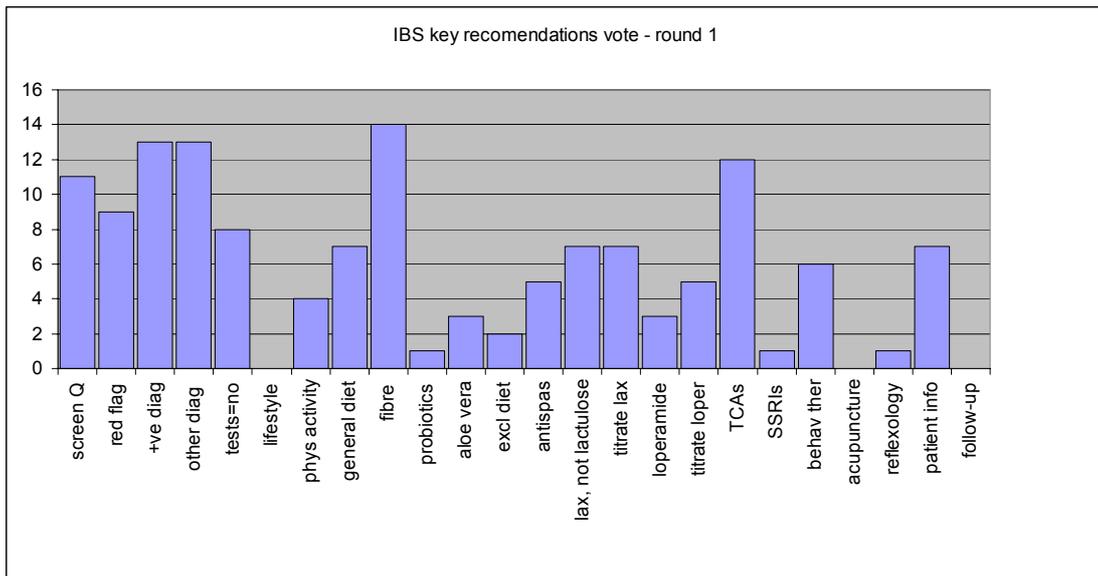
All recommendations with more than 50% of the vote were selected automatically as key recommendations; i.e. screening question, red flags, positive diagnosis, other diagnostic tests, tests that should not be done, fibre, and tricyclics. This gave seven recommendations, but the next highest results gave 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.

2nd round voting

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 psychological interventions 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:
    - o Relief by defaecation
    - o Onset associated with a change in frequency of stool
    - o 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. Healthcare professionals need to be sensitive to and take into consideration cultural, ethnic and communication needs of people for whom English is not a first language or who may have cognitive and/or behavioural disabilities. Appropriate action should be taken to facilitate effective consultation.

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 a positive predictive value (PPV) of 37% (Dawes et al, 2005). 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) of 63%. 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. This is extremely useful to clinicians in trying to establish risk and/or probability of a disease in a particular individual, enabling them to articulate this to the person seeking consultation.

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. Studies included in the table illustrate how research has validated the use of these tools, where the criteria used is matched against a clinical reference standard (in all studies this was an expert gastroenterologist) by using specificity, sensitivity and positive predictive value data. 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).

[1]**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?

---

[1] In the absence of a gold standard, the reference standard was expert gastroenterologist diagnosis.

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 (reference standard was expert gastroenterologist diagnosis)

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 = 0.190 p=>0.05	N=347	Se= 80% Sp= 97%	PPV=96%
Manning (3/6)	Jeong et al 1990	N=172	Se= 67% Sp= 70%	
	Smith 1992 Manning ≥ ¾	N=109	Se= 63% Sp= 85%	
Manning (>3/6)	Talley et al 1990	N= ??	Se= 84% Sp= 76%	
	Kruis et al 1984	N=479	Se= 64% Sp= 99%	94%
Rome	Saito et al 2003a, USA	Prevalence Cohort study	Prevalence rates by criteria:	

		1st survey 1987 n=1121 2nd survey 1989 3rd survey 1992 n=892 response n=643 (72%)	Rome (1989) 27.6 per 100 (95%CI:23.6-31.5) Rome (1990) 5.1 per 100 (95%CI:3.2-7.1)	
	Vanner et al 1999	N=384 (retrospective) N=95	Se= 63% Sp= 100%	98%
Rome I	Saito et al 2003b, USA		Rome I (1992) 6.8 per100 (95%CI 4.7-8.9)	Good agreement between Rome I and II (>95% Kappa >0.68)
	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	Se=83% Sp= not given Se/Sp = not given	
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) 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.
	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	
BDQ (Talley et al) Validated q'aire for	Bijkerk et al 2003, Netherlands	N= 99	All patients had diagnosis of IBS but only18% (n=14) met Rome II	

identifying IBS			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). 168/229 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 (Table 1) 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 Form Scale 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 by using a clinical reference standard (expert gastroenterologist diagnosis), 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. It was not considered necessary to carry out further economic analysis to estimate the cost-effectiveness of routine laboratory investigations, such as FBC, ESR and CRP. They are unlikely to result in a significant cost burden for the NHS and it would be difficult to estimate their cost-effectiveness due to their non-specific nature which means that an abnormal result may be due to a variety of causes. Therefore, further economic modelling focused on the cost-effectiveness of 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
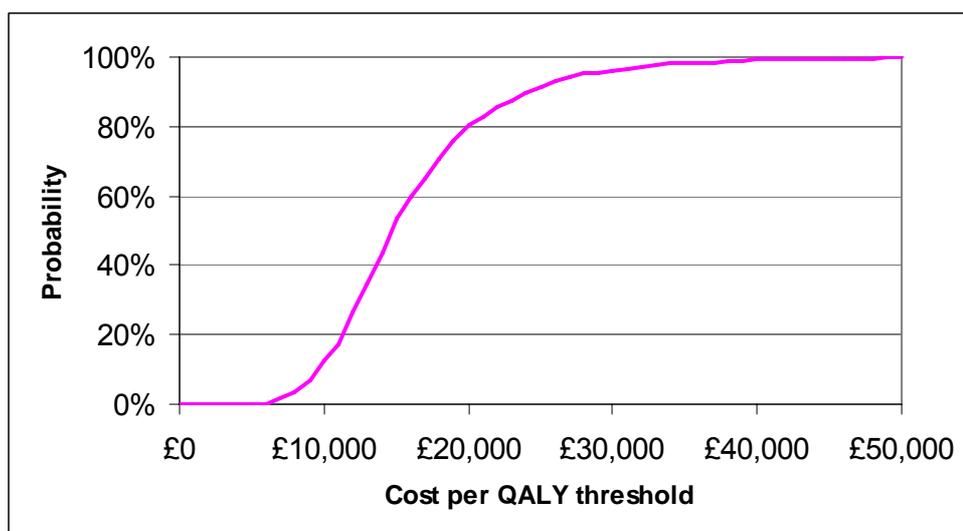
The univariate sensitivity analysis in Figure 2, shows that the cost per QALY estimate was not particularly sensitive to the age of the patient at presentation. This may be because younger patients have a longer life-expectancy, but this increases both their lifetime cost of care and their survival gain 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
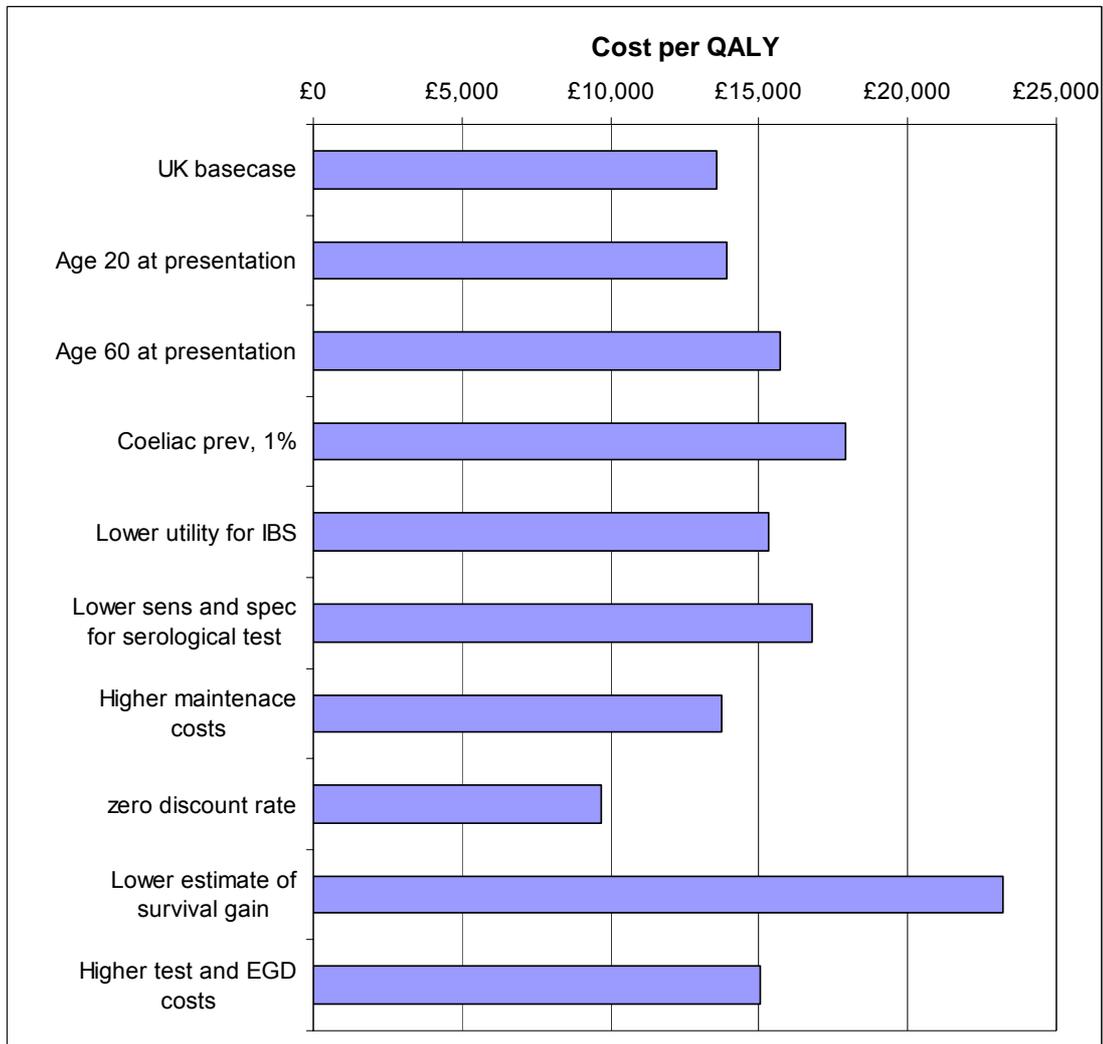
SUMMARY OF THE EVIDENCE

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 is illustrated in Table 1. 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.

There was limited evidence on the clinical utility of routine laboratory investigations such as a FBC, ESR or CRP. However, GDG consensus was that these low cost tests were clinically useful in supporting a positive diagnosis and were unlikely to significantly increase the cost burden to the NHS.

For all other diagnostic investigations (ultrasound, sigmoidoscopy, colonosocopy, barium enema, thyroid function test, faecal ova and parasite tests, faecal occult blood test, hydrogen breath test) the GDG felt that there was insufficient evidence of their clinical utility to support the routine use of these tests in individuals meeting the IBS diagnostic criteria who did not have any red flag symptoms. They also felt that repeated testing could undermine confidence in the positive diagnostic approach. Whilst the cost-effectiveness of these tests was not explicitly estimated, the GDG felt that they were unlikely to be cost-effective given the lack of evidence of clinical utility in this population. The GDG recognized the importance of early and appropriate investigations in individuals with red flag symptoms. They were also mindful that the recommendations should not discourage the use of investigations in patients who do not meet the IBS diagnostic criteria or who have symptoms suggestive of organic disease.

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. Contemporised language previously used in the Manning criteria aligns closely to ROME III criteria. The diagnostic criteria recommended in this guideline implicitly acknowledges their contribution, in strengthening a contemporary IBS diagnostic approach.

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.

RECOMMENDATION

Healthcare professionals should consider assessment for IBS if the person reports having had any of the following symptoms for at least 6 months:

- Abdominal pain or discomfort
- Bloating
- Change in bowel habit.

RECOMMENDATION

All people presenting with possible IBS symptoms should be asked if they have any of the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected).

- Unintentional and unexplained weight loss
- Rectal bleeding
- A family history of bowel or ovarian cancer
- A change in bowel habit to looser and / or more frequent stools persisting for more than 6 weeks in a person aged over 60 years.

RECOMMENDATION

All people presenting with possible IBS symptoms should be assessed and clinically examined for the following "red flag" indicators and should be referred to secondary care for further investigation if any are present (see 'Referral guidelines for suspected cancer', NICE clinical guideline 27, for detailed referral criteria where cancer is suspected).

- Anaemia
- Abdominal masses
- Rectal masses

- Inflammatory markers for inflammatory bowel disease.

~~If there is significant concern that symptoms may suggest ovarian cancer, a pelvic examination should also be considered.~~

RECOMMENDATION

A diagnosis of IBS should be considered only if the person has abdominal pain or discomfort that is either relieved by defaecation or associated with altered bowel frequency or stool form. This should be accompanied by at least two of the following four symptoms:

- Altered stool passage (straining, urgency, incomplete evacuation)
- Abdominal bloating (more common in women than men), distension, tension or hardness
- Symptoms made worse by eating
- Passage of mucus.

Other features such as lethargy, nausea, backache and bladder symptoms are common in people with IBS, and may be used to support the diagnosis.

RECOMMENDATION

In people who meet the IBS diagnostic criteria, the following tests should be undertaken to exclude other diagnoses:

- Full blood count (FBC)
- Erythrocyte sedimentation rate (ESR) or plasma viscosity
- C-reactive protein (CRP)
- Antibody testing for coeliac disease (endomysial antibodies [EMA] or tissue transglutaminase [TTG]).

RECOMMENDATION

The following tests are not necessary 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).

RECOMMENDATION

Follow-up should be agreed between the healthcare professional and the person with IBS, based on the response of the person's symptoms to interventions. This should form part of

the annual patient review. The emergence of any 'red flag' symptoms during management and follow-up should prompt further investigation and/or referral to secondary care.

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	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	8	13.2
		Modified Kruis (male) ≥44	Clinical	Organic GI	93	15	67	90	18.7		
		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

Study	Rome Criteria (%)							
	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

Study	No. of Subjects with Organic Disease	Upper GI Disorders	Lower GI Disorders					Other GI Disorders
			IBD	Other Colitis	Colonic Cancer	Diverticulitis	
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, aged 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, aged 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 – very 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, aged 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, aged 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.

7 DIET AND LIFESTYLE

<table>
<tr><td>

Clinical Questions

1. What associations are there between diet and IBS?

2. What dietary interventions improve symptoms/quality of life?

3. Does Aloe Vera have a role in managing symptoms?

4. What associations are there between physical activity and IBS?

5. Does physical activity improve IBS or related symptoms?

</td></tr>
</table>

BACKGROUND

Diet and lifestyle may be factors that trigger or exacerbate symptoms of IBS so they are factors that need to be given due consideration both at the initial and later stages of management. This chapter includes all the reviews in the guideline pertaining to diet and lifestyle interventions.

DIET

A healthy diet, as based on the 'Balance of Good Health', is promoted for the UK population. Some aspects of this are appropriate for people with IBS, e.g. regular meals, drinking plenty of fluid (e.g. 8 cups of non-caffeine based fluid per day) and encouraging a wide variety of foods. However, people with IBS often find that following healthy eating advice exacerbates symptoms and, in particular, this may relate to dietary fibre and lactose (milk and dairy foods). Wheat, resistant starch, caffeine, fructose, sorbitol, alcohol and fizzy drinks have also been reported to commonly affect symptoms. Potential beneficial components of the diet include probiotics and prebiotics and water soluble dietary fibre. Diet and nutrition are fundamental in the management of IBS to avoid malnutrition and to contribute to achieving optimal symptom control.
Food products have been reported as causing, contributing to and perpetuating Irritable Bowel Syndrome. The term 'food intolerance' includes effects of pharmacologically active constituents (e.g. caffeine in coffee); enzyme deficiencies (e.g. lactose intolerance) and demonstrable immunological response (allergy or hypersensitivity to peanut, cow's milk, gluten, soya bean). The notion of food intolerance and food allergy is not new and many IBS patients give a history of food intolerance, although few clinicians consider food hypersensitivity to be a cause of IBS. There are no objective tests available to identify food intolerance and few to confirm food allergy. Data from dietary elimination and food challenge studies are contradictory.

Dietary intolerance is defined as a non-immunologically mediated response to particular foods, which resolve following dietary elimination and re-occur with food challenge. An exclusion diet is defined as a diet in which specific food products are totally excluded for a specified period of

time. The excluded food products are then gradually re-introduced one by one to confirm response.

Diagnostic testing for food intolerance includes hydrogen breath testing and diagnostic testing for Coeliac Disease. Hydrogen breath tests are based on the fact that the only source of hydrogen gas in humans comes from the bacterial metabolism of carbohydrates. Different carbohydrates are given orally to patients and the amount of hydrogen in the expired air is measured. Patients need to be fasted and to have had at least one day of a low fibre diet. Smoking and exercise alters the hydrogen concentrations so are not permitted during the tests.

Potential sources of error are:
- Carbohydrate malabsorption in chronic pancreatitis and Coeliac disease
- False positive for small intestinal bacterial overgrowth due to colonic fermentation
- Delayed gastric emptying may cause false negative
- Oral bacterial flora, failure to follow low fibre diet and rapid transit through small intestine may produce false positive.

Testing for Coeliac disease involves a blood test for immunoglobulin A (IgA) antigliadin antibodies; endomysial antibodies (EMA) and TTG anti-tissue transglutaminase antibodies. The sensitivity and specificity for IgA, EMA and TTG are 95% and 89%; 100% and 97%; and 100% and 97%, respectively in patients with GI symptoms. For general population screening, EMA and TTG have a positive predictive value of 15.7% and 21.8%. A positive blood result requires an endoscopy with duodenal biopsy to confirm a diagnosis of Coeliac disease.

People with IBS may alter their diet to alleviate symptoms of IBS. Guidance may either be sought from inadequately qualified nutritionists or be self directed. Excluding individual foods or complete food groups without appropriate dietetic supervision can lead to inadequate nutrient intake and ultimately malnutrition, e.g. calcium. In addition, symptoms often remain unresolved leading to further inappropriate dietary restriction. The gold standard diagnosis for intolerance to a food is by elimination and reintroduction. Intolerance would be demonstrated if symptoms resolved on elimination and reappeared on reintroduction. Importantly, dietary advice will vary depending on symptoms, e.g. diarrhoea and/or constipation, abdominal bloating and therefore needs to be tailored to the individual to manage symptoms. Expert professional advice on diet and nutrition for IBS should be obtained from a registered dietitian or an appropriately qualified nutritionist.

Dietary Fibre

Fibre is defined as non-starch polysaccharides in agreement with FAO/WHO/DOH measurement methods. Dietary fibre is food material that is not hydrolysed by enzymes secreted by the human gastrointestinal tract. Soluble fibre dissolves in water forming a gel and may be digested by the colonic microbiota increasing bacterial numbers and thus faecal bulk. It

includes β-glucans, pectins, gums, mucilages and some hemicelluloses. Dietary sources include oats, psyllium, ispaghula, nuts and seeds, some fruit and vegetables and pectins. Insoluble fibre is not readily broken down by the gastrointestinal microbiota and it increases faecal bulk, shortening colonic transit. It includes celluloses, some hemicelluloses and lignin and is chiefly found in corn (maize) and wheat bran and some fruit and vegetables.

An increase in fibre has often been suggested as an initial treatment for IBS, although more recently there are conflicting data to support its effectiveness and a range of views on its usefulness. The dietary reference value for non-starch polysaccharides (fibre) is 18g per day for adults. A high fibre diet is defined as 18g or more per day in recognition of the fact that many people in the UK eat on average 10 to 12g per day. Dietary manipulation of the fibre content in practice is dependent on the presenting symptom profile (constipation dominant, diarrhoea dominant or alternating symptoms) and whether abdominal bloating and flatus is present.

Wheat
Wheat is a grass and is cultivated worldwide as a food grain, ranking second in total production as a cereal crop behind maize. Whole wheat is made up of 14% bran, 2.5% germ and the rest is starchy endosperm. Wheat bran has a faecal bulking effect, delays gastric emptying and accelerates small bowel transit (McIntyre 1997). Wheat is found in bread, many breakfast cereals, pasta, cakes and biscuits and is one of the major cereals consumed in the UK. In IBS, wheat consumption is often associated with increased symptoms which may be due to the content of fibre, fructans or resisitant starch. Increasing the variety of other cereals and reducing, but not necessarily, excluding wheat may be beneficial in IBS.

Resistant Starches
Resistant starch comprises starch polymers that are not readily digested in the stomach or small intestine Resistant starches are the total amount of starches, and the products of starch degradation that resist digestion in the small intestine of healthy people (Asp 1982) and therefore reach the colon intact. The extent of resistance is influenced by the structure of naturally occurring starch polymers and food processing methods employed, e.g. how starch changes during cooking and cooling. People with IBS may benefit from a reduction of foods high in resistant starch to alleviate symptoms of wind and bloating.Common dietary sources of resistant starch are cold or re-heated potatoes, bread, cereal products containing modified starch (e.g. cakes, biscuits and breakfast cereals.

Lactose
Lactose is a sugar found in milk of all mammalian varieties including cow, goat, sheep and human and it is also used in processed foods, particularly slimming products. Approximately 10% of people with IBS have lactose intolerance (BSG Guidelines). The symptoms of IBS are brought on by undigested lactose passing into the small intestine causing an increase in the

secretion of fluid into the small bowel through osmotic mechanisms. It then passes into the colon undigested and is available for colonic fermentation as described above (Mascolo 1998).

Removing lactose from the diet may not lead to complete symptom relief in IBS and exclusion needs careful monitoring due to other nutritional inadequacies in the diet e.g. calcium. Often people with lactose intolerance can manage 10 to 12g lactose per day if spread throughout the day. Milk contains the highest level of lactose (cow's milk 5g per 100ml), foods that are lower include butter (trace), cheese (cottage cheese: 1g per tablespoon, processed cheese: 1g per slice, Cheddar, edam, brie, Danish blue: trace), yoghurts (trace – 4g per pot) and low lactose milk. It is therefore relatively easy to include a sufficient amount of dairy foods to maintain a balanced diet in diagnosed and self reported people with lactose intolerance.

Fructose

Fructose intake has increased considerably as a result of an increased consumption of high fructose corn syrup, fruits and juices and crystalline fructose. Fructose is almost twice as sweet as normal table sugar (sucrose). Fructose is an important source of energy for humans, but incomplete absorption in the small bowel can lead to colonic fermentation causing diarrhoea, wind and bloating.

Up to 80% of healthy subjects incompletely absorb 50g of fructose (Scoog 2004). In real terms 25g fructose is equivalent to that found in 200ml apple juice or 2 bananas. A regular consumption of dried fruit and high juice squash will easily add another 25g.

Sorbitol

Sorbitol is a natural component of fruits and significant amounts are found in dried apple and apricots, prunes, cherries and pears. Produced from maize it is also used as an artificial, low calorie sweetener for its low cariogenicity, e.g. in sugar-free chewing gum, mints and cough syrups and as a humectant and thickener in confectionary, frozen desserts and toothpaste. It is poorly absorbed in the small bowel and in the colon has a laxative effect if consumed in quantities of around 30g/day, although some individuals, particularly people with IBS may be sensitive to much less (Thomas 1992).

Caffeine

Caffeine is found naturally in many plant-derived foodstuffs and beverages, chiefly coffee, tea, cocoa and chocolate confectionary, cola and other stimulant drinks. It is also found in many pharmacological agents. Caffeine has many reported effects on the body: negative effects include raised blood pressure, increased heart rate, arrhythmias, dehydration, anxiety, insomnia, headaches and heartburn. Caffeine can also stimulate the central nervous system, improve alertness and mental efficiency and improve athletic performance (Thomas 2003). There is a general consensus that a moderate intake of caffeine (up to 300mg/day in adults) is not harmful.

Caffeine has stimulatory effects on the digestive system but there is little evidence that it will cause gastrointestinal dysfunction (Thomas 2003). Heartburn is the most commonly reported symptom from drinking coffee. It may promote gastrointestinal reflux and stimulate gastrin release and gastric acid secretion but does not appear to affect gastric emptying or small bowel transit.

Probiotics and Prebiotics

In IBS, the gastrointestinal flora may undergo both qualitative and quantitative changes and the most common finding is a decrease in the population of 'good bacteria' such as *Bifidobacteria* and *Lactobacilli* and the faecal microflora has increased numbers of facultative organisms (Madden and Hunter 2002; Quigley 2007). Probiotics may be useful in the management of IBS, however dose and specific bacterial strain are important. In vivo studies have identified some of the variables that determine the survival of probiotics through the GI tract, and some have attempted to quantify the degree of survival of the dose administered. This was found to vary from 10 to 50% depending on the probiotic species used and the dose administered.

For the purposes of this guideline probiotics are defined as microbial food supplements which, when administered in adequate amounts, have a beneficial effect on the host. Prebiotics are defined as a non-digestible food ingredient that affects the host by selectively targeting growth and/or activity of one or more bacteria in the colon that can improve health. Synbiotics are defined as a combination of pre and probiotics which beneficially affects the host by improving survival and implantation of live microbial dietary supplements in the gastrointestinal tract.

Fermented milks and yoghurts have been the most commonly used carrier of probiotics. The probiotic organism is added at the end of the milk fermentation process. The range of probiotic products is expanding to include cheese, frozen yoghurt, ice cream and non-dairy foods, liquids, powders, capsules and drinks. It should be noted that many available probiotics have not had their health benefits identified or been scientifically proven to be beneficial to the host (Reid 2001).

In vivo studies have identified some of the variables that determine the survival of probiotics through the GI tract, and some have attempted to quantify the degree of survival of the dose administered. This was found to vary from 10 to 50% depending on the probiotic species used and the dose administered. The GDG defined the minimum acceptable dose to be 1×10^6 (one million) bacteria per day. The duration of the intervention is also considered important. To avoid concerns regarding possible effects during the menstrual cycle, four weeks was thought to be the minimum duration of intervention.

Colonic Fermentation

Some of the symptoms of IBS (e.g. abdominal bloating, flatus and diarrhoea) may be due to colonic fermentation by intestinal microflora of certain dietary constituents to short chain fatty acids (acetate, butyrate and propionate) and gases (hydrogen, carbon dioxide and methane). The short chain fatty acids have been shown to stimulate ileal and colonic smooth muscle contractility (Barbara 2005). Watery diarrhoea may also happen due to the increased osmotic load. The dietary constituents include non absorbed lactose (as in lactose intolerance), dietary fibre/non-starch polysaccharides, resistant starches and oligosacchaides from wheat and other grains.

Aloe vera

Aloe vera (*Aloe barbadensis* Miller) belongs to the Liliaceal family of which there are approximately 360 species. Aloe vera is a cactus like plant; cosmetic and medicinal products are derived from the leaf tissue and called aloe vera gel. Aloe sap or juice, often referred to as aloes, are derived from the peripheral bundle sheath cells of aloe vera. Aloe vera sap contains anthraquinones that are known to have laxative effects. These are not found in the gel but may be present in total leaf extracts (Vogler and Ernst 1999). The use of aloe vera is being promoted for many conditions including IBS. Most of the evidence is based on anecdotal, historical use rather than scientific evidence.

PHYSICAL ACTIVITY

The relationship between physical activity and chronic disease

There is strong evidence from observational studies that moderate to high levels of physical activity can have a substantial impact on major non communicable diseases, such as coronary heart disease (CHD), hypertension, diabetes and certain types of cancer (US Department of Health and Human Services, 1996; Department of Health, 2004a; WHO, 2004). People who are physically active typically experience 30 to 50% reductions in relative risk of CHD compared with people who are sedentary, after adjustment for other risk factors (Murphy 2003).

The Chief Medical Officer (CMO) recently published a report stating the importance of physical activity for health (Department of Health, 2004a). As well as linking chronic disease with physical inactivity the report also described how physical activity can reduce the risk of musculoskeletal health conditions, including osteoporosis, back pain and osteoarthritis. It stated that regular physical activity can reduce the risk of depression and promotes many other positive mental health benefits including reducing anxiety and promoting self esteem (Department of Health, 2004a). The CMO's report also presented a series of recommendations for the amount of physical activity that should be undertaken by different population groups. These recommendations mimicked similar recommendations from other international bodies (Pate et al, 1995; US Department of Health and Human Services, 1996; Department of Health, 2004a). The report advised that adults should undertake at least 30 minutes of moderate intensity

physical activity on at least five days of the week (Department of Health, 2004a). In 2002 the cost of physical inactivity was estimated to be £8.2 billion annually in terms of mortality, morbidity and quality of life (Department for Culture Media and Sports and London Strategy Unit, 2002). A more accurate estimate of the direct costs of physical inactivity to the UK health service was £1.06 billion annually (Allender et al, 2006a). Physical activity has been described as a good investment for public health, not only because of the great potential for benefit, but also because 'it is inexpensive and has few side-effects' (Morris 1992, in Marmot and Elliot 1992).

In 2006, NICE published guidance (Public Health Intervention Guidance No. 2) on exercise interventions in primary care, pedometers, exercise referral schemes and community-based exercise programmes for walking and cycling to increase physical activity. Two specific recommendations were made for primary health care professionals:

Recommendation 1

Primary care practitioners should take the opportunity, whenever possible, to identify inactive adults and advise them to aim for 30 minutes of moderate activity on 5 days of the week (or more)*. They should use their judgement to determine when this would be inappropriate (for example, because of medical conditions or personal circumstances). They should use a validated tool, such as the Department of Health's forthcoming general practitioner physical activity questionnaire (GPPAQ), to identify inactive individuals.

* The practitioner may be a GP or another professional with specific responsibility for providing encouragement or advice. This will depend on local conditions, professional interest and resources. Health trainers are likely to have a role in offering brief advice. 'Inactive' is used as shorthand for those failing to reach the CMO's recommendation. 'Advise' is used as shorthand for 'encourage, advise, discuss, negotiate'.

Recommendation 2

When providing physical activity advice, primary care practitioners should take into account the individual's needs, preferences and circumstances. They should agree goals with them. They should also provide written information about the benefits of activity and the local opportunities to be active. They should follow them up at appropriate intervals over a 3 to 6 month period.

The NICE public health intervention advisory committee determined that there was insufficient evidence to recommend the use of exercise referral schemes to promote physical activity, other than as part of research studies where their effectiveness can be evaluated.

This guidance aims to help practitioners deliver effective interventions that will increase people's physical activity levels and therefore benefit their health.

The use of physical activity as part of a non-pharmacological therapy for IBS is described as "reasonable" despite the relationship between exercise and gastrointestinal system being unclear (Bi and Triadafilopoulos 2003). For example, moderate physical activity (e.g. brisk walking) is reported to improve gut transit time, whereas vigorous physical activity (e.g. running) can result in "runners trots" (Oettle, 1991). Physical activity has been associated with improved outcomes in uncontrolled studies (Colwell et al, 1998).

7.1 General dietary and lifestyle advice

This section is concerned with the effect of diet and lifestyle on IBS and its management. Five reviews are addressed, fibre, probiotics, aloe vera, exclusion diets and physical activity. In addition, the GDG made some consensus recommendations, partly informed by dietary advice leaflets. These are listed below.

RECOMMENDATION

Healthcare professionals should encourage people with IBS to identify and make the most of their available leisure time and to create relaxation time.

RECOMMENDATION

Diet and nutrition should be assessed for people with IBS and the following general advice given.

- Have regular meals and take time to eat.
- Avoid missing meals or leaving long gaps between eating.
- Drink at least eight cups of fluid per day, especially water or other non-caffeinated drinks, for example herbal teas.
- Restrict tea and coffee to three cups per day.
- Reduce intake of alcohol and fizzy drinks.
- It may be helpful to limit intake of high-fibre food (such as wholemeal or high-fibre flour and breads, cereals high in bran, and whole grains such as brown rice).
- Reduce intake of 'resistant starch' (starch that resists digestion in the small intestine and reaches the colon intact), which is often found in processed or re-cooked foods.
- Limit fresh fruit to three portions per day (a portion should be approximately 80g).
- People with diarrhoea should avoid sorbitol, an artificial sweetener found in sugar-free sweets (including chewing gum) and drinks, and in 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).

7.2 Physical activity

SELECTION CRITERIA

The selection criteria described in the general methodology section were to be used, but some were specific to the physical activity review and are reported below.

Types of studies

For intervention studies, randomised trials (RCTs) examining the use of physical activity for the treatment or management of IBS were to be included. In the absence of randomised trials, quasi randomised studies were to be considered. Crossover trials with a washout period of less than 2 weeks were to be excluded. All study designs were to be included for adverse effects, but specific searches for adverse effects will not be carried out. Studies were restricted to the English language.

Types of intervention

Studies were included if they had one or more of the following interventions:

- The use of physical activity alone or in combination with other therapies
- 12 weeks minimum length of intervention

A physical activity intervention is defined as the use of physical activity or exercise as a therapeutic and/or preventative medical procedure used to support the management and treatment of IBS. Physical activity is usually defined as *any force exerted by skeletal muscles that results in energy expenditure above resting level* whereas exercise is defined as *a subset of physical activity, which is volitional, planned, structured, repetitive and aimed at improvement or maintenance of any aspects of fitness or health* (Casperson et al, 1985). The GDG defined the minimum acceptable dose of physical activity to be at least 30 minutes per week of at least moderate intensity physical activity. The duration of the intervention is also considered important, and the minimum duration of intervention was set at twelve weeks.

Types of comparisons

The following comparisons were to be included

- Physical activity versus attention control
- Combination of physical activity with another non-pharmacological intervention (e.g. diet advice) versus control.

Types of participants

Studies were to be included if the participants were:

- Adults (18 years and over)
- Had symptoms of IBS
- No serious diseases (e.g. cancer, heart disease) other than IBS
- Did not have a single symptom of IBS only (e.g. not constipation only)

In the absence of studies in patients with IBS, we extended the review to cover studies in people with single symptoms such as constipation or diarrhoea. Studies in these participants were regarded as indirect as far as the population was concerned.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL and *The Cochrane Library (*1966 to current day with guidance from the GDG). An additional database searched for this review only was SPORTS DISCUS. The search strategies are listed in Appendix B.

Subgroup analyses

Subgroup analyses were proposed to examine any heterogeneity as follows:

- Dose
- Type of physical activity
- Symptom severity.

Sensitivity analyses

The following sensitivity analyses may be considered:

- Setting (primary/secondary care).

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

The search strategy identified 2608 studies. The titles and abstracts of these studies were assessed. Of these, 19 that were potentially relevant to the review were identified on the basis of the title and abstract – these papers were retrieved in full. All reference lists of these studies were inspected for potential papers for inclusion in the review, but none was found.

None of the studies identified met the primary inclusion criteria. Therefore, we included some studies with indirect evidence, and considered other studies to aid GDG discussions. One systematic review was identified (Bi and Triadafilopoulos 2003). This review examined the relationship between exercise and gastrointestinal function for eight disease types.

RESULTS

Evidence from Systematic Reviews

Bi and Triadafilopoulos (2003) reviewed the relationship between exercise and gastrointestinal function for eight disease types. The authors described their review as systematic but provided no methods in the paper.

1. Gastroesophageal reflux disease
2. Gastric emptying and gastric acid production
3. Peptic ulcer disease
4. Inflammatory bowel disease

5. Constipation and gastrointestinal motility
6. Colorectal cancer
7. Gastrointestinal bleeding
8. Liver disease

The authors attempted to identify if there were any differential effects of physical activity (by intensity or type) on gastrointestinal function. IBS was not identified as a separate class, but it may be possible to extrapolate from the indirect evidence in section 5 of the Bi and Triadafilopoulos (2003) review. Participants in these studies tended to be young, fit and active males, rather than typical clinical populations. The review found that physical activity could improve gastric emptying and lower the risk of bowel cancer. However, there was insufficient evidence to suggest that exercise can relieve chronic constipation. The authors also noted consistent improvements in aerobic fitness and general health for all subjects participating in regular physical activity programmes and that this outcome is a notable behavioural goal for sedentary patients. The majority of risks to gastrointestinal organs relate to very high levels of sustained physical activity (performed at elite levels). However these risks do not outweigh the benefits of light and moderate physical activity.

Evidence from intervention studies

One randomised trial was included as an indirect study as it examined the impact of physical activity upon adults with chronic constipation only (De Schryver 2005). However this study was not included in the Bi and Triadafilopoulos (2003) review.

The aim of the study was to investigate the effect of regular physical activity on colonic transit time and defecation in middle aged inactive patients suffering from chronic idiopathic constipation. Forty three adults aged 51 to 61 were recruited from general practice lists and pharmacies. Using Rome I criteria all were categorised as suffering from constipation, with IBS patients excluded. Participants' physical activity levels were also assessed using a self report measure and all the participants were categorised as sedentary if they failed to reach the current physical activity recommendation (under 30 minutes or more of moderate physical activity on most days of the week).

Other baseline measures included food consumption, assessed by self report using a diary, in order to determine the average daily fibre and water intake. Defecations patterns were recorded in a 7-day diary, at the start of the study and at 12 weeks follow up. Colonic transit time was measured using radiopaque markers and x-rays. Transit time was calculated based on the number of markers visible in the colon, segmented into three (right colon, left colon and rectosigmoid).

Participants were randomised in to two groups, physical activity versus waiting list control. Group A maintained their normal lifestyle for 12 weeks, and then started their 12 week physical

activity programme. Group B started their physical activity programme immediately after randomisation. Both groups were given dietary advice by a dietician concerning the consumption of fluid and fibre at the start of the study. Group A received a second dietary advice after 12 weeks, before starting the physical activity programme. This programme consisted of both aerobic and strength/flexibility exercises. Brisk walking was chosen for aerobic training and strength/ flexibility exercises were chosen for a home based programme. Brisk walking was performed at least twice a week for at least 30 minutes per session, performed at 70 to 80% of the subject's maximal heart rate. Participants were able to monitor their heart rates using a Polar sports tester (a heart rate monitor worn on the wrist, in conjunction with a chest sensor). Maximal heart rate was assessed for all patients at baseline using a maximal heart test performed on a cycle ergometer. Participants were also asked to perform a walking test on a treadmill at 70% of their heart rate for 5 minutes to establish an average heart rate for their brisk walking.

The number of defecations did not change in either of the study groups (Table 1). However in Group B the percentage of incomplete stools decreased significantly, compared to Group A at 12 weeks (Group A from 58.8% to 39.5% whereas in Group B from 54.3% to 27.4%).

Table 1. Defecation patterns at baseline and after 12 week physical activity programme for 41 adult participants aged 51-61 years old (De Schryver et al, 2005)

	Group A (12 weeks inactive, 12 weeks PA)			Group B (12 weeks PA)	
	Week 0	Week 12	Week 24	Week 0	Week 12
No. of defecations/wk	7.1 ±0.8	7.5 ±1.1	7.8 ±1.1	7.5 ±1.1	7.8 ±1.0
% Hard stools	53.8 ±8.5	51.9 ±9.5	35.1 ±9.2	59.5 ±8.7	39.5 ±6.8*
% Straining at defecation	65.7 ±7.7	69.2 ±7.9	54.3 ±9.8	71.2 ±4.6	40.4 ±6.4*
% Incomplete stools	51.3 ±7.9	58.8 ±8.5	39.5 ±8.9	54.3 ±7.2	27.4 ±6.0*
No. of Rome criteria	2.3 ±0.1	2.6 ±0.2	1.7 ±0.3*	2.7 ±0.1	1.7 ±0.2*
% Patients with ≥2 Rome criteria	100	89	61*	100	64*

PA = physical activity.
Data are given as means ± SEM.
*$p < 0.05$

Despite randomisation, there were considerable differences between right and total colonic transit times at baseline between groups (Table 2). No significant changes in right or left colonic transit time were observed in either group at the end of the physical activity programme. In Group B there was an observed acceleration in rectosigmoid mean transit time compared to Group A. Total colonic transit time also improved with a significant reduction in Group B. The

authors reported that there was no correlation between fibre intake and improvements in defecation patterns and colonic transit times.

The GDG noted that the normal total colonic transit time is 72 hours and concluded that group B was significantly different from group A, so that the study was considered to be at least partially confounded.

The evidence from this study was assessed to be low, using GRADE criteria. Limitations included (i) the study was conducted in secondary care (ii) there were important differences in baseline characteristics, (iii) IBS patients were excluded. This study was also limited because the participants had relatively high levels of baseline physical activity, which equates to over 2 hours of walking per week, and may not be representative. The study did show that moderate physical activity could deliver a consistent reduction in total colonic transit time and improvements in ROME I symptoms amongst older adults with chronic constipation.

Table 2. Colonic transit times (hours) at baseline and after 12 week physical activity programme for 41 adult participants aged 51-61 years old (De Schryver et al, 2005)

	Group A (12 weeks inactive, 12 weeks PA)			Group B (12 weeks PA)	
	Week 0	Week 12	Week 24	Week 0	Week 12
RCTT	15.1 ±2.2	14.0 ±2.7	13.8 ±2.1	27.5 ±4.7	22.2 ±2.8
LCTT	27.5 ±4.9	29.5 ±6.1	33.9 ±6.9	33.8 ±5.0	27.6 ±4.9
RSTT	16.9 ±3.0	18.9 ±3.0	14.3 ±3.2	17.5 ±2.5	9.6 ±1.6
Total CTT	59.5 ±8.4	62.4 ±9.5	61.0 ±9.9	79.2 ±9.1	58.4 ±7.7

RCTT = right colonic transit time; LCTT = left colonic transit time; RSTT = rectosigmoid transit time; PA = physical activity.
Data are given as means ± SEM.
p < 0.05

Studies used to aid GDG discussions

One pre-post intervention study examined the impact of a lifestyle education programme upon IBS symptoms. This study design was judged inadequate to make recommendations on interventions, but was considered useful to inform GDG discussions, and does illustrate a suitable approach for evaluating a lifestyle intervention for IBS patients.

Colwell et al (1998) assessed the impact at one and six months of a patient education class, that included exercise, on 52 adult outpatients with IBS (definition of IBS not stated). Patients were advised to increase their physical activity by walking or basic stretching exercises during one 3 hour structured class, delivered by a specialist nurse. Pre-class data was compared with results for physical activity levels at follow up. Exercise scores increased significantly at one month but

not at 6 months, compared with baseline, using a self-rating scale. It is difficult to assess if this increase was clinically significant because the physical activity variable was assessed using a categorical scale, and so the physical activity change scores were not adjusted for baseline values. Pain scores at 1 and 6 months reduced significantly (see Table 3). The Manning score also decreased significantly, on a scale of 0 to 6 using Manning criteria.

Table 3. Symptom scores at 1 and 6 months for 57 adult participants in an IBS educational class aged 21-79 years old (Colwell et al, 1998, page 903)

Score	Median Scores (ranges)		
	Baseline	1 month	6 months
Pain*	3.0 (1.9-3.9)	2.4 (0.0-3.7)§	2.6 (0.0-4.0)§
Manning∇	4.0 (1.0-6.0)	3.0 (0.0-6.0)§	3.0 (0.0-6.0)§

* Pain score: a weighted average of severity, frequency, and duration of pain on a scale from 0-4
∇ Manning score: On a scale of 0 to 6 Manning criteria: pain relief with defecation; looser stools with pain onset; abdominal distension; mucus in the stool; and a feeling of incomplete evacuation (2)
§ $p < 0.01$.

Evidence from Epidemiological studies

Three observational studies reported the prevalence and association between IBS and physical activity. In a case-control study, Kim and Ban (2006) reported a small, non-significant difference in the mean number of hours of exercise per week for students with IBS compared to students without IBS (defined by ROME 2 criteria) (students with IBS 2.38 h/week, SD 3.2 versus students with non-IBS: 2.69 h/week, SD 6.3).

Figure 1:
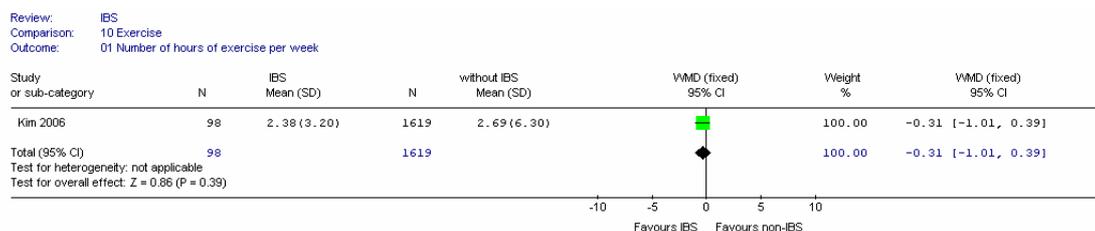
Lustyk et al (2001) compared prevalence and severity of IBS symptoms between active and sedentary women with IBS (defined by ROME I criteria). They found that active and inactive women reported the same level of recalled psychological and somatic symptoms as well as daily reports of GI and psychological distress. Active women (those who took at least 2.5 hours per week of moderate physical activity and meeting recommended physical activity levels) reported significantly less fatigue than sedentary women. This outcome was assessed by combining frequency and severity of fatigue using categorical scale. No differences were observed for

other somatic symptoms including backache, headache and insomnia between active and sedentary women with IBS.

Dancey et al (2002) examined gender differences in the prevalence and severity of IBS symptoms. They used a cross sectional survey to compare the prevalence and impact of IBS symptoms between 117 male and female IBS patients. IBS was assessed by self report measure with respondents rating severity of abdominal pain, constipation, diarrhoea, incomplete evacuation after bowel movement, bloating and flatulence on a 7 point severity scale (0 = no symptoms to 7 = extremely severe). Illness intrusiveness ratings were assessed across 13 life domains, using a 7 category Likert scale. Respondents were asked to rate the extent to which their illness interfered with each life domain important to quality of life (e.g. health, diet, financial situation, relationship with partner etc.). One life domain was active recreation (e.g. sports). The authors reported that in response to this item, men and women scored the interference of IBS similarly (i.e. moderate interference), with no significant differences between genders. They found that IBS inference was higher in diet, health and self expression domains than active recreation. Other domains reporting less interference than active recreation were social relations, work, community/civic life, sex life, relationship with spouse, family relations, financial situation, passive recreation and religious expression.

Two studies examined the relationship between physical activity and bowel frequency in the general population. In a cohort study, Sanjoaquin et al (2004) investigated the association between mean number of bowel movements and physical activity, adjusted for other confounding variables (e.g. age, BMI, diet, fibre intake) in 20,630 EPIC-Oxford cohort participants. The EPIC-Oxford cohort is a cohort study forming part of the European Prospective Investigation into Cancer and Nutrition (EPIC). Participants were recruited from general practice surgeries, vegetarian and health food magazines, the Vegetarian Society, the Vegan Society and from friends and relatives of participants. In a follow up study a short questionnaire was sent to all participants and included two questions relating to bowel movements, (i) "About how many bowel movements do you have each week? And (ii) How often do you take laxatives?" The number of bowel movements was counted for each participant. Respondents were then dichotomised into one of two groups, either above of below 7 movements per week.

The authors reported a positive association between increasing amounts of vigorous physical activity and mean number of bowel movements per week for both men and women. However only highly active women (more than 7 hours per week of vigorous physical activity) had a greater likelihood of reporting more than 7 bowel movements per week (OR; 1.70 [95%CI 1.42, 2.03]) compared to women who reported no vigorous physical activity. Curtin et al (1996) conducted a population survey of bowel habits in urban Swiss men but found no relationship between physical activity status and bowel habits.

EVIDENCE STATEMENTS

1. There is poor evidence to show that the percentage of incomplete stools decreased significantly in non-IBS constipated people given an exercise programme.

2. There is weak evidence that IBS Manning and pain scores at one and six months were reduced significantly in comparison with pre-intervention scores following a patient education class that included exercise for people with IBS.

3. There is mixed evidence on whether there is a positive association between physical activity and bowel habits in the general population.

GDG DISCUSSION

The GDG considered the evidence and discussed whether exercise effects were related to stress reduction. It was noted that some people may have increased stress levels because of exercise, depending on their liking for exercise. The GDG thought that exercise would not necessarily be beneficial for people with IBS-D. It was also noted that attendance at exercise classes might prove difficult for patients and a gentle exercise programme that could be carried out at home (e.g. Tai Chi, yoga, stretching) might be more beneficial.

The GDG discussed whether it was useful to recommend taking more fluid after exercise, but concluded that this would not necessarily be appropriate for people with IBS, since many have bladder problems, and taking more fluid does not help constipation.

EVIDENCE TO RECOMMENDATION

The GDG took into consideration the limited evidence and also referred to NICE public health guidance and the Chief Medical Officers report on physical activity. This led to a general recommendation for practice.

Recommendations for active living throughout the lifecourse (DoH 2004)

- For general health benefit, adults should achieve a total of at least 30 minutes a day of at least moderate intensity physical activity on 5 or more days of the week.
- The recommended levels of activity can be achieved either by doing all the daily activity in one session, or through several shorter bouts of activity of 10 minutes or more. The activity can be lifestyle activity* or structured exercise or sport, or a combination of these.
- More specific activity recommendations for adults are made for beneficial effects for individual diseases and conditions. All movement contributes to energy expenditure and is important for weight management. It is likely that for many people, 45-60 minutes of

---

* Lifestyle activity: activities that are performed as part of everyday life, such as climbing stairs, walking (for example to work, school or shops) and cycling. They are normally contrasted with 'programmed' activities such as attending a dance class or fitness training session.

moderate intensity physical activity a day is necessary to prevent obesity. For bone health, activities that produce high physical stresses on the bones are necessary.

- The recommendations for adults are also appropriate for older adults.
  Older people should take particular care to keep moving and retain their mobility through daily activity. Additionally, specific activities that promote improved strength, co-ordination and balance are particularly beneficial for older people.
- People with disabilities will know their abilities and should modify their physical activity accordingly e.g. chair-based exercises.

The GDG was also interested to know if exercise affects IBS symptoms and quality of life for people with IBS, and whether the type of IBS was important.

RECOMMENDATION

Healthcare professionals should assess the physical activity levels of people with IBS, ideally using the General Practice Physical Activity Questionnaire (GPPAQ; see Appendix J). People with low activity levels should be given brief advice and counselling to encourage them to increase their activity levels.

7.3 Fibre

SELECTION CRITERIA

The selection criteria described in the general methodology section were to be used, but some were specific to the fibre review and are reported below.

Types of studies

The GDG decided that the washout period for crossover studies in this review should be at least 4 weeks. Trials with shorter washout periods were not to be included in the analysis.

Types of intervention

Studies were to include the following interventions:

- Insoluble fibre (corn, wheat, fruit and vegetables)
- Soluble fibre (pectins, fruit and vegetables, oats, nuts and seeds, psyllium, ispaghula)
- Bran.

It was to be noted if the fibre was provided as a food or as a capsule/supplement. In addition, the total amount of fibre in the diet for each intervention was to be recorded where possible.

The following comparisons were to be included:

- Fibre + normal diet versus normal diet (fibre versus nothing)
- Fibre versus low fibre diet or placebo (fibre versus placebo)
- Bran versus placebo
- Insoluble fibre versus soluble fibre
- Insoluble fibre + soluble fibre versus soluble fibre
- Insoluble fibre + soluble fibre versus insoluble fibre
- Fibre level 1 versus fibre level 2
- Duration of treatment 1 versus duration 2
- Fibre versus another type of intervention
- Fibre plus another type of intervention versus another type of intervention.

In spite of the large placebo effect associated with IBS, comparisons with no treatment were to be included.

The fibre review was to be concerned only with longer-term maintenance treatment. The GDG decided that there should be a minimum duration of treatment of four weeks for this review. Studies of shorter durations were to be excluded.

Outcomes

In addition to the outcomes discussed in the general methods section, the GDG were interested in the number of people with global deterioration, other than those who withdrew because of the treatment.

Data extraction

In addition to the items given in the general section, we also extracted information on the total amount of fibre (i.e. the sum of the intervention and the fibre in the diet).

Subgroup analyses

We planned to carry out subgroup analyses by type of fibre (soluble, insoluble, mixed), dose (both intervention and total amount), duration of intervention, and, post-hoc, by means of ingestion (supplement or dietary).

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

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. The search strategies are given in Appendix B.

The titles and abstracts from the search strategy were assessed. Sixty-four were identified to be potentially relevant to the review and these papers were retrieved in full. Twenty studies met the inclusion criteria for the review. The reference lists of the retrieved studies were inspected for further potential papers, but none were identified. The forty-four excluded studies are listed in the Appendix, along with reasons for exclusion.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

There were 20 included studies (Aller 2004; Arthurs 1983; Chapman 1990; Cook 1990; Dettmar 1999; Fielding 1984; Fowlie 1992; Kruis 1986; Longstreth 1981; Lucey 1987; Manning 1977; Parisi 2002; Parisi 2005; Prior and Whorwell 1987; Rees 2005; Ritchie 1979; Ritchie 1980; Soltoft 1976; Tarpila 2004; Vilagrasa 1991). Nine studies were conducted in the UK (Chapman 1990; Dettmar 1999; Fowlie 1992; Lucey 1987; Manning 1977; Prior and Whorwell1987; Rees 2005; Ritchie 1979; Ritchie 1980); two in Ireland (Arthurs 1983; Fielding 1984); seven in the rest of Europe, and two in the USA and Canada.

One study (Cook 1990) had fewer than 20 participants (n=14). This was a crossover study so fewer participants were required to achieve adequate power. Five studies had more than 100 participants in total (Chapman 1990; Dettmar 1999; Kruis 1986; Parisi 2002; Villagrasa 1991).

Study Design

Setting: The majority of studies took place in secondary care; one was in primary care (Dettmar 1999) and one study did not report the setting (Tarpila 2004).

There were two crossover studies (Cook 1990; Lucey 1987) in which participants were allocated to receive both the intervention and control treatments during the course of the study, in a random order. The GDG defined the minimum washout period to be four weeks for crossover studies in this review, so the only crossover study eligible was Cook (1990). However, a second crossover study (Lucey 1987) became eligible because individual patient data were reported, allowing calculation of first period results. This gave the study a 'pseudo-parallel' design, although the power was reduced. The remaining studies had a parallel design. One study had more than two arms: Kruis (1986) compared bran with mebeverine (anti-spasmodic) and placebo.

Population

The definition of IBS varied between studies: two used the Manning criteria (Chapman 1990; Cook 1990); two used Rome I (Parisi 2002; Rees 2005); two used Rome II (Aller 2004; Parisi 2005) and two met criteria defined by the authors that were similar to the above (Fielding 1984; Tarpila 2004). In five studies, the authors stated that the participants had IBS, with no further explanation (Lucey 1987; Manning 1977; Ritchie 1979; Ritchie 1980; Vilagrasa 1991). The remaining seven studies (Arthurs 1983; Dettmar 1999; Fowlie 1992; Kruis 1986; Longstreth 1981; Prior and Whorwell 1987; Søltoft 1976) did not use a formal definition but described a range of symptoms consistent with IBS.

Most studies included a combination of IBS types. Four specified constipation-predominant IBS (Cook 1990; Fielding 1984, Rees 2005; Tarpila 2004) and three were unclear (Arthurs 1963; Dettmar 1999; Fowlie 1992).

None of the studies stated that any participants had IBS as result of gastrointestinal infection. The majority of studies (13) did not state the number of participants with bloating. Four studies reported that some people had bloating (Aller 2004; Kruis 1986; Longstreth 1981; Vilagrasa 1991). Two studies (Prior and Whorwell 1987; Tarpila 2004) stated that all people had bloating.

Most of the studies did not describe symptom severity. Six studies stated that participants had symptoms of mixed severity (Dettmar 1999; Fowlie 1992; Longstreth 1981; Parisi 2002; Parisi 2005; Prior and Whorwell 1987).

The age range of participants across studies was 14 to 82 years, with the mean age (where given) ranging from 25.8 to 45 years. No study particularly identified elderly people. All studies had more women than men.

Interventions

The studies varied in the type of fibre used: six had insoluble fibre (wheatbran); eight had soluble fibre (six ispaghula, one partially hydrolysed guar gum ['PHGG'], one psyllium); five had mixed fibres: studies used a combination of fruit, vegetables and cereal.

One study gave the fibre in a capsule form (Fowlie 1992), eight gave the fibre as a supplement (Arthurs 1983; Chapman 1990; Dettmar 1999; Fowlie 1992; Longstreth 1981; Prior and Whorwell 1987; Ritchie 1979; Ritchie 1980); and the rest added fibre to the diet with food (e.g. bran-containing biscuits).

A fibre level of 18g per day is regarded as a threshold dose. When assessing dose we considered both the amount of additional fibre and the amount of total fibre (intervention plus that in the diet). The amount of additional fibre ranged from 7g per day (Dettmar 1999), although a third 3.5g sachet could be added if needed, to 40g per day (Fielding 1984). Ten studies gave additional fibre as amounts of less than 18g (Chapman 1990; Dettmar 1999; Fowlie 1992; Kruis 1986; Lucey 1987; Parisi 2005; Prior and Whorwell 1987; Rees 2005; Ritchie 1979; Ritchie 1980). Nine studies gave more than 18g (Aller 2004; Arthurs 1983; Cook 1990; Fielding 1984; Longstreth 1981; Parisi 2002; Manning 1977; Søltoft 1976; Villagrasa 1991). One study (Tarpila 2004) gave 12 to 24g daily.

Eight studies reported the total fibre in the intervention arm (Aller 2004; Arthurs 1983; Cook 1990; Fielding 1984; Fowlie 1992; Prior and Whorwell 1987; Tarpila 2004; Villagrasa 1991).

The duration of the intervention ranged from four weeks (Arthurs 1983; Dettmar 1999; Fielding 1984; Parisi 2002) to two years (Villagrasa 1991). One study reported follow-up after the end of the trial (Parisi 2005; 3 months follow-up).

Comparisons

The included studies covered the following comparisons:
- Eleven comparisons of fibre versus placebo, including one versus usual diet (Kruis 1986); and one versus reduced fibre (Manning 1977):
    - Four gave soluble fibre (Arthurs 1983; Longstreth 1981; Prior and Whorwell 1987; Ritchie 1979)
    - Six gave insoluble fibre (Cook 1990; Kruis 1986; Lucey 1987; Manning 1977; Rees 2005; Søltoft 1976)
    - One gave mixed fibre (Fowlie 1992);
- Three studies compared different classes of fibre:
    - Two studies compared soluble versus insoluble fibre
        - PHGG versus bran (Parisi 2002)

- - Ispaghula versus bran (Ritchie 1980)
  - One study compared mixed versus soluble fibre
    - Ground flax seed (containing 20% flaxseed oil) versus psyllium (Tarpila 2004);
- One study compared different types of fibre in the same class (mixed):
  - One study compared different combinations of fruit and cereal fibre (Fielding 1984);
- Two studies compared different doses of fibre:
  - One compared 30.5g with 10.4g of mixed fibre. However, the proportion of soluble fibre differed between the two groups (13% versus 19%) (Aller 2004)
  - One study compared 5 and 10g of PHGG (Parisi 2005);
- Two studies compared fibre + mebeverine versus mebeverine + dietary advice (Chapman 1990; Dettmar 1999)
- Two studies compared fibre with an antispasmodic (Kruis 1987; Villagrasa 1991).

OUTCOMES

The studies measured a range of outcomes.

1.Global symptoms

a) Number of people with improvement in global symptoms

Ten studies recorded the participants' assessment of improvement (Fowlie 1992; Kruis 1986; Longstreth 1981; Lucey 1987; Parisi 2002; Prior and Whorwell 1987; Rees 2005; Ritchie 1979; Ritchie 1980; Søltoft 1976) and one (Arthurs 1983) appeared to record a clinician's assessment.

b) Number of people with deterioration in global symptoms

Four studies recorded the participants' assessment of deterioration (Longstreth 1981; Lucey 1987; Parisi 2002; Søltoft 1976).

c) Global symptom score (mean)

Global symptom scores combined pain, bowel habits, flatulence and bloating. This outcome was recorded by five studies (Cook 1990; Fowlie 1992; Longstreth 1981; Lucey 1987; Parisi 2005). Longstreth (1981) recorded how symptoms interfered with normal activity.

2. Individual symptoms

a) Pain

Pain was reported in several ways: the number of people with pain at the end of the study; the number of people whose pain improved or worsened compared with the baseline, and; pain scores. The pain score recorded a range of features, including severity, frequency and duration, or a combination of these. In addition, studies recorded the final scores, mean daily scores or the change from baseline. The studies reporting the following outcomes are listed below:

  i. Number of people with pain: three studies (Parisi 2002; Prior and Whorwell 1987; Villagrasa 1991)

ii. Number of people with more pain: one study (Chapman 1990)

iii. Number of people with no pain: two studies (Prior and Whorwell 1987; Villagrasa 1991)

iv. Number of people with less pain: four studies (Chapman 1990; Dettmar 1999; Fielding 1984; Kruis 1986)

v. Pain score (change and final): six studies (Aller 2004; Cook 1990; Fowlie 1992; Longstreth 1981; Manning 1977; Parisi 2005):

    a. Three studies reported pain severity at the end of the study (Cook 1990; Fowlie 1992; Parisi 2005)

    b. Two studies reported pain severity from daily diary readings (Longstreth 1981; Manning 1977)

    c. One study reported a combined score for pain frequency and severity (Aller 2004) and this study also reported change scores. In all cases the highest rating meant worst symptoms, although the scales used were not the same.

b) Bloating

i. Number of people with bloating: two studies (Prior and Whorwell 1987; Villagrasa 1991)

ii. Number of people with more bloating: one study (Tarpila 2004)

iii. Number of people with no bloating: two studies (Prior and Whorwell 1987; Villagrasa 1991)

iv. Number of people with less bloating: one study (Tarpila 2004)

v. Bloating score (change and final): no studies reported this outcome.

c) Combined bloating and flatulence score

Three studies measured end of study scores (Aller 2004; Longstreth 1981; Parisi 2005).

d) Bowel habits

i. Number of people with improved bowel habits

Eight studies recorded the number of people with improved bowel habits (Chapman 1990; Dettmar 1999; Fielding 1984; Kruis 1986; Manning 1977; Parisi 2002; Tarpila 2004; Villagrasa 1991). Of these, two reported normalisation of bowel habits (Parisi 2002; Villagrasa 1991), and the rest reported the patient's assessment of improvement.

ii. Stool score (aggregate)

Three studies (Aller 2004; Fowlie 1992; Longstreth 1981) measured an aggregate of frequency, consistency and straining. Fowlie (1992) reported the sum of number of stools x consistency score (1=hard; 5=watery), for people whose IBS type was unclear; we regarded this outcome as unhelpful. Longstreth (1981) reported the number of normal stools and this study was included in the analysis.

e) Quality of life

Two studies reported a measure of quality of life (Fielding 1984; Parisi 2005). Parisi (2005) reported the social functioning item on the SF-36 scale.

f) Adverse events

Two studies reported adverse effects (Chapman 1990; Villagrasa 1991).

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D. The method of randomisation was reported in one study, classified as partially adequate (Manning 1977; drawing a randomly numbered card). The other studies did not state the method of randomisation.

Allocation concealment was reported in two studies (Parisi 2002; Parisi 2005), both of which reported a partially adequate method in which randomisation and analysis were said to be 'supervised by a statistician'.

Nine studies reported that the outcome assessors were blinded to the interventions (Cook 1990; Fielding 1984; Longstreth 1981; Manning 1977; Prior and Whorwell 1987; Ritchie 1979; Ritchie 1980; Søltoft 1976; Tarpila 2004). One study stated that the outcome assessors were not blinded (Parisi 2002). The remaining studies did not report blinding of outcome assessors.

Eleven studies reported that the participants were blinded to the interventions (Arthurs 1983; Cook 1990; Fowlie 1992; Longstreth 1981; Lucey 1987; Prior and Whorwell 1987; Rees 2005; Ritchie 1979; Ritchie 1980; Søltoft 1976; Tarpila 2004). Eight studies stated that the participants were not blinded (or this was deduced from intervention differences) (Aller 2004; Chapman 1990; Dettmar 1999; Fielding 1984; Kruis 1986; Manning 1977; Parisi 2002; Villagrasa 1991). One study (Parisi 2005) was unclear about patient blinding.

Only one study (Cook 1990) described an *a-priori* power calculation. Several studies included in the review demonstrated baseline comparability of the groups, but eight did not give baseline characteristics (Arthurs 1983; Dettmar 1999; Longstreth 1981; Lucey 1987; Manning 1977; Ritchie 1979; Ritchie 1980; Søltoft 1976).

Six studies reported no withdrawals (Aller 2004; Dettmar 1999; Lucey 1987; Parisi 2002; Ritchie 1979; Ritchie 1980). Four studies reported that more than 20% of people in at least one arm (or overall) were not analysed or were lost to follow-up (attrition bias):

- Cook (1990): 5/14 (36%) of participants withdrew from the study

- Longstreth (1981): 6/40 (15%) on placebo and 11/37 (30%) on psyllium did not complete the study. 3/6 and 7/11 respectively dropped out because of dislike for the study preparation or failure to improve; 1/6 and 1/7 dropped out because their symptoms improved
- Prior and Whorwell (1987): 8/40 (20%) withdrew from ispaghula group; 15/40 (38%) withdrew from placebo group. This study reported most recent data carried forward in the analysis, but this is not an approved method of handling missing data. The study also stated that 4/8 and 10/15 withdrawals, respectively, were because of treatment failure.
- Rees (2005): 2/14 (14%) did not complete the study in the intervention arm and 4/14 (29%) on placebo. There were no further details.

Thus, Cook (1990), Longstreth (1981), Prior and Whorwell (1987) and Rees (2005) were treated with caution and examined in sensitivity analyses.

The risk of bias was assessed for each included study. Four studies were assessed as being at higher risk of bias (Cook 1990; Longstreth 1981; Prior and Whorwell 1987; Rees 2005 – attrition bias) and were treated with caution. The eight studies that reported that the participants were not blinded (Aller 2004; Chapman 1990; Dettmar 1999; Fielding 1984; Kruis 1986; Manning 1977; Parisi 2002; Villagrasa 1991) were also treated more cautiously.

RESULTS

A. Fibre versus Placebo

There were eleven studies that compared fibre with placebo (Arthurs 1983; Cook 1990; Fowlie 1992; Kruis 1986; Longstreth 1981; Lucey 1987 first period only; Manning 1977; Prior and Whorwell 1987; Rees 2005; Ritchie 1979; Søltoft 1976). Two of these studies were in people with constipation-predominant IBS (Cook 1990; Rees 2005); three did not specify the type of IBS (Arthurs 1983; Fowlie 1992; Ritchie 1979) and the remainder had mixed IBS types. Therefore the studies were not stratified by IBS type. Similarly, there was too little information to separate by severity, post-infective cause or bloating status.

Where outcomes were measured at different times during the study, we took the end-study results unless there were significant numbers of withdrawals or problems with compliance. Therefore, for the Kruis (1986) study we took the values at four weeks. The results in Rees (2005) were collected between week 8 and week 12 (11 people were assessed at week 8; six at week 9; three at week 10; one at week 11, and; one at week 12).

1. Global symptoms

a) Number of people with improvement in global symptoms

Nine studies with 545 participants reported this outcome. Overall the relative risk was 1.18 (95% CI 1.03 to 1.35), i.e. statistically significant, in favour of fibre.

Figure 1

Review: IBS MW 080507
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 01 Global improvement of IBS symptoms

Study or sub-category	Fibre n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 parallel studies					
Rees 2005	8/12	7/10		5.19	0.95 [0.54, 1.68]
Ritchie 1979	23/48	13/48		8.84	1.77 [1.02, 3.07]
Kruis 1986(usldiet)	20/40	16/40		10.87	1.25 [0.77, 2.04]
Soltoft 1976	15/29	15/23		11.37	0.79 [0.50, 1.26]
Longstreth 1981	20/26	24/34		14.14	1.09 [0.81, 1.47]
Arthurs 1983	29/40	24/38		16.73	1.15 [0.84, 1.56]
Lucey 1987	11/14	10/14		6.80	1.10 [0.72, 1.69]
Prior 1987	33/40	21/40		14.27	1.57 [1.13, 2.18]
Fowlie 1992	15/25	17/24		11.79	0.85 [0.56, 1.28]
Subtotal (95% CI)	**274**	**271**		**100.00**	**1.18 [1.03, 1.35]**

Total events: 174 (Fibre), 147 (Placebo)
Test for heterogeneity: Chi² = 11.34, df = 8 (P = 0.18), I² = 29.5%
Test for overall effect: Z = 2.33 (P = 0.02)

| **Total (95% CI)** | **274** | **271** | | **100.00** | **1.18 [1.03, 1.35]** |

Total events: 174 (Fibre), 147 (Placebo)
Test for heterogeneity: Chi² = 11.34, df = 8 (P = 0.18), I² = 29.5%
Test for overall effect: Z = 2.33 (P = 0.02)

0.2   0.5   1   2   5
Favours Placebo   Favours Fibre

Subgroup analysis into soluble and insoluble fibres (Figure 2) gave some suggestion that soluble fibre was more effective than insoluble, however, this conclusion was fairly reliant on the Prior and Whorwell (1987) study, which had some attrition bias and was analysed using the last measurement carried forward method. A sensitivity analysis without Prior and Whorwell (1987), Longstreth (1981), Rees (2005 - attrition bias) and Kruis (1986 - which did not have a placebo comparator) showed little difference in global improvement between fibre and placebo overall, although the results for soluble fibre were still significant (Figure 3a).

Figure 2:

Review: IBS MW 080507
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 02 Global improvement by fibre type

Study or sub-category	Fibre n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Soluble Fibre					
Ritchie 1979	23/48	13/48		8.86	1.77 [1.02, 3.07]
Longstreth 1981	20/26	24/34		14.18	1.09 [0.81, 1.47]
Arthurs 1983	29/40	24/38		16.78	1.15 [0.84, 1.56]
Prior 1987	33/40	21/40		14.32	1.57 [1.13, 2.18]
Subtotal (95% CI)	**154**	**160**		**54.15**	**1.35 [1.12, 1.61]**

Total events: 105 (Fibre), 82 (Placebo)
Test for heterogeneity: Chi² = 4.71, df = 3 (P = 0.19), I² = 36.3%
Test for overall effect: Z = 3.24 (P = 0.001)

02 Insoluble Fibre					
Rees 2005	8/12	7/10		5.21	0.95 [0.54, 1.68]
Kruis 1986(usldiet)	20/40	16/40		10.91	1.25 [0.77, 2.04]
Soltoft 1976	15/32	15/27		11.09	0.84 [0.51, 1.39]
Lucey 1987	11/14	10/14		6.82	1.10 [0.72, 1.69]
Subtotal (95% CI)	**98**	**91**		**34.03**	**1.04 [0.81, 1.35]**

Total events: 54 (Fibre), 48 (Placebo)
Test for heterogeneity: Chi² = 1.37, df = 3 (P = 0.71), I² = 0%
Test for overall effect: Z = 0.31 (P = 0.75)

03 Mixed fibre					
Fowlie 1992	15/25	17/24		11.83	0.85 [0.56, 1.28]
Subtotal (95% CI)	**25**	**24**		**11.83**	**0.85 [0.56, 1.28]**

Total events: 15 (Fibre), 17 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

| **Total (95% CI)** | **277** | **275** | | **100.00** | **1.18 [1.03, 1.36]** |

Total events: 174 (Fibre), 147 (Placebo)
Test for heterogeneity: Chi² = 10.29, df = 8 (P = 0.25), I² = 22.2%
Test for overall effect: Z = 2.39 (P = 0.02)

0.1 0.2   0.5   1   2   5   10
Favours Placebo   Favours Fibre

Figure 3: Sensitivity analysis


```
Review:         IBS MW 080507
Comparison:     01 Fibre vs Placebo (all IBS types)
Outcome:        02 Global improvement by fibre type

Study                    Fibre          Placebo        RR (fixed)      Weight       RR (fixed)
or sub-category          n/N            n/N            95% CI          %            95% CI

01 Soluble Fibre
 Ritchie 1979            23/48          13/48                          16.00        1.77 [1.02, 3.07]
 Arthurs 1983            29/40          24/38                          30.30        1.15 [0.84, 1.56]
 Subtotal (95% CI)       88             86                             46.31        1.36 [1.02, 1.81]
Total events: 52 (Fibre), 37 (Placebo)
Test for heterogeneity: Chi² = 2.05, df = 1 (P = 0.15), I² = 51.2%
Test for overall effect: Z = 2.13 (P = 0.03)

02 Insoluble Fibre
 Soltoft 1976            15/32          15/27                          20.03        0.84 [0.51, 1.39]
 Lucey 1987              11/14          10/14                          12.31        1.10 [0.72, 1.69]
 Subtotal (95% CI)       46             41                             32.34        0.94 [0.67, 1.33]
Total events: 26 (Fibre), 25 (Placebo)
Test for heterogeneity: Chi² = 0.69, df = 1 (P = 0.41), I² = 0%
Test for overall effect: Z = 0.35 (P = 0.73)

03 Mixed fibre
 Fowlie 1992             15/25          17/24                          21.35        0.85 [0.56, 1.28]
 Subtotal (95% CI)       25             24                             21.35        0.85 [0.56, 1.28]
Total events: 15 (Fibre), 17 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

Total (95% CI)           159            151                            100.00       1.12 [0.92, 1.35]
Total events: 93 (Fibre), 79 (Placebo)
Test for heterogeneity: Chi² = 5.67, df = 4 (P = 0.22), I² = 29.5%
Test for overall effect: Z = 1.11 (P = 0.27)

        0.1   0.2      0.5    1    2    5    10
              Favours Placebo    Favours Fibre
```

Sensitivity analysis by method of ingestion

A further sensitivity analysis was carried out on the studies that were not at risk of bias, to investigate if there was an effect of supplementary fibre compared with dietary fibre. This was examined in a subgroup analysis (Figure 3b). There was heterogeneity (I^2=58%, p=0.09) in the supplement group, which was probably caused by different types of fibre.

Figure 3b


```
Review:         IBS
Comparison:     01 Fibre vs Placebo (all IBS types)
Outcome:        28 Global improvement of IBS symptoms - supplemental or dietary

Study                    Fibre          Placebo        RR (fixed)      Weight       RR (fixed)
or sub-category          n/N            n/N            95% CI          %            95% CI

01 supplemental
 Arthurs 1983            29/40          24/38                          30.13        1.15 [0.84, 1.56]
 Fowlie 1992             15/25          17/24                          21.23        0.85 [0.56, 1.28]
 Richie 1978             23/48          13/48                          15.91        1.77 [1.02, 3.07]
 Subtotal (95% CI)       113            110                            67.28        1.20 [0.95, 1.52]
Total events: 67 (Fibre), 54 (Placebo)
Test for heterogeneity: Chi² = 4.76, df = 2 (P = 0.09), I² = 58.0%
Test for overall effect: Z = 1.52 (P = 0.13)

03 dietary (using foods)
 Lucey 1987              11/14          10/14                          12.24        1.10 [0.72, 1.69]
 Soltoft 1976            15/29          15/23                          20.48        0.79 [0.50, 1.26]
 Subtotal (95% CI)       43             37                             32.72        0.91 [0.66, 1.25]
Total events: 26 (Fibre), 25 (Placebo)
Test for heterogeneity: Chi² = 1.10, df = 1 (P = 0.30), I² = 8.8%
Test for overall effect: Z = 0.59 (P = 0.56)

Total (95% CI)           156            147                            100.00       1.10 [0.91, 1.34]
Total events: 93 (Fibre), 79 (Placebo)
Test for heterogeneity: Chi² = 6.47, df = 4 (P = 0.17), I² = 38.1%
Test for overall effect: Z = 1.02 (P = 0.31)

        0.1   0.2      0.5    1    2    5    10
              Favours Placebo    Favours Fibre
```

b) Number of people with deterioration in global symptoms

Three studies reported this outcome, and included 140 participants (Figure 4). The numbers of events were few and there was too much uncertainty (wide confidence interval) to draw conclusions.

Figure 4:

Review: IBS 31 August 2006
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 15 Global deterioration of IBS symptoms

Study or sub-category	Fibre n/N	Placebo n/N	Peto OR 95% CI	Weight %	Peto OR 95% CI
01 parallel studies					
Soltoft 1976	1/29	0/23		13.75	6.01 [0.12, 310.87]
Longstreth 1981	1/26	1/34		26.91	1.32 [0.08, 22.13]
Lucey 1987	2/14	3/14		59.35	0.63 [0.09, 4.18]
Subtotal (95% CI)	69	71		100.00	1.04 [0.24, 4.51]

Total events: 4 (Fibre), 4 (Placebo)
Test for heterogeneity: Chi² = 1.06, df = 2 (P = 0.59), I² = 0%
Test for overall effect: Z = 0.06 (P = 0.95)

0.1 0.2 0.5 1 2 5 10
Favours Fibre   Favours Placebo

c) Global symptom score (mean)

This outcome was recorded by four studies (Cook 1990; Fowlie 1992; Longstreth 1981; Lucey 1987), and different scales were used. Fowlie (1992) did not give scores for the two groups and Cook (1990) was a crossover design (and had some attrition bias). In view of the different scales it was not possible to meta-analyse the parallel and crossover studies using the generic inverse variance method, so the two remaining parallel studies and the crossover study were analysed separately using the standardised mean difference. The results were inconclusive (Figure 5).

Figure 5

Review: IBS 31 August 2006
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 05 Global Symptom Score (mean) Low score=good

Study or sub-category	Fibre N	Mean (SD)	Placebo N	Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Parallel							
Longstreth 1981	26	0.36 (0.56)	34	0.23 (0.35)		68.13	0.28 [-0.23, 0.80]
Lucey 1987	14	1.00 (4.77)	14	3.50 (6.47)		31.87	-0.43 [-1.18, 0.32]
Subtotal (95% CI)	40		48			100.00	0.06 [-0.37, 0.48]

Test for heterogeneity: Chi² = 2.35, df = 1 (P = 0.13), I² = 57.4%
Test for overall effect: Z = 0.26 (P = 0.79)

02 Crossover							
Cook 1990 (Constip)	9	16.90 (3.30)	9	15.00 (4.50)		100.00	0.46 [-0.48, 1.40]
Subtotal (95% CI)	9		9			100.00	0.46 [-0.48, 1.40]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.96 (P = 0.34)

-4 -2 0 2 4
Favours Fibre   Favours Placebo

2. Individual symptoms

a) Pain

The following studies measured pain:

   i. Number of people with no pain: one study (Prior and Whorwell 1987)

   ii. Number of people with less pain: three studies (Kruis 1986)

   iii. Pain score (final): four studies (Cook 1990; Fowlie et al 1992; Longstreth et al 1981; Manning 1977);

   vi. Two studies reported pain severity at the end of the study (Cook 1990; Fowlie 1992)

   iv. Two studies reported pain severity from daily diary readings (Longstreth 1981; Manning 1977).

Figure 6 shows the number of people with less pain and the number of people with no pain, in two single studies. The confidence intervals were too wide to draw conclusions.

Figure 6

Review: IBS 31 August 2006
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 14 Pain (No of patients)

Study or sub-category	Fibre n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Number with no pain (soluble fibre)					
Prior 1987	7/40	3/40		100.00	2.33 [0.65, 8.39]
Subtotal (95% CI)	40	40		100.00	2.33 [0.65, 8.39]
Total events: 7 (Fibre), 3 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.30 (P = 0.19)					
02 Number with less pain (insoluble fibre)					
Kruis 1986(usldiet)	5/40	11/40		100.00	0.45 [0.17, 1.19]
Subtotal (95% CI)	40	40		100.00	0.45 [0.17, 1.19]
Total events: 5 (Fibre), 11 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.61 (P = 0.11)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours Fibre

Fowlie (1992) only gave the difference in the mean change score from baseline and its 95%CI, which was 1 (95%CI -1.5, 4), i.e. not statistically significant.

Combining the other three studies recording pain score, using the standardised mean difference (Figure 7), showed little difference between fibre and placebo, but the data was limited.

Figure 7

Review: IBS 31 August 2006
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 06 Pain Score (Low - good)

Study or sub-category	N	Fibre Mean (SD)	N	Placebo Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Parallel studies							
Longstreth 1981	26	0.59(0.51)	34	0.62(0.52)		73.01	-0.06 [-0.57, 0.45]
Manning 1977	12	1.95(1.76)	10	1.75(1.56)		26.99	0.12 [-0.73, 0.96]
Subtotal (95% CI)	38		44			100.00	-0.01 [-0.45, 0.43]
Test for heterogeneity: Chi² = 0.12, df = 1 (P = 0.73), I² = 0%							
Test for overall effect: Z = 0.05 (P = 0.96)							
03 Crossover studies							
Cook 1990 (Constip)	9	2.90(0.60)	9	2.30(1.20)		100.00	0.60 [-0.35, 1.55]
Subtotal (95% CI)	9		9			100.00	0.60 [-0.35, 1.55]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.24 (P = 0.21)							

-4 -2 0 2 4
Favours Placebo   Favours Fibre

b) Bloating

Only one study (Prior and Whorwell 1987) reported bloating (Figure 8). This showed that statistically significantly more people had bloating when they took fibre (soluble) compared with placebo. It should be noted that this was a last measurement carried forward analysis, but that a large proportion withdrew from the study in the ispaghula group.

Figure 8


```
Review:        IBS 31 August 2006
Comparison:    01 Fibre vs Placebo (all IBS types)
Outcome:       13 Bloating/Gas reduction

Study                       Fibre        Placebo      RR (fixed)      Weight     RR (fixed)
or sub-category             n/N          n/N          95% CI          %          95% CI

01 No of patients with bloating
 Prior 1987                 40/40        33/40                        100.00     1.21 [1.05, 1.40]
Subtotal (95% CI)           40           40                          100.00     1.21 [1.05, 1.40]
Total events: 40 (Fibre), 33 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.64 (P = 0.008)

Total (95% CI)              40           40                          100.00     1.21 [1.05, 1.40]
Total events: 40 (Fibre), 33 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.64 (P = 0.008)

                                           0.1  0.2  0.5  1   2   5   10
                                           Favours Fibre    Favours Placebo
```

c) Combined bloating and flatulence score

One study reported this outcome (Longstreth 1981). The results showed a small non-significant difference (0.31 on a scale of 0 to 4) in favour of placebo. We noted that this study had attrition bias.

Figure 9


```
Review:        IBS 31 August 2006
Comparison:    01 Fibre vs Placebo (all IBS types)
Outcome:       11 Bloating and flatus combined

Study                 Fibre            Placebo         WMD (fixed)     Weight     WMD (fixed)
or sub-category    N  Mean (SD)     N  Mean (SD)       95% CI          %          95% CI

 Longstreth 1981   26 1.19(0.76)    34 0.88(0.70)                      100.00     0.31 [-0.07, 0.69]

Total (95% CI)     26              34                                  100.00     0.31 [-0.07, 0.69]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.62 (P = 0.11)

                                           -4    -2    0    2    4
                                           Favours Fibre   Favours Placebo
```
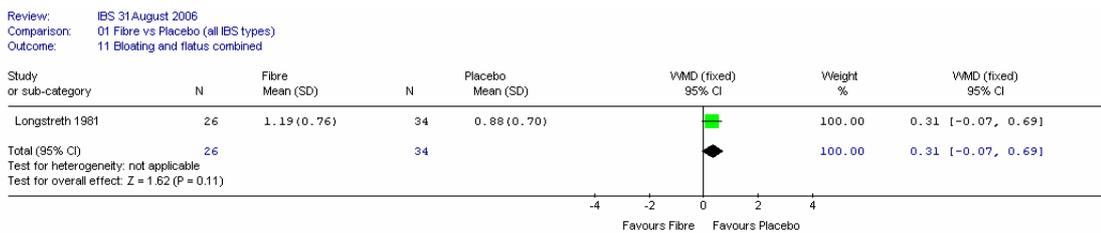
d) Bowel habits

i. Number of people with improved bowel habits

Two studies, with 106 participants, recorded the number of people with improved bowel habits (Kruis 1986; Manning 1977). Meta-analysis showed some heterogeneity between studies and a wide confidence interval. Each study was a comparison with a non-placebo comparator (low fibre or usual diet).

Figure 10


```
Review:        IBS 31 August 2006
Comparison:    01 Fibre vs Placebo (all IBS types)
Outcome:       20 Improvement in bowel habit

Study                       Fibre        Placebo      RR (fixed)      Weight     RR (fixed)
or sub-category             n/N          n/N          95% CI          %          95% CI

01 parallel studies
 Kruis 1986(usldiet)        9/40         10/40                        82.28      0.90 [0.41, 1.98]
 Manning 1977               7/14         2/12                         17.72      3.00 [0.76, 11.80]
Subtotal (95% CI)           54           52                          100.00     1.27 [0.66, 2.46]
Total events: 16 (Fibre), 12 (Placebo)
Test for heterogeneity: Chi² = 2.25, df = 1 (P = 0.13), I² = 55.6%
Test for overall effect: Z = 0.72 (P = 0.47)

                                           0.1  0.2  0.5  1   2   5   10
                                           Favours Placebo  Favours Fibre
```

ii. Stool score (aggregate)

Longstreth (1981) reported the number of normal stools per week. The confidence interval was fairly wide (-2.0 to 2.6), but there was little difference between fibre and placebo. We noted that this study had attrition bias.

Figure 11

Review: IBS 31 August 2006
Comparison: 01 Fibre vs Placebo (all IBS types)
Outcome: 10 Stool score - number of normal stools

Study or sub-category	N	Fibre Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 parallel							
Longstreth 1981	26	6.30(4.50)	34	6.00(4.60)		100.00	0.30 [-2.02, 2.62]
Subtotal (95% CI)	26		34			100.00	0.30 [-2.02, 2.62]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.25 (P = 0.80)

-10 -5 0 5 10
Favours Placebo   Favours Fibre

B. Fibre type 1 versus Fibre type 2

B1. Insoluble versus soluble fibre

Two studies compared insoluble and soluble fibre: Parisi (2002) compared wheat bran (insoluble; 30g/day) with guar gum (soluble; 5g/day) in people with a mixture of IBS types; Ritchie (1980) compared coarse natural bran (insoluble; 20g/day) with ispaghula (soluble; Fibogel 7g/day).

1. Global outcomes

a) Global improvement of symptoms

Meta-analysis of two studies (Parisi 2002; Ritchie 1980) in 281 people, found a statistically significant increase in the number of people reporting improved global symptoms in favour of the soluble fibre (RR 0.61, 95% CI 0.51 to 0.73), with no heterogeneity. This corresponded to a number needed to harm of 3 (95%CI 2, 4), for a soluble group rate of 69 to 88%.

Figure 12

Review: IBS MW 080507
Comparison: 02 Fibre Type 1 vs Fibre Type 2
Outcome: 01 Global Improvement of symptoms

Study or sub-category	Insoluble n/N	Soluble n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Insoluble vs soluble fibre					
Ritchie 1980	23/48	33/48		28.84	0.70 [0.49, 0.99]
Parisi 2002	47/93	81/92		71.16	0.57 [0.46, 0.71]
Subtotal (95% CI)	141	140		100.00	0.61 [0.51, 0.73]

Total events: 70 (Insoluble), 114 (Soluble)
Test for heterogeneity: Chi² = 0.86, df = 1 (P = 0.35), I² = 0%
Test for overall effect: Z = 5.29 (P < 0.00001)

0.1 0.2 0.5 1 2 5 10
Soluble   Insoluble

b) Global deterioration in symptoms

One study (Parisi 2002) showed a wide confidence interval for this outcome and conclusions could not be drawn.

Figure 13

Review: IBS 31 August 2006
Comparison: 02 Fibre Type 1 vs Fibre Type 2
Outcome: 02 Global deterioration of symptoms

Study or sub-category	Insoluble n/N	Soluble n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Insoluble vs soluble					
Parisi 2002	5/93	1/92		100.00	4.95 [0.59, 41.52]
Subtotal (95% CI)	93	92		100.00	4.95 [0.59, 41.52]
Total events: 5 (Insoluble), 1 (Soluble)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.47 (P = 0.14)					

0.1 0.2 0.5 1 2 5 10
Insoluble   Soluble

2. Individual symptoms

a) Pain

One study (Parisi 2002) showed little difference between the interventions for the number of people with pain.

Figure 14

Review: IBS 31 August 2006
Comparison: 02 Fibre Type 1 vs Fibre Type 2
Outcome: 03 Number of patients with pain (mild, moderate and severe)

Study or sub-category	Insoluble n/N	Soluble n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Insoluble vs soluble					
Parisi 2002	56/93	58/92		100.00	0.96 [0.76, 1.20]
Subtotal (95% CI)	93	92		100.00	0.96 [0.76, 1.20]
Total events: 56 (Insoluble), 58 (Soluble)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.40 (P = 0.69)					

0.1 0.2 0.5 1 2 5 10
Insoluble   Soluble

b) Bowel habits

There was no significant difference in the number of people with improved bowel habits.

Figure 15

Review: IBS
Comparison: 02 Fibre Type 1 vs Fibre Type 2
Outcome: 07 Number of patients with improved bowel habit

Study or sub-category	Type 1 fibre n/N	Soluble n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Insoluble vs soluble (mixed IBS patients)					
Parisi 2002	23/93	26/92		100.00	0.88 [0.54, 1.42]
Subtotal (95% CI)	93	92		100.00	0.88 [0.54, 1.42]
Total events: 23 (Type 1 fibre), 26 (Soluble)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.54 (P = 0.59)					
Total (95% CI)	93	92		100.00	0.88 [0.54, 1.42]
Total events: 23 (Type 1 fibre), 26 (Soluble)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.54 (P = 0.59)					

0.1 0.2 0.5 1 2 5 10
type 1 fibre   Soluble

B2. Mixed fibre versus soluble fibre

Tarpila (2004) compared 6 to 24g/day flax seed (mixed fibre: 33% insoluble, 11% soluble, 20% flaxseed oil) with 6 to 24g/day psyllium (soluble), in 55 people with IBS-C.

a) Bloating

There were significantly more people with a reduction in bloating for the mixed fibres (flax seeds) group, compared to psyllium.

Figure 16a

```
Review:       IBS 31 August 2006
Comparison:   02 Fibre Type 1 vs Fibre Type 2
Outcome:      04 Number of patients with less bloating

Study                           Mixed           Soluble              RR (fixed)         Weight          RR (fixed)
or sub-category                  n/N              n/N                  95% CI              %              95% CI

01 Mixed vs soluble
  Tarpila 2004                   24/26           15/29                               ▪       100.00      1.78 [1.23, 2.58]
  Subtotal (95% CI)               26              29                               ◆        100.00      1.78 [1.23, 2.58]
Total events: 24 (Mixed), 15 (Soluble)
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.08 (P = 0.002)

                                                         0.1   0.2   0.5   1   2   5   10
                                                                    Soluble    Mixed
```

The number of people with more bloating also significantly favoured the mixed fibre, although the confidence interval was very wide.

Figure 16b

```
Review:       IBS 31 August 2006
Comparison:   02 Fibre Type 1 vs Fibre Type 2
Outcome:      05 Number of patients with more bloating

Study                           Mixed           Soluble              Peto OR            Weight          Peto OR
or sub-category                  n/N              n/N                  95% CI              %              95% CI

01 Mixed vs soluble
  Tarpila 2004                    0/26            5/29             ◀▪                     100.00      0.13 [0.02, 0.80]
  Subtotal (95% CI)               26              29              ◆                       100.00      0.13 [0.02, 0.80]
Total events: 0 (Mixed), 5 (Soluble)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.20 (P = 0.03)

                                                         0.1   0.2   0.5   1   2   5   10
                                                                    Mixed    Soluble
```

b) Bowel habits

There was no significant difference in the number of people with improved bowel habits.

Figure 17

```
Review:       IBS
Comparison:   02 Fibre Type 1 vs Fibre Type 2
Outcome:      08 Number of patients with improved bowel habit

Study                         Type 1 fibre      Soluble              RR (fixed)         Weight          RR (fixed)
or sub-category                  n/N              n/N                  95% CI              %              95% CI

02 Mixed vs soluble (constipation predominant IBS)
  Tarpila 2004                   24/26           22/29                              ▪        100.00      1.22 [0.96, 1.54]
  Subtotal (95% CI)               26              29                              ◆         100.00      1.22 [0.96, 1.54]
Total events: 24 (Type 1 fibre), 22 (Soluble)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.65 (P = 0.10)

Total (95% CI)                    26              29                              ◆         100.00      1.22 [0.96, 1.54]
Total events: 24 (Type 1 fibre), 22 (Soluble)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.65 (P = 0.10)

                                                         0.1   0.2   0.5   1   2   5   10
                                                                  type 1 fibre    Soluble
```

C. Mixed fibre 1 versus mixed fibre 2

One study (Fielding 1984) compared 40g of mixed fibre diet with different proportions of cereal and fruit/vegetables 75% cereal versus 25% cereal. The study recorded a state of well being score and individual symptom outcomes.

1. Number of people with an improved state of well being

There was little difference between interventions, although the confidence interval was fairly wide.

Figure 18

Review: IBS 31 August 2006
Comparison: 04 Mixed Fibre type 1 (75% cereal / 25% fruit/veg) vs Mixed Fibre type 2 (25% cereal / 75% fruit/veg)
Outcome: 06 Number with improved / much improved state of well being

Study or sub-category	High cereal n/N	High fruit/veg n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Fielding 1984	19/28	18/27		100.00	1.02 [0.70, 1.47]
Total (95% CI)	28	27		100.00	1.02 [0.70, 1.47]

Total events: 19 (High cereal), 18 (High fruit/veg)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.09 (P = 0.93)

0.1 0.2 0.5 1 2 5 10
Favours Hi fruit/veg   Favours high cereal

2. Individual symptoms

a) Number of people with less pain

There was little difference in pain incidence between the two types of mixed fibre.

Figure 19

Review: IBS 31 August 2006
Comparison: 04 Mixed Fibre type 1 (75% cereal / 25% fruit/veg) vs Mixed Fibre type 2 (25% cereal / 75% fruit/veg)
Outcome: 01 Number of patients with less pain

Study or sub-category	High cereal n/N	High fruit/veg n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Fielding 1984	20/28	19/27		100.00	1.02 [0.72, 1.42]
Total (95% CI)	28	27		100.00	1.02 [0.72, 1.42]

Total events: 20 (High cereal), 19 (High fruit/veg)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.09 (P = 0.93)

0.1 0.2 0.5 1 2 5 10
Favours Hi fruit/veg   Favours high cereal

b) Number of people with improved bowel habit

There was little difference between interventions.

Figure 20

Review: IBS 31 August 2006
Comparison: 04 Mixed Fibre type 1 (75% cereal / 25% fruit/veg) vs Mixed Fibre type 2 (25% cereal / 75% fruit/veg)
Outcome: 05 Number of pts with improved bowel habit

Study or sub-category	High cereal n/N	High fruit/veg n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Fielding 1984	18/28	21/27		100.00	0.83 [0.59, 1.16]
Total (95% CI)	28	27		100.00	0.83 [0.59, 1.16]

Total events: 18 (High cereal), 21 (High fruit/veg)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.09 (P = 0.27)

0.1 0.2 0.5 1 2 5 10
Favours Hi Fruit/veg   Favours High cereal

D. Fibre dose 1 versus fibre dose 2

Two studies compared different doses of fibre (Aller 2004; Parisi 2005). In the former, the comparison was 30.5 versus 10.4g /day of mixed fibre over 12 weeks (i.e. above versus below the 18g/day threshold). The latter compared 10 and 5g/day of partially hydrolysed guar gum over 12 weeks, which was then followed up for a further 12 weeks.

a) Global symptom score

There was little difference between interventions in a single study in 96 patients, and the further 12 weeks follow-up did not change this conclusion

Figure 21
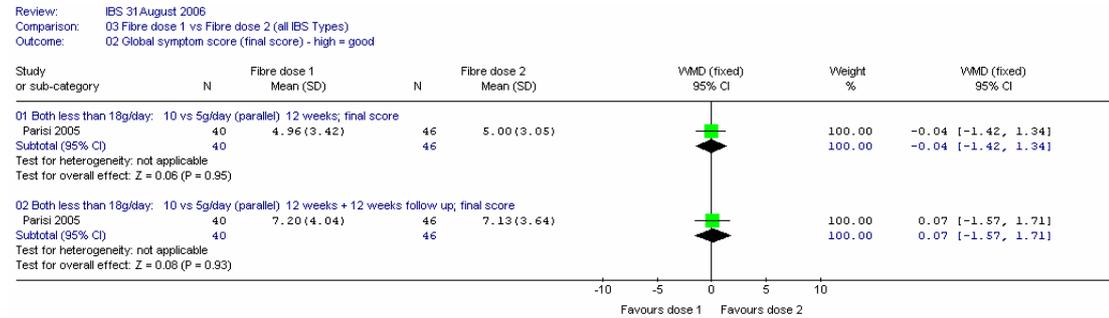
b) Pain score

There was a small, non-significant difference between interventions, favouring the lower dose of soluble fibre in Parisi (2005) at 12 weeks, which decreases to zero after a further 12 weeks. There was no significant difference in the two doses (above and below the threshold) for the Aller (2004) study.

Figure 22
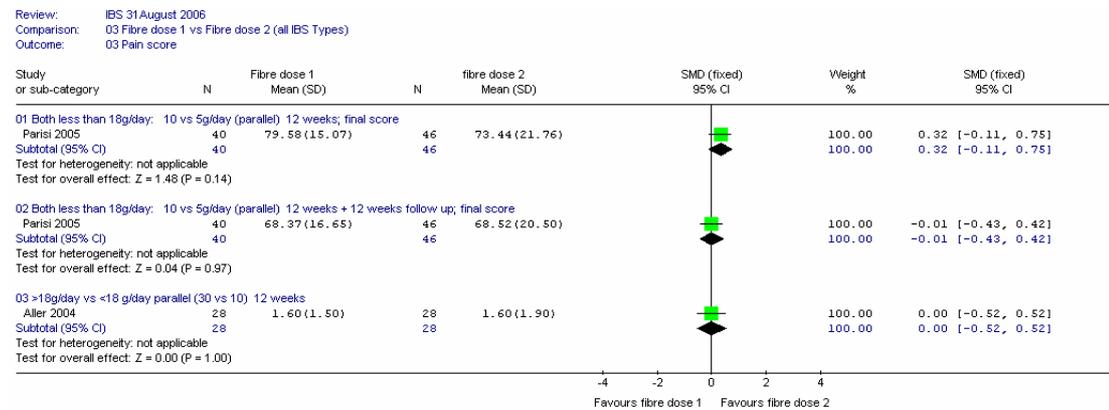
c) General bloating and flatus score

There is little difference between dose levels in either study.

Figure 23

Review: IBS 31 August 2006
Comparison: 03 Fibre dose 1 vs Fibre dose 2 (all IBS Types)
Outcome: 07 Bloating and flatus score

Study or sub-category	Fibre dose 1 N	Mean (SD)	fibre dose 2 N	Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Both less than 18g/day: 10 vs 5g/day (parallel) 12 weeks; final score							
Parisi 2005	40	2.90(1.72)	46	2.71(1.62)		100.00	0.11 [-0.31, 0.54]
Subtotal (95% CI)	40		46			100.00	0.11 [-0.31, 0.54]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.52 (P = 0.60)							
03 >18g/day vs <18 g/day parallel (30 vs 10) 12 weeks							
Aller 2004	28	1.50(0.90)	28	1.30(0.90)		100.00	0.22 [-0.31, 0.74]
Subtotal (95% CI)	28		28			100.00	0.22 [-0.31, 0.74]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.82 (P = 0.41)							

-10   -5   0   5   10
Favours fibre dose 1   Favours fibre dose 2

d) Bowel scores

There is little difference between doses for the Aller (2004) study.

Figure 24

Review: IBS 31 August 2006
Comparison: 03 Fibre dose 1 vs Fibre dose 2 (all IBS Types)
Outcome: 08 Stool score aggregate

Study or sub-category	Fibre dose 1 N	Mean (SD)	fibre dose 2 N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
03 >18g/day vs <18 g/day parallel (30 vs 10) 12 weeks							
Aller 2004	28	1.00(0.80)	28	1.30(1.30)		100.00	-0.30 [-0.87, 0.27]
Subtotal (95% CI)	28		28			100.00	-0.30 [-0.87, 0.27]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.04 (P = 0.30)							

-10   -5   0   5   10
Favours fibre dose 1   Favours fibre dose 2

E. Fibre plus another intervention versus another intervention alone

Two studies (Dettmar 1999; Chapman 1990) assessed ispaghula plus mebeverine (anti-spasmodic) versus mebeverine plus high fibre dietary advice. Each study reported the number of people improved in terms of abdominal pain, and in terms of improvements in bowel habit, at 4 weeks.

Figure 25

Review: IBS MW 080507
Comparison: 38 Fibre + antispasmodic vs. antispasmodic + high fibre diet advice
Outcome: 01 Improvement in abdominal pain (no. of patients)

Study or sub-category	Fibre+mebeverine n/N	Mebev+diet advice n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Dettmar 1999	34/56	24/54		41.57	1.37 [0.95, 1.97]
Chapman 1990	33/50	34/49		58.43	0.95 [0.72, 1.25]
Total (95% CI)	106	103		100.00	1.12 [0.90, 1.40]
Total events: 67 (Fibre+mebeverine), 58 (Mebev+diet advice)					
Test for heterogeneity: Chi² = 2.54, df = 1 (P = 0.11), I² = 60.6%					
Test for overall effect: Z = 1.03 (P = 0.30)					

0.2   0.5   1   2   5
Favours meb+diet   Favours fibre+meb

Figure 26

Review: IBS MW 080507
Comparison: 38 Fibre + antispasmodic vs. antispasmodic + high fibre diet advice
Outcome: 02 Improvement in bowel function (no. of patients)

Study or sub-category	Fibre+mebeverine n/N	Mebev + diet advice n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Dettmar 1999	31/56	18/54		41.10	1.66 [1.06, 2.59]
Chapman 1990	25/50	26/49		58.90	0.94 [0.64, 1.38]
Total (95% CI)	106	103		100.00	1.24 [0.93, 1.65]
Total events: 56 (Fibre+mebeverine), 44 (Mebev + diet advice)					
Test for heterogeneity: Chi² = 3.63, df = 1 (P = 0.06), I² = 72.5%					
Test for overall effect: Z = 1.45 (P = 0.15)					

0.2   0.5   1   2   5
Favours meb+diet   Favours fibre+mebev

F. Protective effects of fibre for the prevention of colorectal adenomas and carcinomas, coronary heart disease and breast cancer

1. Colorectal Cancer

The role of diet in the development of colorectal cancer has long been hypothesised. Although there are many studies investigating the relationship between diet and colorectal cancer, the exact relationship remains unclear.

In the 1970's epidemiological studies first suggested an inverse relationship between foods rich in dietary fibre and the incidence of colorectal cancer. However, many of these studies were case-control designs, which were subject to selection bias and recall bias. Evidence from two large cohort studies (the Nurses Health Study in 88,757 women and the Health Professionals' Follow-up Study in 47,325 men) found that dietary fibre had no significant effect on the risk of colorectal cancer. A further cohort study in 61,463 people, however, found a weak association between fruit consumption and reduction in risk, but no association between cereal intake and risk. More recently a Cochrane review (Asano 2002) of five large randomised trials showed no significant protective effect of fibre on the development of colorectal adenomas within two to four years.

2. Coronary Heart Disease (CHD)

Prior to 2000, a number of reviews investigated the relationship between diet and CHD and Stroke. Since 2000 several studies have concentrated on the relationship between wholegrain dietary intake and CHD, and there is a body of evidence to support a 20 to 40% risk reduction of CHD for those who consume a diet rich in wholegrains compared to those who do not. However many studies have not shown an independent effect of fibre alone. The only RCT in secondary prevention of CHD that advised participants to eat more cereal fibre showed no reduction in the reinfarction rate, but there was no data on primary prevention. There was strong evidence to suggest that wheat fibre does not lower cholesterol.

Cereal products provide around 30% of total energy intake in British adults. Several nutrients contained in cereals have the potential to reduce the risk factors for CHD (linoleic acid, fibre vitamin E, selenium and folate, phytoestrogens of the lignan family, phenolic acids with antioxidant properties). It should be noted that some processed cereal foods are high in salt and could contribute to raising blood pressure.

Over 40 human trials have shown that oat fibre tends to lower plasma total and LDL cholesterol but wheat fibre does not. Rice bran and barley may also lower cholesterol but intake of barley tends to be too low to have an effect.

There is no clear association, negative or positive, between total cereal consumption and CHD.

The intake of wholegrain foods may protect against heart disease and stroke but the exact mechanism is not clear. Fibre, magnesium, folate and vitamins B6 and E may be important.

The Joint Health Claims Initiative states that the evidence supports the association between regular consumption of wholegrains and a healthy heart but that it is insufficient to demonstrate cause and effect.

3. Breast Cancer

In the mid 1980's the role of fibre in breast cancer was suggested. There have been many studies including case control studies in several populations reporting a reduced risk for breast cancer for individuals with a high intake of dietary fibre. Other studies were contradictory and the positive effect of fibre for breast cancer risk reduction was not confirmed by prospective cohort studies in the US (Holmes 2004; Terry 2002). A recent study (Cade 2007) investigated the relationship between dietary fibre intake and breast cancer in a large cohort of British women. The conclusions were that total fibre of more than 30g/day was protective against breast cancer in pre-menopausal women relative to an intake of less than 20g/day, but was not significant in post-menopausal women. After assessing this study we had some reservations.

- The population were highly selected and not necessarily representative
- Lower levels of fibre intake were not protective and subgroup analysis according to fruit, vegetable and cereal fibre showed no significant effect
- There was no data available on the effects of soluble and insoluble fibre (Cade, personal communication to GDG).

A recent large RCT (Pierce 2007) in 3088 women investigated the effect on prognosis, following treatment for breast cancer, of a diet very high in vegetables, fruit and fibre and low in fat, plus telephone counselling, in comparison to dietary guidelines. The trial found no reduction of breast cancer events (recurrence or new primary) or any improvement in survival over a 7.5 year follow-up period.

There is currently insufficient evidence to demonstrate a causal relationship between total cereal consumption and breast cancer prevention. Studies have not investigated the specific effects of soluble and insoluble fibre.

In summary, the protective effects of fibre for the prevention of colorectal adenomas and carcinomas, coronary heart disease and breast cancer remain uncertain.

GDG DISCUSSION

The GDG discussed the use of fibre at some length, also taking into account a survey of the use of bran in people with IBS in primary and secondary care (Miller 2006). This paper suggested that bran was not especially effective in primary care, improving symptoms in 27/100 people with IBS, with 22/100 reporting an exacerbation of symptoms. This was significantly fewer than found in people in secondary care. The effects of soluble fibre were similar in both primary care and secondary care. The study highlighted the issues of extrapolating the response to treatment in IBS from different care settings.

The GDG unanimously agreed that the practice in primary care of recommending high fibre diets to people with IBS should cease. They suggested that GPs should investigate the person's usual fibre intake with a view to modifying fibre levels to suit the symptom profile and they should monitor the person's response to dietary modification. GDG consensus was that wheat bran should not be recommended for people with IBS as it is ineffective in the management of symptoms and may even increase symptoms in some people. It may be preferential for the dietary fibre intake to be closer to 12g/day rather than 18g/day. If an increase in fibre were needed, this should be in the form of soluble fibre. Although the RCT evidence for the beneficial effect of soluble fibre was based on trials using supplements such as ispaghula, the GDG wished to give an example of a dietary food that is high in soluble fibre in their recommendation. They also took into consideration the protective effect of oats on cholesterol levels. GDG consensus was that oats should be given as an example of a food high in soluble fibre.

The GDG noted that any protective effect of fibre is from food rich in dietary fibre as opposed to supplemental fibre, because the former contain other nutrients and phytochemicals and the roles these play may be more important than the fibre alone.

HEALTH ECONOMIC EVIDENCE

The cost effectiveness of fibre was not estimated as fibre is not prescribed but purchased by people with IBS as part of their food or as an over the counter food supplement.

EVIDENCE STATEMENTS

1. There is a moderate amount of weak evidence to show that significantly more patients have improved global symptoms when taking soluble fibre compared with placebo, and that there is no significant difference for insoluble fibre compared with placebo.

2. There is weak evidence to show no significant effect on global symptoms of the means of delivery of fibre, whether given as a food or as a supplement.

3. There is good evidence to show that significantly more patients have improved global symptoms when taking soluble fibre compared with insoluble fibre; however there is no significant difference in pain or in improvement in bowel habits.

4. There is a fair evidence to show that flax seed containing flaxseed oil gave significantly less bloating than psyllium in people with IBS, but there was no significant difference in the number of people with improved bowel habits.

5. There is a moderate amount of fair evidence to show no significant difference in the state of well being and the number of patients with reduced pain, or improved bowel habit, when comparing a mixed diet containing 25 % or 75% cereal.

6. There is limited evidence to show little effect of fibre dose on pain, bloating and bowel scores in people with IBS.

7. There is inconsistent evidence of a protective effect of fibre on colorectal cancer, breast cancer and coronary heart disease, and a causal protective relationship has not been demonstrated.

EVIDENCE TO RECOMMENDATION

The GDG took into consideration the clinical evidence on the effectiveness of high fibre diets, together with their clinical experience of deleterious effects of a high fibre diet; they balanced these with a consideration of the protective nature of fibre against cancers and heart disease, as determined in the general population. The GDG was unanimous that the practice of recommending that people with IBS eat a diet high in fibre should cease, and recommended that the first stage in improving a person's diet was to review the fibre intake and adjust accordingly. The improvement in IBS symptoms due to soluble fibre was noted, and its possible protective effect against heart disease, so that the GDG recommended soluble fibre if an increase in fibre was required. Soluble fibre should be either in the form of supplements (as in the RCT evidence) or as foods high in soluble fibre, such as oats (from GDG consensus).

RECOMMENDATION

Healthcare professionals should review the fibre intake of people with IBS, adjusting (usually reducing) it while monitoring the effect on symptoms. People with IBS should be discouraged from eating insoluble fibre (for example, bran). If an increase in dietary fibre is advised, it should be soluble fibre such as ispaghula powder or foods high in soluble fibre (for example, oats).

7.4 Probiotics and prebiotics

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the probiotics review and are reported below.

Types of studies

The GDG decided that crossover studies should not be included in this review because it was unclear whether probiotics effected longer term changes or how long they were retained in the gut.

Types of intervention

Studies should include the following interventions:

- Single probiotics
- Combination probiotics
- Single prebiotics
- Synbiotics.

Probiotics may be given as a food or as an enteric coated capsule. Prebiotics should fulfil three criteria: (a) resistance to gastric acidity, hydrolysis by mammalian enzymes and gastrointestinal absorption; (b) fermentation by intestinal microflora; (c) selective stimulation of the growth and/or activity of intestinal bacteria associated with health and well-being. Acceptable prebiotics are mainly fructo-oligosaccharides, galacto-oligosaccharides and lactulose.

The following comparisons were included:

- Single probiotic versus placebo
- Combination probiotic versus placebo
- Single prebiotic versus placebo
- Synbiotics versus placebo
- Probiotic 1 versus probiotic 2
- Probiotic dose 1 versus dose 2
- Intervention duration 1 versus duration 2.

The probiotics review was concerned only with longer-term maintenance treatment.

In spite of the large placebo effect associated with IBS, comparisons with no treatment were included, and the minimum duration of treatment was four weeks.

Stratification and Subgroup analyses

Pre and probiotics were to be treated separately. We planned to carry out subgroup analyses as follows:

- o Type of probiotic (single, combination)
- o Nature of bacteria, including the strain (e.g. Lactobacillus salivarius, Bifidobacterium infantis, Streptococcus faecium)
- o Dose (above and below 1 x 10^6 bacteria per day; this was later revised to 10^6, 10^8, 10^{10} subgroups and the GDG later excluded studies with levels below 1 x 10^6)
- o Duration of intervention (5-8, 9-12, 13-16, 16+ weeks).

We also planned to investigate the effect of enteric coated capsules compared with the addition of probiotics as a food.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL and *The Cochrane Library* (1966 to current day with guidance from the GDG). The search strategies are listed in Appendix B.

The titles and abstracts from the search strategy were assessed. Thirty-seven were identified to be potentially relevant to the review and these papers were retrieved in full. Thirteen studies met the inclusion criteria for the review. The reference lists of these were inspected for further potential papers, but none were identified. The excluded studies are listed in Appendix E, along with reasons for exclusion.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

Thirteen studies met the inclusion criteria for the review (Bittner 2005; Gade 1989; Kajander 2005; Kim 2003; Kim 2005; Niedzielin 2001; Niv 2005; Nobaek 2000; Olesen 2000; O'Mahony 2004; Saggioro 2004; Tsuchiya 2004; Whorwell 2006). One study was conducted in the UK (Whorwell 2006) and one was carried out in Ireland (O'Mahony 2004). Two were conducted in Italy, three in the USA, two in Denmark and one each in Finland, Sweden, Poland and Israel.

The majority of studies had fewer than 100 patients. Two studies (Kajander 2005; Whorwell 2006) had more than 100 patients in total.

Study Design

Setting: The majority of studies took place in secondary care, but three were carried out in primary care (Bittner 2005; Gade 1989; Whorwell 2006) and one was assumed to be primary care (Nobaek 2000; recruited by newspaper advertisement). One study had patients from both primary and secondary care (O'Mahony 2004; patients from gastroenterology clinics and newspaper advertisement).

All the studies included in the review had a parallel design.

One study (O'Mahony 2004) had 3 arms comparing Lactobacillus salivarius UCC4431, Bifidobacterium infantis 35624 and placebo malted milk. Whorwell (2006) compared three different doses of encapsulated Bifidobacterium infantis 35624 with placebo in women with IBS. This gave a total of 20 comparisons in the review.

Population

The definition of IBS varied between studies: two used the Manning criteria (Niedzielin 2001; Olesen 2000); one used the Rome I criteria (Kajander 2005); one met criteria defined by the authors that were similar to the above (Gade 1989); and the rest used the Rome II criteria.

All studies but one included patients who had a range of IBS types; the other study specified diarrhoea predominant IBS symptoms (Kim 2003). Only one study (Niv 2005) stated that the participants had IBS as result of gastrointestinal infection.

The majority of studies (12) did not state the number of participants with bloating. Five studies had some patients with bloating measured as an outcome (Kajander 2005; Kim 2003; Olesen 2000; Tsuchiya 2004; Whorwell 2006). Two studies (Kim 2005; Niedzielin 2001) identified all patients as having bloating.

Most of the studies described symptom severity as mixed; one study described the symptoms as mild (Nobaek 2000). Two studies did not state symptom severity (Kajander 2005; O'Mahony 2004). Three studies suggested that the patients had refractory IBS: Saggioro (2004) reported that the patients had been treated with drugs without success; Niedzielin (2001) stated that all patients had been referred to secondary care because of problems with management; Tsuchiya (2004) reported that all patients had undergone a number of treatments without significant and lasting benefit.

The age range of participants across studies was 19 to 78 years, with the mean age (where given) ranging from 34 to 48 years. No study particularly identified elderly participants.
All studies had a ratio of women to men greater than one. Whorwell (2006) included only women participants.

Interventions

The studies varied in the type of probiotics used:

- Six used a single probiotic (Gade 1989; Niedzielin 2001; Niv 2005; Nobaek 2000; O'Mahony 2004; Whorwell 2006)
- Five used a combination of probiotics (Kajander 2005; Kim 2003; Kim 2005; Saggioro 2004; Tsuchiya 2004)
- One study used a prebiotic (Olesen 2000)
- One gave a pre/probiotic combination (Bittner 2005).

A range of different bacteria was used, and included various strains of Lactobacillus, Bifidobacterium and Streptococcus. Further details are available in the Appendix.

Three studies gave the probiotic in a capsule form (Kajander 2005; Whorwell 2006; Bittner 2005), two gave a tablet (Gade and Thorn 1989; Niv 2005) and the remainder used a food source or solution as the means of ingestion. The food sources used included milk or milk products, yoghurt, oatmeal soup and fruit drinks. The GDG considered the medium in which the probiotics were ingested to be an important difference and decided to consider, as subgroups, capsules versus other delivery routes. Whether the intervention was given with food was also considered important because of the increased levels of bile salts, as a result of digestive process, which are a serious obstacle to probiotic survival. However only three studies gave details as to when the probiotics were taken: Niedzielin (2001) directed that they should be taken before breakfast and two hours after the evening meal; in Gade (1989) the dose was given in the morning and evening with meals; and Olesen (2000) required the dose of prebiotics be taken with breakfast.

The GDG defined the minimum dose of probiotic as 1×10^6. The doses of probiotic varied considerably, and ranged from 8×10^6 (Gade 1989) to 4×10^{10} (Niedzielin 2001) for single probiotics, and 5×10^9 to 5×10^{11} for combination probiotics. It is noted that the activity of the probiotics vary according to strain.

The duration of the intervention ranged from four weeks (Gade 1989; Niedzielin 2001; Nobaek 2000; Saggioro 2004; Whorwell 2006) to six months (Kajander 2005; Niv 2005). Three studies had durations of eight weeks (Kim 2003; Kim 2005; O'Mahony 2004), and two studies had interventions lasting 12 weeks (Olesen 2000; Tsuchiya 2004). One study followed the patients for 12 months (Nobaek 2000).

Comparisons

The included studies covered the following comparisons:

- Nine comparisons of a single probiotic versus placebo (Gade 1989; Niedzielin 2001; Niv 2005; Nobaek 2000; O'Mahony 2004 x 2; Whorwell 2006 x3)
  - Lactobacillus salivarius UCC4331 (1×10^{10}) in malted milk drink (O'Mahony 2004)
  - Lactobacillus plantarum DSM 9843 (5×10^7 CFU) in oatmeal soup (Nobaek 2000)
  - Lactobacillus plantarum 299V (5×10^7) in oatmeal soup (Niedzielin 2001)
  - Lactobacillus reuteri ATCC 55730 (1×10^8) tablet (Niv 2005)
  - Two used Bifidobacterium infantis 35624 (O'Mahony 2004 (1×10^{10}) in malted milk drink; Whorwell (2006) 1×10^6, 1×10^8, 1×10^{10} CFU in capsule)
  - Streptococcus faecium (dose estimated as 8×10^6) tablet (Gade 1989);
- Five comparisons of a combination of probiotics versus placebo:

- - Two studies used VSL3 powder sachet (Bifidobacterium 3 strains, Lactobacillus 4 strains, Streptococcus 1 strain) (Kim 2003; Kim 2005)
  - SCM-III solution (Lactobacillus acidophilus1.25×10^6 CFU; Lactobacillus helveticus 1.3×10^9; bifidobacterium 4.95×10^9) (Tsuchiya 2004)
  - Lactobacillus rhamnosus GG, L. rhamnosus LC705, Bacillus breve Bb99, P.freudenreichii ssp.shermanii JS capsule (Kajander 2005)
  - Lactobacillus plantarum LPO1 & Bifidobacterium Breve BRO 5×10^9 CFU sachet dissolved in water (Saggioro 2004);
- One study compared two different probiotics (Lactobacillus salivarius UCC4331 versus Bifidobacterium infantis 35624) (O'Mahony 2004)
- One study compared three doses of probiotics (Whorwell 2006; 3 comparisons)
- One comparison of Prebiotic versus placebo (Oleson and Hoyer 2000)
- One comparison of a pre/probiotic capsule versus placebo, but this study contained no analysable data (Bittner 2005).

OUTCOMES

The studies measured a range of outcomes.

1. Global symptoms

a) Number of patients with an improvement in global symptoms

Seven studies recorded the patients' assessment of improvement (Gade 1989; Kajander 2005; Kim 2003; Niedzielin 2001; Olesen 2000; Tsuchiya 2004 – overall clinical effectiveness; Whorwell 2006 – adequate symptom relief) and two (Gade 1989; Tsuchiya 2004) also recorded a clinician assessment.

b) Global symptom score (mean)

The global symptom score was recorded by seven studies (Kajander 2005; Kim 2003; Niv 2005; Nobaek 2000; O'Mahony 2004; Saggioro 2004; Whorwell 2006), but Saggioro (2004) recorded the percentage change in global symptom score.

c) Global improvement in symptoms score (mean)

This outcome was recorded by two studies (Kim 2003; Olesen 2000).

d) Number of patients with deterioration in global symptoms

This outcome was recorded by three studies (Gade 1989; Olesen 2000; Tsuchiya 2004).

2. Individual symptoms

a) Pain

Pain was reported in several ways, either giving the number of patients with pain at the end of the study, the number of patients whose pain improved or worsened compared with the

baseline, and pain scores. The latter recorded a range of features, including severity, frequency and duration, or a combination of these. In addition, studies recorded the final scores, mean daily scores or the change from baseline. The studies reporting the following outcomes are listed below:

- Number of patients with pain: three studies (Gade 1989; Olesen 2000; Niedzielin 2001)
- Pain score: eight studies recorded a pain score (Kajander 2005; Kim 2003; Kim 2005; Nobaek 2000; O'Mahony 2004; Saggioro 2004; Tsuchiya 2004; Whorwell 2006), although Saggioro (2004) recorded the percentage change in pain score.
  In all cases the highest rating meant worst symptoms, although the scales used were not the same.
- Number of patients with less pain: one study (Niedzielin 2001).

b) Bloating

- Number of patients with more bloating (Olesen 2000)
- Number of patients with less bloating (Kim 2005; Olesen 2000)
- Bloating score (Kajander 2005; Kim 2003; Kim 2005; O'Mahony 2004; Tsuchiya 2004; Whorwell 2006).

c) Bowel habits

- Stool frequency (change and final) (Kajander 2005; Kim 2003; Kim 2005; Tsuchiya 2004). We decided that stool frequency was an unreliable measure of improvement if the type of IBS was not given. Only one of these studies specified the type of IBS (Kim 2003, which was in patients with diarrhoea predominant IBS), and the other studies were disregarded for this outcome.
- Stool score which was an aggregate score including stool frequency, consistency, ease of passage and completeness of evacuation (Kim 2003; Kim 2005; Nobaek 2000; O'Mahony 2004).

3. Quality of Life

- Two studies reported quality of life as an outcome (Niv 2005; Whorwell 2006).

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in Appendix D. The method of randomisation was reported in four studies, all of which gave an adequate method: computer generated numbers (Gade 1989; Kajander 2005; Olesen 2000) and one picking a card from a pack (O'Mahony 2004). The other studies did not state the method of randomisation (Kim 2003; Kim 2005; Niedzielin 2001; Niv 2005; Nobaek 2000; Tsuchiya 2004; Whorwell 2006).

Allocation concealment was reported in three studies (Kim 2005; Olesen 2000; O'Mahony 2004), one of which reported a partially adequate method (O'Mahony 2004), in which

randomisation and analysis were said to be 'supervised by a person independent from the study'. The other two were classified as having adequate concealment because the sequence was retained by a third party.

All the studies reported that the outcome assessors and the patients were blinded to the interventions. All described in detail the appearance and taste of the placebo and active intervention.

Four studies (Kajander 2005; Kim 2003; Kim 2005; Tsuchiya 2004) described an *a-priori* power calculation. Five studies used an intention to treat analysis (Kim 2003; Kim 2005; Olesen 2000; O'Mahony 2004; Whorwell 2006). All studies included in the review demonstrated some level of baseline comparability of the groups, but two provided limited data regarding baseline characteristics (Gade 1989; Nobaek 2000).

One study had no loss to follow-up (Niedzielin 2001). Three studies reported that more than 20% of patients in at least one arm (or overall) were not analysed or were lost to follow-up (Kim 2005; Niv 2005; Olesen 2000). For the Kim (2005) study we used four week data instead. In Niv (2005), 9/27 (33%) in the control group withdrew; and in Olesen (2000), 14/52 (27%) did not complete the 12 week comparative phase in FOS group.

The risk of bias was assessed for each included study and only Niv (2005) and Olesen (2000) were considered to be at higher risk of bias. These were considered, where possible, in sensitivity analyses.

RESULTS

A. Probiotics versus placebo

Ten studies compared probiotics (singly or in combination) with placebo (Gade 1989; Kajander 2005; Kim 2003; Kim 2005; Nobaek 2000; Niedzielin 2001; Niv 2005; O'Mahony 2004; Tsuchiya 2004; Whorwell 2006).

1. Global symptoms

a) Number of patients with an improvement in global symptoms

Six studies (eight comparisons), with 629 patients, recorded the patients' assessment of global improvement (Gade 1989; Kajander 2005, Kim 2003, Niedzielin 2001, Tsuchiya 2004, Whorwell 2007). Meta-analysis showed significant heterogeneity (I^2=65%; p=0.006).

Figure 1
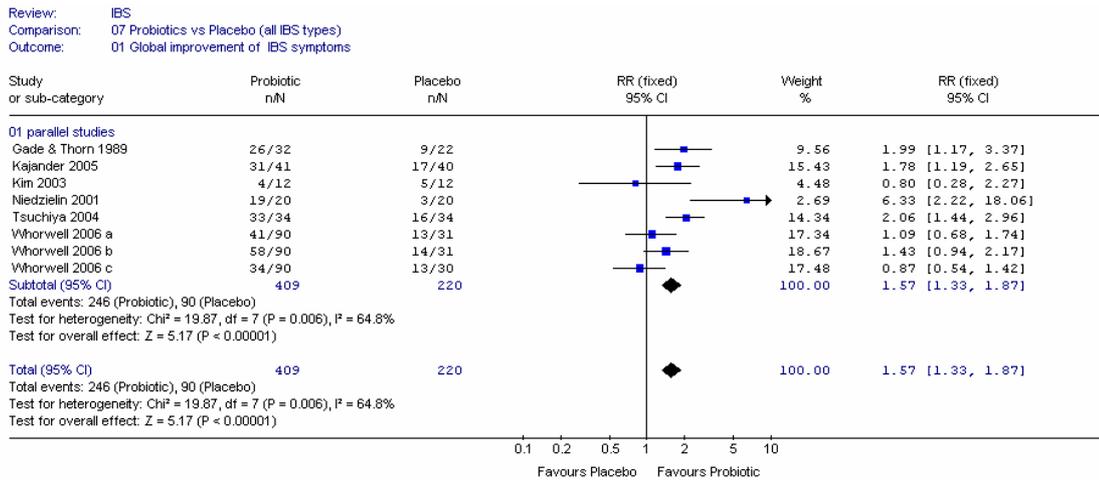
This heterogeneity was investigated in terms of the pre-specified subgroup analyses: by type of probiotic (single, combination), by duration and by dose and strain of bacterium (Figures 2 to 3).

Type of probiotic (Figure 2)

There was still significant heterogeneity in the single probiotic group, but it was not significant in the combination probiotic group (I^2=31%, p=0.23). Meta-analysis of three studies in 173 patients showed a statistically significant improvement in global symptoms. This corresponded to an NNT of 3 (95%CI 3, 5), for a control group rate of 42 to 47%. We noted there was significant heterogeneity for the risk difference (I^2=74%, p=0.02), which may have been an indication that the particular combination of probiotics was important.

Figure 2: Subgroup analysis by type of probiotic
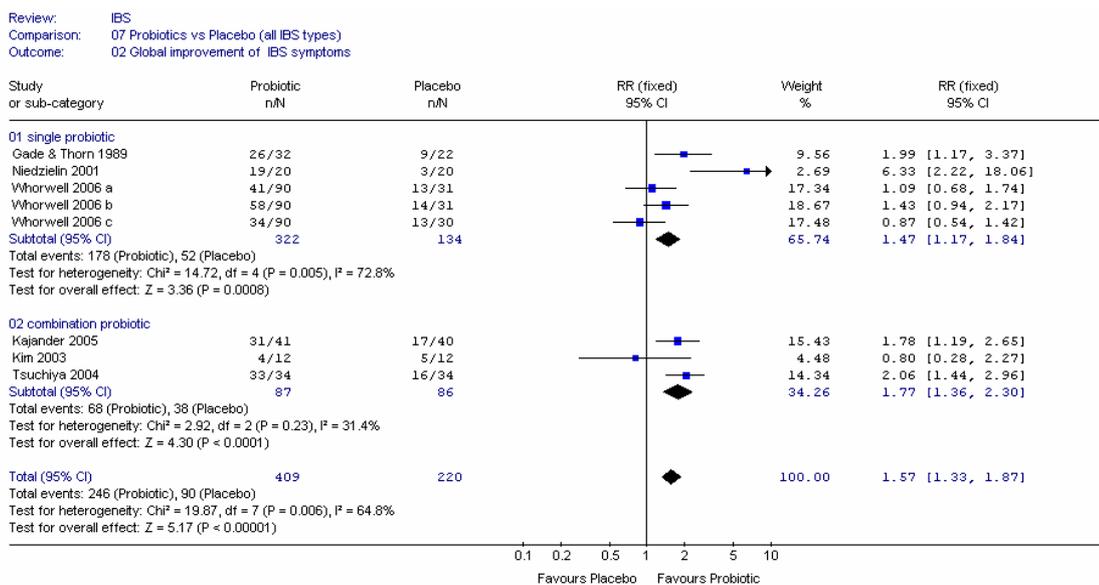
Duration

The Kajander (2005) study was six months duration, Tsuchiya (2004) was 12 weeks and the others were four or eight weeks. This did not account for the heterogeneity amongst studies.

Strain and dose of probiotic (Figure 3)

All the studies had different strains and/or doses, and the confidence intervals are wide in some cases. The heterogeneity may be indicative of different efficacies of the different probiotics; most of the probiotics tested in the trials gave a greater improvement in symptoms than placebo, but there were some exceptions. Whorwell (2006) showed a maximum in the improvement of global symptoms with increasing dose, with only the 10^8 dose of Bifidobacterium infantis being significant at four weeks. The authors attributed this effect to dissolution problems of the capsule for particular concentrations (see GDG discussion at the end of this review).

Figure 3: By type and dose of bacterium
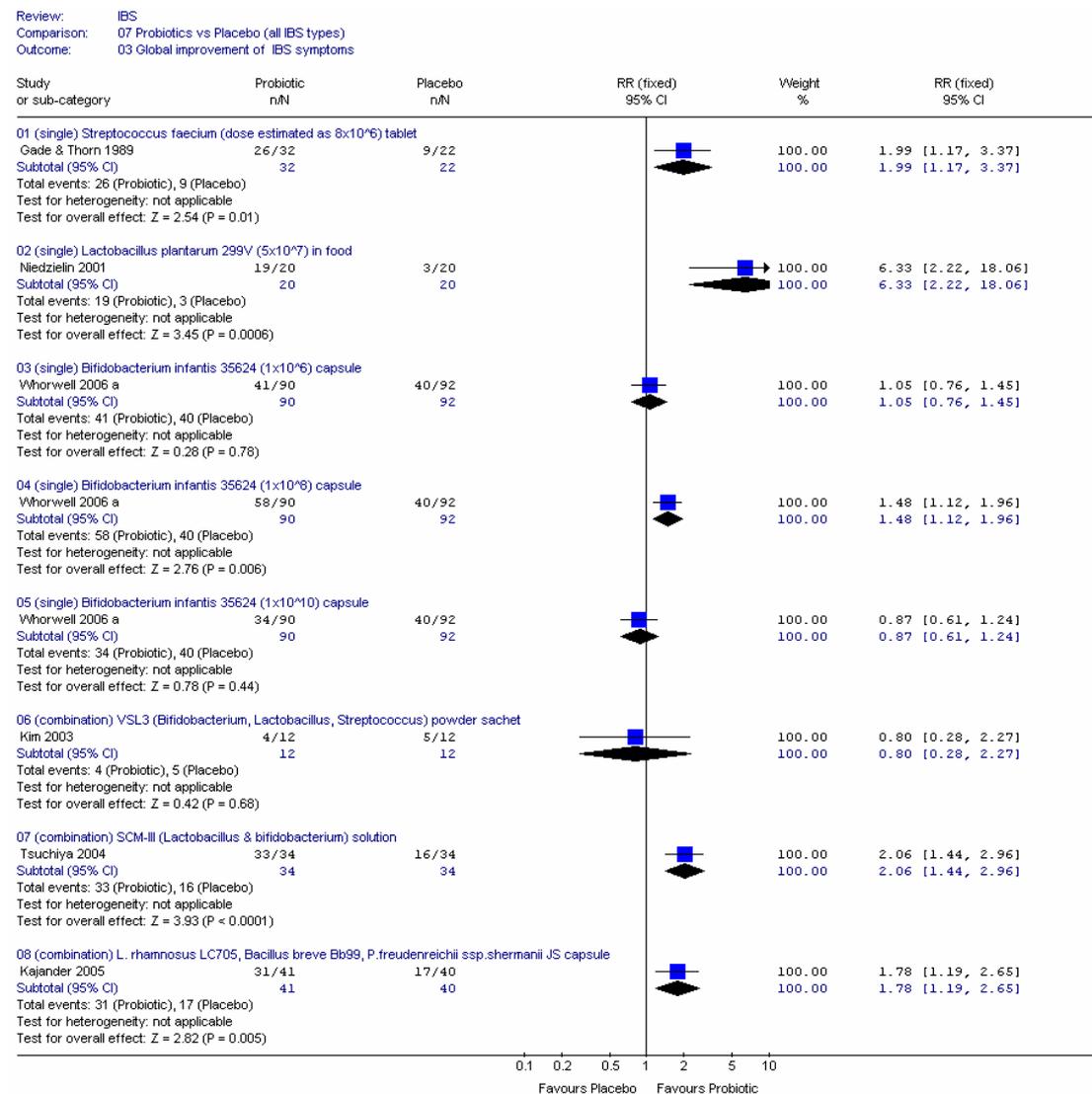
b) Number of patients with deterioration in global symptoms

This outcome was recorded by two studies (Gade 1989; Tsuchiya 2004). The confidence intervals were very wide, although Tsuchiya (2004) was statistically significantly in favour of probiotics.

Figure 4: (NB 0.01 to 100 scale)

Review:	Probiotics
Comparison:	12 Probiotic versus placebo
Outcome:	04 Global deterioration in IBS symptoms

Study or sub-category	Probiotics n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Single probiotic					
Gade & Thorn 1989	1/32	1/22		6.18	0.69 [0.05, 10.42]
Subtotal (95% CI)	32	22		6.18	0.69 [0.05, 10.42]
Total events: 1 (Probiotics), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.27 (P = 0.79)					
02 Combination Probiotics					
Tsuchiya 2004	1/34	18/34		93.82	0.06 [0.01, 0.39]
Subtotal (95% CI)	34	34		93.82	0.06 [0.01, 0.39]
Total events: 1 (Probiotics), 18 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.90 (P = 0.004)					
Total (95% CI)	66	56		100.00	0.09 [0.02, 0.41]
Total events: 2 (Probiotics), 19 (Placebo)					
Test for heterogeneity: Chi² = 2.33, df = 1 (P = 0.13), I² = 57.1%					
Test for overall effect: Z = 3.13 (P = 0.002)					

0.01　0.1　1　10　100
Favours Probiotics　Favours Placebo

c) Global symptom score (mean)

This outcome was recorded by seven studies (Kajander 2005; Kim 2003; Niv 2005; Nobaek 2000; O'Mahony 2004; Saggioro 2004; Whorwell 2006). One study (Saggioro 2004) reported percentage change scores, with p values, so these results are given separately. The other studies reported the global symptom score, but on different scales, so the standardised mean difference was used to analyse the data (Figure 5). The Niv (2005) values were taken from a graph and it was assumed that the standard error was given. This study also had some attrition bias, so a sensitivity analysis was repeated excluding this study (Figure 6).

Figure 5

Review:	IBS
Comparison:	07 Probiotics vs Placebo (all IBS types)
Outcome:	05 Global Symptom Score (mean) Low score=good

Study or sub-category	N	Probiotic Mean (SD)	N	Placebo Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Single probiotic							
Niv 2005	21	268.00(134.27)	18	224.31(124.31)		7.21	0.33 [-0.30, 0.96]
Nobaek 2000	25	3.90(1.28)	27	4.31(1.87)		9.72	-0.25 [-0.80, 0.30]
O'Mahony 2004 a	25	5.87(3.33)	12	5.47(4.00)		6.11	0.11 [-0.58, 0.80]
O'Mahony 2004 b	25	4.40(3.33)	12	5.47(4.00)		6.06	-0.29 [-0.99, 0.40]
Whorwell 2006 a	90	2.15(0.85)	31	2.09(0.96)		17.41	0.07 [-0.34, 0.48]
Whorwell 2006 b	90	1.76(0.85)	31	2.09(0.96)		17.18	-0.37 [-0.78, 0.04]
Whorwell 2006 c	90	2.13(0.85)	30	2.09(0.96)		16.99	0.05 [-0.37, 0.46]
Subtotal (95% CI)	366		161			80.69	-0.07 [-0.26, 0.12]
Test for heterogeneity: Chi² = 5.44, df = 6 (P = 0.49), I² = 0%							
Test for overall effect: Z = 0.72 (P = 0.47)							
02 Combined probiotic (final score)							
Kajander 2005	40	20.40(13.88)	40	26.88(13.88)		14.69	-0.46 [-0.91, -0.02]
Kim 2003	12	103.00(48.50)	13	121.00(46.90)		4.62	-0.37 [-1.16, 0.43]
Subtotal (95% CI)	52		53			19.31	-0.44 [-0.83, -0.05]
Test for heterogeneity: Chi² = 0.04, df = 1 (P = 0.83), I² = 0%							
Test for overall effect: Z = 2.22 (P = 0.03)							
Total (95% CI)	418		214			100.00	-0.14 [-0.31, 0.03]
Test for heterogeneity: Chi² = 8.30, df = 8 (P = 0.41), I² = 3.6%							
Test for overall effect: Z = 1.62 (P = 0.10)							

-4　-2　0　2　4
Favours Probiotic　Favours Placebo

Meta-analysis of eight comparisons in 624 patients showed no significant difference between probiotics and placebo overall, with little heterogeneity (I^2=4%, p=0.41). However, there was a statistically significant difference for the combined probiotic subgroup in 105 patients. Meta-analysis of seven single probiotics showed no significant difference between probiotic and placebo, with no heterogeneity. Sensitivity analysis without Niv (2005) made a small difference.

The Saggioro (2004) study in 40 patients reported a statistically significant difference in the percentage change in IBS symptom severity (-44% versus -8.5% after 28 days; p<0.001 for the combined probiotic versus placebo).

Figure 6: Sensitivity analysis without Niv (2005)
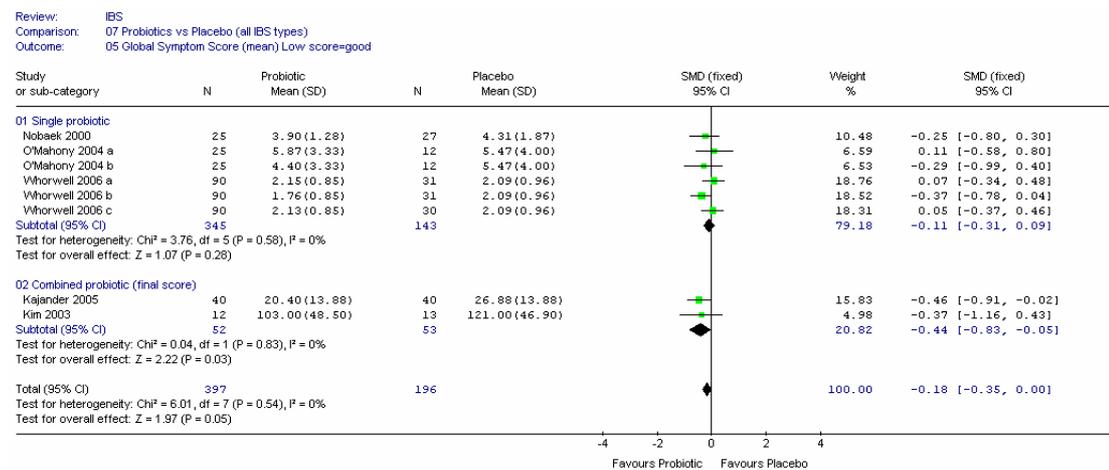
Subgroup analyses were carried out by strain and dose of bacterium because the GDG was uncertain whether studies using different bacteria should be combined. Figure 7 shows the studies by bacterium and dose. Most comparisons showed no significant difference compared with placebo, including the meta-analysis of two studies in 232 patients, receiving Bifidobacterium infantis at a dose of 1×10^{10} CFU; there was no significant heterogeneity for these two studies (I^2=6%, p=0.30). There were only two statistically significant comparisons:

- Encapsulated Bifidobacterium infantis at a dose of 1×10^8 CFU versus placebo, in 182 patients. This had a mean difference of -0.33 (95%CI -0.59, -0.07) on a scale of 0 to 15 (i.e. a fairly small effect)
- Encapsulated combined probiotic versus placebo in 81 patients. This had a mean difference of -6.48 (95%CI -12.56, -0.40) on a scale of 0 to 112.

This may, however, be a size effect; most of the non-significant studies had around 50 patients or fewer.

Figure 7a: Global symptom score
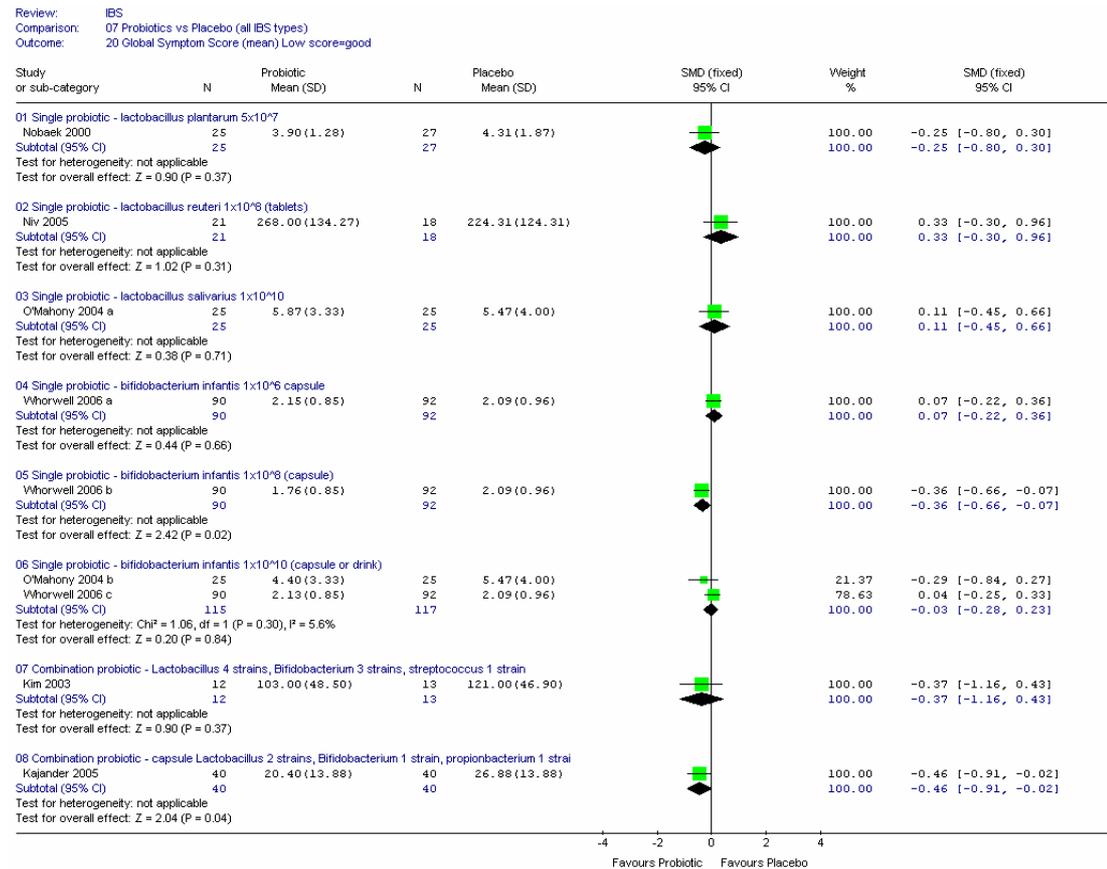
Nobaek (2000) also reported 12 month follow-up, but it was unclear if the patients continued to modify their diet after the trial had ended. There was a borderline significant effect of probiotics at 12 months, but not at 5 to 6 weeks. The scale was 0 to 10.

Figure 7b: Global symptom score – 5/6 weeks and 12 month follow-up
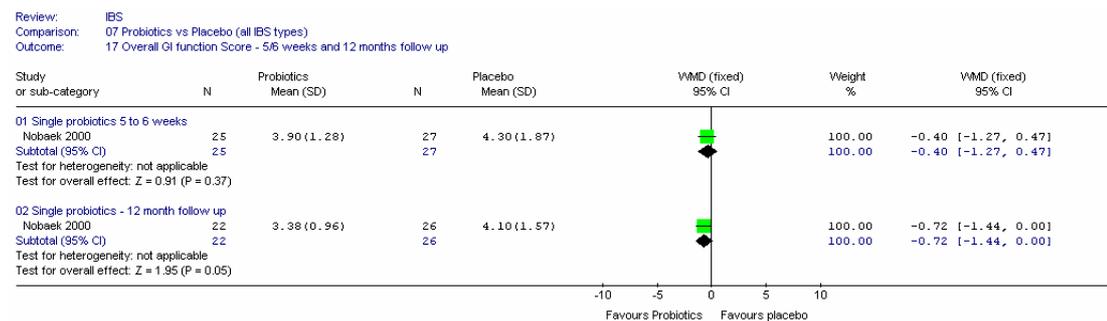
e) Global improvement in symptoms score (mean)

This outcome was recorded by one study (Kim 2003), which showed too much uncertainty to draw conclusions. It is unclear what scale is used.

Figure 8
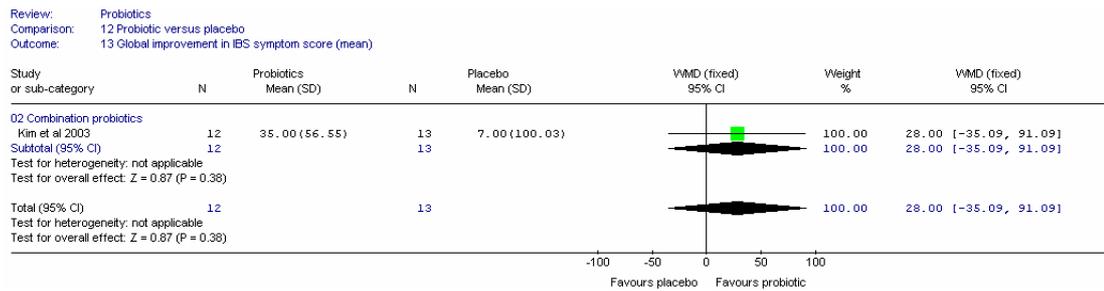
2. Individual symptoms

a) Pain

i. Number of patients with pain

Two studies measured this outcome (Gade 1989; Niedzielin 2001) and both separately were statistically significantly in favour of probiotic. The relative risk ranged from 0.03 to 0.2 (i.e. 5 to 33 times less risk of pain with the probiotic). However, combining the studies gave heterogeneity (I^2=80%; p=0.03).

Figure 9
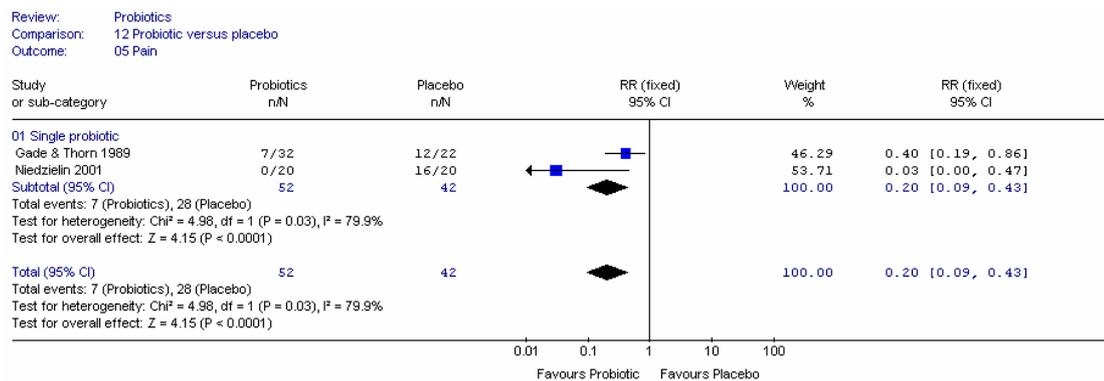
ii. Pain score

Eight studies (Kajander 2005; Kim 2003; Kim 2005; Nobaek 2000; O'Mahony 2004; Saggioro 2004; Tsuchiya 2004; Whorwell 2006) reported a pain score. Saggioro (2004) reported percentage change scores, with p values, so these results are given separately.

This outcome showed no significant difference between interventions for the single probiotic group (although it is difficult to estimate the width of the confidence interval because the standardised mean difference was used) and a highly heterogeneous result for the combined probiotics group (I^2=93%, p=0.00001), attributable to the Tsuchiya (2004) study, from which data were extracted from a graph, which may not have been to scale for the standard deviations. In the absence of this study the meta-analysis of three studies gave a statistically significant reduction in pain for the combined probiotic group, and no heterogeneity (I^2=0%; p=0.94).

Saggioro (2004) reported a statistically significant difference in the percentage change in pain score (-38% versus -18% after 28 days; p<0.05 for the combined probiotic versus placebo).

Figure 10a

					SMD (fixed)	Weight	SMD (fixed)

Review: IBS
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 10 Pain Score

Study or sub-category	Probiotics N	Mean (SD)	Placebo N	Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Single probiotics							
Nobaek 2000	25	2.81(0.66)	27	3.20(1.37)		8.78	-0.35 [-0.90, 0.20]
O'Mahony 2004 a	25	1.50(1.17)	12	1.50(1.17)		5.58	0.00 [-0.69, 0.69]
O'Mahony 2004 b	25	1.03(1.17)	12	1.50(1.17)		5.47	-0.39 [-1.09, 0.30]
Whorwell 2006 a	90	1.89(0.95)	31	1.73(0.96)		15.81	0.17 [-0.24, 0.58]
Whorwell 2006 b	90	1.43(0.95)	31	1.73(0.96)		15.71	-0.31 [-0.72, 0.10]
Whorwell 2006 c	90	1.84(0.95)	30	1.73(0.96)		15.45	0.11 [-0.30, 0.53]
Subtotal (95% CI)	345		143			66.80	-0.09 [-0.28, 0.11]

Test for heterogeneity: Chi² = 5.27, df = 5 (P = 0.38), I² = 5.2%
Test for overall effect: Z = 0.85 (P = 0.40)

02 Combined probiotics							
Kajander 2005	41	4.24(4.41)	51	5.80(4.36)		15.38	-0.35 [-0.77, 0.06]
Kim 2003	12	23.00(13.86)	13	26.00(14.42)		4.26	-0.20 [-0.99, 0.58]
Kim 2005	24	23.00(10.70)	24	26.90(10.70)		8.11	-0.36 [-0.93, 0.21]
Tsuchiya 2004	34	1.51(0.47)	34	2.91(0.47)		5.44	-2.94 [-3.64, -2.25]
Subtotal (95% CI)	111		122			33.20	-0.76 [-1.04, -0.48]

Test for heterogeneity: Chi² = 45.25, df = 3 (P < 0.00001), I² = 93.4%
Test for overall effect: Z = 5.28 (P < 0.00001)

| Total (95% CI) | 456 | | 265 | | | 100.00 | -0.31 [-0.47, -0.15] |

Test for heterogeneity: Chi² = 65.15, df = 9 (P < 0.00001), I² = 86.2%
Test for overall effect: Z = 3.73 (P = 0.0002)

-4 -2 0 2 4
Favours Probiotics Favours placebo

Figure 10b: Without Tsuchiya 2004

Review: IBS
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 10 Pain Score

Study or sub-category	Probiotics N	Mean (SD)	Placebo N	Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Single probiotics							
Nobaek 2000	25	2.81(0.66)	27	3.20(1.37)		9.28	-0.35 [-0.90, 0.20]
O'Mahony 2004 a	25	1.50(1.17)	12	1.50(1.17)		5.90	0.00 [-0.69, 0.69]
O'Mahony 2004 b	25	1.03(1.17)	12	1.50(1.17)		5.79	-0.39 [-1.09, 0.30]
Whorwell 2006 a	90	1.89(0.95)	31	1.73(0.96)		16.72	0.17 [-0.24, 0.58]
Whorwell 2006 b	90	1.43(0.95)	31	1.73(0.96)		16.61	-0.31 [-0.72, 0.10]
Whorwell 2006 c	90	1.84(0.95)	30	1.73(0.96)		16.34	0.11 [-0.30, 0.53]
Subtotal (95% CI)	345		143			70.64	-0.09 [-0.28, 0.11]

Test for heterogeneity: Chi² = 5.27, df = 5 (P = 0.38), I² = 5.2%
Test for overall effect: Z = 0.85 (P = 0.40)

02 Combined probiotics							
Kajander 2005	41	4.24(4.41)	51	5.80(4.36)		16.27	-0.35 [-0.77, 0.06]
Kim 2003	12	23.00(13.86)	13	26.00(14.42)		4.51	-0.20 [-0.99, 0.58]
Kim 2005	24	23.00(10.70)	24	26.90(10.70)		8.58	-0.36 [-0.93, 0.21]
Subtotal (95% CI)	77		88			29.36	-0.33 [-0.64, -0.02]

Test for heterogeneity: Chi² = 0.12, df = 2 (P = 0.94), I² = 0%
Test for overall effect: Z = 2.11 (P = 0.03)

| Total (95% CI) | 422 | | 231 | | | 100.00 | -0.16 [-0.33, 0.01] |

Test for heterogeneity: Chi² = 7.11, df = 8 (P = 0.52), I² = 0%
Test for overall effect: Z = 1.86 (P = 0.06)

-4 -2 0 2 4
Favours Probiotics Favours placebo

Nobaek (2000) also reported 12 month follow-up data, shown in Figure 10c. The scale is a visual analogue scale of 0 to 10. There was no significant difference at 5 to 6 weeks, but a statistically significant difference after 12 months. It was unclear if the patients in the intervention group changed their dietary habits following the trial or if there was a long-term effect.

Figure 10c

Review: IBS
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 07 Pain Score - 5/6 weeks and 12 months follow up

Study or sub-category	Probiotics N	Mean (SD)	Placebo N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Single probiotics 5 to 6 weeks							
Nobaek 2000	25	2.81(0.66)	27	3.20(1.37)		100.00	-0.39 [-0.97, 0.19]
Subtotal (95% CI)	25		27			100.00	-0.39 [-0.97, 0.19]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.32 (P = 0.19)							
02 Single probiotics - 12 month follow up							
Nobaek 2000	22	2.73(0.80)	26	3.76(1.45)		100.00	-1.03 [-1.68, -0.38]
Subtotal (95% CI)	22		26			100.00	-1.03 [-1.68, -0.38]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.11 (P = 0.002)							

-10 -5 0 5 10
Favours Probiotics   Favours placebo

iii. Number of patients with less pain

Two studies compared Lactobacillus planetarum, given in food, with placebo (Niedzielin 2001; Nobaek 2000) and reported the number of patients with reduced pain. There was a statistically significant reduction in pain; RR 1.67 (95%CI 1.09, 2.56), with no heterogeneity (I^2=0, p=0.64). This corresponded to a number needed to treat of 5 (95%CI 3, 20) for a control group rate of 19 to 55%.

Figure 11

Review: IBS
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 07 Number of patients with less pain

Study or sub-category	Probiotics n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 single probiotic (Lactobacillus plantarum) given as food					
Niedzielin 2001	17/20	11/20		69.59	1.55 [1.00, 2.39]
Nobaek 2000	9/25	5/27		30.41	1.94 [0.75, 5.02]
Subtotal (95% CI)	45	47		100.00	1.67 [1.09, 2.56]
Total events: 26 (Probiotics), 16 (Placebo)					
Test for heterogeneity: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%					
Test for overall effect: Z = 2.34 (P = 0.02)					
Total (95% CI)	45	47		100.00	1.67 [1.09, 2.56]
Total events: 26 (Probiotics), 16 (Placebo)					
Test for heterogeneity: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%					
Test for overall effect: Z = 2.34 (P = 0.02)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours probiotics

b) Bloating

i. Number of patients with less bloating (Kim 2005)

In a single study in 48 patients, there was no significant difference between probiotics and placebo in the number of patients with less bloating, although the confidence interval was fairly wide.

Figure 12

Review: Probiotics
Comparison: 12 Probiotic versus placebo
Outcome: 15 Number of patients with less bloating

Study or sub-category	Probiotic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Combination probiotic					
Kim et al 2005	11/24	8/24		100.00	1.38 [0.67, 2.81]
Subtotal (95% CI)	24	24		100.00	1.38 [0.67, 2.81]
Total events: 11 (Probiotic), 8 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.87 (P = 0.38)					
Total (95% CI)	24	24		100.00	1.38 [0.67, 2.81]
Total events: 11 (Probiotic), 8 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.87 (P = 0.38)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours probiotic

ii. Bloating score (Kajander 2005; Kim 2003; Kim 2005; O'Mahony 2004; Tsuchiya 2004; Whorwell 2006)

This outcome showed no significant difference between interventions for the single probiotic group and a highly heterogeneous result for the combined probiotics group, attributable to the Tsuchiya (2004) study, from which data were extracted from a graph, which may not be to scale for the standard deviations. The authors of this study reported that there was no significant difference between groups, which belies the data on the graph, suggesting that the standard deviations on the graph were inaccurate. In the absence of this study, meta-analysis of the three studies in 73 patients gave a statistically significant reduction in bloating score for the combined probiotic group, MD -0.42 (95%CI -0.73, -0.10), with no heterogeneity (I^2=0%; p=0.87).

Figure 13a: Bloating score (final scores)
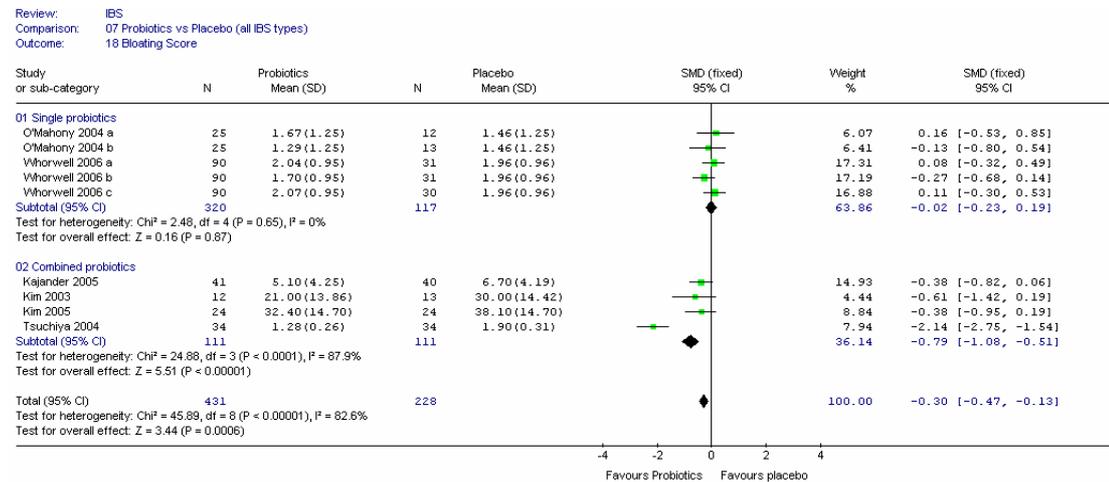
Figure 13b: Sensitivity analysis without Tsuchiya (2004)
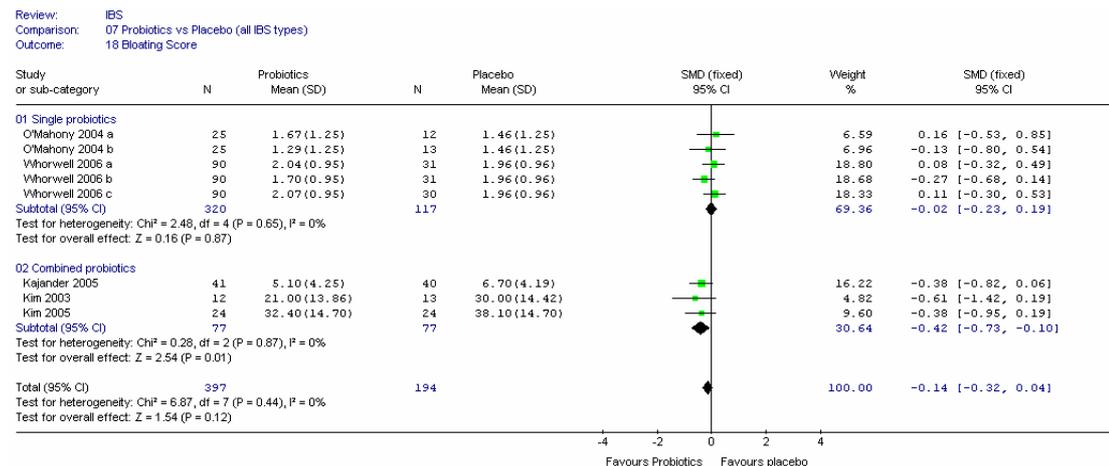
c) Bowel habits

i. Stool frequency

Only one study specified the type of IBS (Kim 2003), which was in patients with diarrhoea predominant IBS. For this study, the frequency was seen as a negative outcome and there was no significant difference between probiotic and placebo.

Figure 14
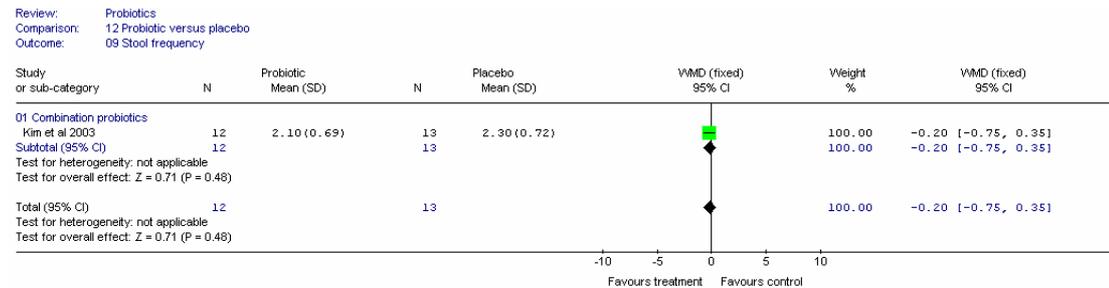
ii. Stool score

This was aggregated to include stool frequency, consistency, ease of passage and completeness of evacuation (Kim 2003; Kim 2005; O'Mahony 2004; Nobaek 2000; Whorwell 2006 – bowel habit satisfaction). Tsuchiya (2004) also reported assessment of bowel habits, but these values were not included in the meta-analysis in view of the uncertainties in the standard deviation described above.

In the meta-analysis of eight comparisons (562 patients) there was no significant difference between probiotics and placebo for this outcome, either overall, or for single or combined probiotics, and there was no significant heterogeneity.

Figure 15
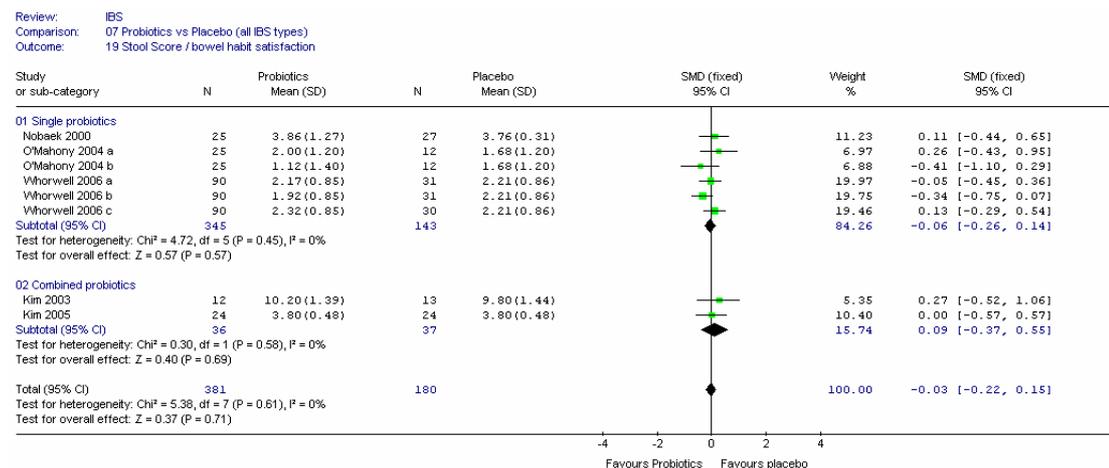
3. Quality of Life

Only one study reported quality of life as an outcome (Niv 2005). This showed no significant difference in the quality of life score. The scale used was unclear, however: the study reported

that there were 26 questions, each rated from mild (1) to severe (7), and the sum of all of them yielded the total QoL score, but the baseline scores for the total were similar to the individual components and were about 4 to 5 points.

Figure 16

Review: IBS MW 080507
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 18 Quality of life

Study or sub-category	N	Probiotic Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Niv 2005	27	4.04(1.30)	27	3.79(1.30)		100.00	0.25 [-0.44, 0.94]
Total (95% CI)	27		27			100.00	0.25 [-0.44, 0.94]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.71 (P = 0.48)

Favours treatment   Favours control

Examination of the two studies using Bifidobacterium infantis 35624

Two studies compared Bifidobacterium infantis 35624, at a dose of 1×10^{10} CFU per day, versus placebo. One study (Whorwell 2006) gave the probiotic in a capsule and the other (O'Mahony 2004) in a malted drink.

The outcomes are summarised in Figure 17. There was some heterogeneity between studies for the outcomes of pain and stool score, with the encapsulated probiotic having less effect. This is discussed further in the next section.

Figure 17

Review: IBS
Comparison: 07 Probiotics vs Placebo (all IBS types)
Outcome: 21 Symptom Scores, global and individual, bifidobacterium infantis 10^10

Study or sub-category	N	Probiotic Mean (SD)	N	Placebo Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Global symptom score							
O'Mahony 2004 b	25	4.40(3.33)	12	5.47(4.00)		6.56	-0.29 [-0.99, 0.40]
Whorwell 2006 c	90	2.13(0.85)	30	2.09(0.96)		18.38	0.05 [-0.37, 0.46]
Subtotal (95% CI)	115		42			24.94	-0.04 [-0.40, 0.31]

Test for heterogeneity: Chi² = 0.68, df = 1 (P = 0.41), I² = 0%
Test for overall effect: Z = 0.24 (P = 0.81)

02 Pain score							
O'Mahony 2004 b	25	1.03(1.17)	12	1.50(1.17)		6.50	-0.39 [-1.09, 0.30]
Whorwell 2006 c	90	1.84(0.95)	30	1.73(0.96)		18.36	0.11 [-0.30, 0.53]
Subtotal (95% CI)	115		42			24.87	-0.02 [-0.37, 0.34]

Test for heterogeneity: Chi² = 1.52, df = 1 (P = 0.22), I² = 34.0%
Test for overall effect: Z = 0.10 (P = 0.92)

03 Bloating score							
O'Mahony 2004 b	25	1.29(1.25)	13	1.46(1.25)		6.97	-0.13 [-0.80, 0.54]
Whorwell 2006 c	90	2.07(0.95)	30	1.96(0.96)		18.36	0.11 [-0.30, 0.53]
Subtotal (95% CI)	115		43			25.34	0.05 [-0.31, 0.40]

Test for heterogeneity: Chi² = 0.38, df = 1 (P = 0.54), I² = 0%
Test for overall effect: Z = 0.26 (P = 0.80)

04 Stool score							
O'Mahony 2004 b	25	1.12(1.40)	12	1.68(1.20)		6.49	-0.41 [-1.10, 0.29]
Whorwell 2006 c	90	2.32(0.85)	30	2.21(0.86)		18.36	0.13 [-0.29, 0.54]
Subtotal (95% CI)	115		42			24.85	-0.01 [-0.37, 0.34]

Test for heterogeneity: Chi² = 1.69, df = 1 (P = 0.19), I² = 40.9%
Test for overall effect: Z = 0.07 (P = 0.95)

Total (95% CI)	460		169			100.00	-0.01 [-0.18, 0.17]

Test for heterogeneity: Chi² = 4.41, df = 7 (P = 0.73), I² = 0%
Test for overall effect: Z = 0.07 (P = 0.94)

Favours Probiotic   Favours Placebo

B. Probiotic dose 1 versus probiotic dose 2

One study (Whorwell 2006) compared three doses of Bifidobacterium infantis 35624, 1×10^6, 1×10^8, 1×10^{10}, with approximately 90 patients in each arm. The outcomes compared are reported in Figure 18.

Head-to-head comparison of the doses 1×10^8 and 1×10^6 showed there was a significant difference in global symptoms, pain and bloating scores and a borderline difference in stool score, favouring the 1×10^8 dose. However, there was no significant difference between the 1×10^{10} and 1×10^6 doses, an unexpected dose effect.

Figure 18
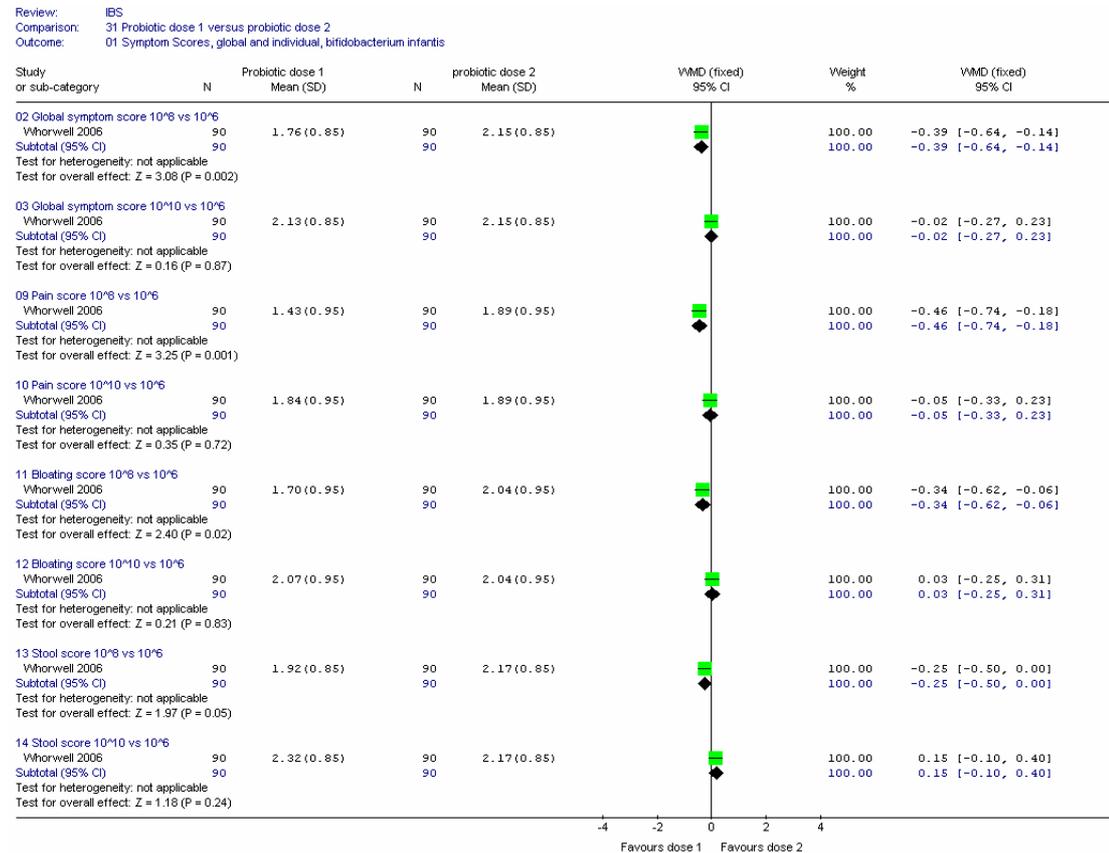
Whorwell (2006) explained this using in-vitro dissolution experiments, showing that the highest concentration of probiotic coagulated on exposure to moisture, making dissolution very difficult, such that the probiotic was not bioavailable to the patient.

This effect also explains the differences between Whorwell (2006) and O'Mahony (2004); in the latter, the probiotic was bioavailable because it was present in a drink.

C. Probiotic 1 versus probiotic 2

One study compared two strains of bacteria, Lactobacillus salivarius UCC4331 versus Bifidobacterium infantis 35624 (O'Mahony 2004) directly in a randomised trial of 50 patients. The results are presented below for the different outcome measures.

1. Global symptom score

There was no significant difference in global symptoms at eight weeks, but the Bifidobacterium is favoured. A 10cm visual analogue scale was used for individual symptoms and combined to give a global score (maximum 30).

Figure 19

Review: IBS
Comparison: 30 Probiotic 1 versus probiotic 2
Outcome: 01 Global Symptom Score (mean) Low score=good

Study or sub-category	N	Probiotic 1 Mean (SD)	N	Probiotic 2 Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Lactobacillus salivarius UCC4331 versus Bifidobacterium infantis 35624							
O'Mahony 2004	25	5.87(3.33)	25	4.40(3.33)		100.00	1.47 [-0.38, 3.32]
Subtotal (95% CI)	25		25			100.00	1.47 [-0.38, 3.32]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.56 (P = 0.12)							
Total (95% CI)	25		25			100.00	1.47 [-0.38, 3.32]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.56 (P = 0.12)							

-10 -5 0 5 10
Favours Probiotic 1   Favours Probiotic 2

2. Individual symptoms

There was no significant difference between the two types of bacteria for pain, bloating or stool score. Likert scales were used for each component with a maximum of 7.

Figure 20

Review: IBS
Comparison: 30 Probiotic 1 versus probiotic 2
Outcome: 02 Individual Symptom Scores (mean) Low score=good

Study or sub-category	N	Probiotic 1 Mean (SD)	N	Probiotic 2 Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Pain score							
O'Mahony 2004	25	1.50(1.17)	25	1.03(1.17)		100.00	0.47 [-0.18, 1.12]
Subtotal (95% CI)	25		25			100.00	0.47 [-0.18, 1.12]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.42 (P = 0.16)							
02 Bloating score							
O'Mahony 2004	25	1.67(1.25)	12	1.29(1.25)		100.00	0.38 [-0.48, 1.24]
Subtotal (95% CI)	25		12			100.00	0.38 [-0.48, 1.24]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.87 (P = 0.39)							
03 Stool score							
O'Mahony 2004	25	2.00(1.20)	12	1.68(1.20)		100.00	0.32 [-0.51, 1.15]
Subtotal (95% CI)	25		12			100.00	0.32 [-0.51, 1.15]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.76 (P = 0.45)							

-10 -5 0 5 10
Favours Probiotic 1   Favours Probiotic 2

D. Prebiotics versus placebo

One study (Olesen 2000) compared a prebiotic (Fructooliogsaccharide given as a 10g sachet for 2 weeks then 20g for 10 weeks) with placebo in 98 patients. The results are given below and generally showed no significant differences between prebiotics and placebo, in either global symptoms or bloating (although the confidence interval was fairly wide in the latter). The confidence interval was too wide to determine if there was a difference for the pain outcome. We noted that there was some attrition bias for this study.

Figure 21

Study or sub-category	Prebiotic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Global improvement of IBS symptoms					
Oleson & Hoyer 2000	29/50	30/46		100.00	0.89 [0.65, 1.22]
Subtotal (95% CI)	50	46		100.00	0.89 [0.65, 1.22]
Total events: 29 (Prebiotic), 30 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.73 (P = 0.47)					
02 Number of patients with less gas					
Oleson & Hoyer 2000	15/50	24/46		100.00	0.58 [0.35, 0.95]
Subtotal (95% CI)	50	46		100.00	0.58 [0.35, 0.95]
Total events: 15 (Prebiotic), 24 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.14 (P = 0.03)					
03 Number of patients with less bloating					
Oleson & Hoyer 2000	21/50	24/46		100.00	0.81 [0.53, 1.23]
Subtotal (95% CI)	50	46		100.00	0.81 [0.53, 1.23]
Total events: 21 (Prebiotic), 24 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.99 (P = 0.32)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours prebiotic

Figure 22

Study or sub-category	Prebiotic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Global deterioration of IBS symptoms					
Oleson & Hoyer 2000	4/50	6/46		100.00	0.61 [0.18, 2.04]
Subtotal (95% CI)	50	46		100.00	0.61 [0.18, 2.04]
Total events: 4 (Prebiotic), 6 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.80 (P = 0.42)					
02 Number of patients with pain					
Oleson & Hoyer 2000	7/50	1/46		100.00	6.44 [0.82, 50.36]
Subtotal (95% CI)	50	46		100.00	6.44 [0.82, 50.36]
Total events: 7 (Prebiotic), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.77 (P = 0.08)					
03 Number of patients with more bloating					
Oleson & Hoyer 2000	11/50	11/46		100.00	0.92 [0.44, 1.92]
Subtotal (95% CI)	50	46		100.00	0.92 [0.44, 1.92]
Total events: 11 (Prebiotic), 11 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.22 (P = 0.82)					
04 Number of patients with more gas					
Oleson & Hoyer 2000	14/50	6/46		100.00	2.15 [0.90, 5.12]
Subtotal (95% CI)	50	46		100.00	2.15 [0.90, 5.12]
Total events: 14 (Prebiotic), 6 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.72 (P = 0.08)					
05 Number of withdrawals					
Oleson & Hoyer 2000	14/52	7/46		100.00	1.77 [0.78, 4.00]
Subtotal (95% CI)	52	46		100.00	1.77 [0.78, 4.00]
Total events: 14 (Prebiotic), 7 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.37 (P = 0.17)					
06 Number of patients with adverse effects					
Oleson & Hoyer 2000	21/50	22/46		100.00	0.88 [0.56, 1.37]
Subtotal (95% CI)	50	46		100.00	0.88 [0.56, 1.37]
Total events: 21 (Prebiotic), 22 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.57 (P = 0.57)					

0.1 0.2 0.5 1 2 5 10
Favours prebiotic   Favours placebo

HEALTH ECONOMIC EVIDENCE

The cost effectiveness of pre and probiotics was not estimated as they are not prescribed, but currently purchased by patients as a food supplement.

GDG DISCUSSION

The GDG discussed the use of pre and probiotics at some length. They were unanimous in their view that different types and doses of probiotic should not be combined together in an analysis because they all have different effects. The main issues raised for discussion were dose, method of ingestion and quality of products available to patients. Probiotics are not generally prescribed by GPs. Patients purchase them and there was concern that sources are not always reliable or safe. There was agreement that there is insufficient information for patients about the quality of products and insufficient information on packaging regarding dose and quality of individual products.

The studies that investigated Bifidobacterium infantis 35624 (Whorwell 2006; O'Mahony 2004) were discussed with regard to the observed maximum in the dose response in Whorwell (2006) and the inconsistencies between the two studies. This was explained by the method of ingestion. In Whorwell (2006), a capsule was used as the means of ingesting the different doses of probiotic. For the $1\text{x}10^{10}$ CFU concentration of probiotic, contact with water led to the probiotic coagulating so that it was no longer bioavailable to the patient. The same dose of probiotic was found to be effective in O'Mahony (2004) because the probiotic was ingested in the form of a milk based drink so the concentration of probiotic was evenly dispersed through the fluid and therefore bioavailable to the patient.

EVIDENCE STATEMENTS

1. There is fair evidence to show that some probiotics (single or combination) give a significantly greater improvement in global symptoms of IBS than placebo. However, this is bacterium dependent, in terms of both dose and strain.

2. There is good evidence to show a significant difference in global symptom score for combined probiotics compared with placebo, favouring probiotics, but no significant difference for single probiotics as a group in people with IBS.

3. There is fair evidence to show a significant reduction in the number of people with pain for those taking single probiotics compared with placebo; there is weak evidence to suggest the extent of this depends on the bacterium strain and/or dose.

4. There is good evidence to show no significant difference in pain score or bloating score for single probiotics, both as a group and individually, compared with placebo. There is a significant difference for combined probiotics, with the probiotic giving significantly less pain and bloating.

5. There is weak evidence to show no significant difference in the number of people with bloating for combined probiotics compared with placebo.

6. There is good evidence to show that the use of probiotics (single or combination) resulted in participants reporting no significant difference in bowel habit.

7. There is good evidence to show that high doses of Bifidobacterium infantis (10^{10} CFU) in capsule form are significantly less effective than moderate doses (10^8 CFU); moderate doses are more effective than low doses (10^6). There is weak indirect evidence to show that this reduction in effect at high doses does not occur when probiotics are delivered in a drink.

8. There is fair evidence to show no significant difference between Lactobacillus salivarius UCC4331 and Bifidobacterium infantis 35624, in global symptoms, pain, bloating or stool scores.

9. There is a moderate amount of weak evidence to show no significant difference in the number of people with improvement in global symptoms or with bloating, between those given the prebiotic, Fructooliogsaccharide, in comparison with placebo.

EVIDENCE TO RECOMMENDATION

The review evidence suggests that some probiotics are effective in people with IBS, but others are not. The effect is dose and strain dependent, and the method of ingestion is also important. Although, there is some evidence from single trials, the GDG did not feel able to recommend named bacteria or probiotic products. On the other hand, it was the view of the GDG that probiotics were not harmful (unless they came from an unreliable source), they were widely available and it might benefit people with IBS if they experimented with probiotics as part of their diet. The GDG agreed there was insufficient evidence to make a recommendation on prebiotics.

RECOMMENDATION

People with IBS who choose to try probiotics should be advised to take the product for at least 4 weeks while monitoring the effect. Probiotics should be taken at the dose recommended by the manufacturer.

7.5 Aloe vera

SELECTION CRITERIA

The selection criteria described in the general methodology section were used.

DESCRIPTION OF STUDIES

Types of Studies

Two randomised trials were included (Odes and Madar 1991; Davis 2006) and two excluded studies are listed in Appendix E, along with reasons for exclusion.

Types of participants

All participants in Davis (2006) had IBS (28% IBS-C, 52% IBS-D, 20% IBS-A); participants had to be between 18 and 65 years, have no other co-morbidities and had to have previously failed conventional management of IBS defined as antispasmodics, bulking agents and dietary interventions. Constipation was defined as per Rome II criteria.

Odes and Madar (1991) had 11/32 people with IBS-C (the rest had simple constipation); people with IBS-D or IBS-A were excluded; participants had to have been receiving laxative therapy for constipation for a minimum of two years as an indication of severity. Participants had previously received other treatments including diet and enemas, but it is not stated if they were refractory to treatment.

Types of intervention

Davis (2006) used aloe vera gel made up in a pink mango flavoured syrup. The dose was 50 ml taken four times a day for one month. The placebo was a matching pink mango flavoured inert syrup.

Odes and Madar (1991) used a capsule laxative preparation made up of celandin, aloe vera and psyllium in ratio of 6:3:1 (total fibre content 47%) given for 28 days. The aloe vera fibre was derived from leaves of *Aloe socotrine* and contains anthraquinone. The dose was one 500mg capsule per day taken with water at bedtime increasing to a maximum of three capsules a day. The placebo capsule was of identical appearance but contained no active ingredient. Participants in Odes and Madar (1991) were given no dietary modification advice. No additional medication was prescribed throughout the treatment period but people could continue with prescribed laxative medication, provided that the dose and frequency were recorded in the study data sheet. Davis (2006) did not state if other medications could be continued.

METHODOLOGICAL QUALITY

Davis (2006) used a computerised random numbers table to generate the randomisation schedule. Allocation concealment was implemented in this study; the pharmacist held the

randomisation code. Both studies were double blind. Davis (2006) carried out an *a-priori* sample size calculation.

In Davis (2006), 58 people were randomised. 49 completed the protocol to one month and 41 to three months (i.e. data missing for 17/58 (29%) overall; 33% in the placebo group and 26% in the active group). In Odes and Madar (1991), 35 people were randomised. Three people (placebo) withdrew citing lack of benefit as reason and were excluded from the analysis because of incomplete data.

The groups in both trials were comparable at baseline as regards age, gender, duration and severity of condition, but Odes and Madar (1991) reported that the treatment group had significantly higher pain scores at baseline.

Overall, neither study was considered to have higher potential for bias.

RESULTS

In view of the differences in population and interventions, these two studies were reported separately.

A. Aloe vera gel versus placebo

One study (Davis 2006) in compared aloe vera gel with placebo in 58 people with IBS.

1. Global improvement of symptoms

The primary outcome was the number of people with an improvement in global symptom score (pain; distension; bowel habit, and; quality of life). The symptom score was derived by adding the scores of individual symptoms and the proportion of days symptoms occurred with a maximum score of 500. A reduction of 50 points was defined as improvement. Participants were assessed at one month and at three months post-intervention. The forest plots below illustrate that there was no significant difference between the active and placebo treatment for global improvement of symptoms, although the confidence interval was fairly wide.

Figure 1a: Global improvement of symptoms at one month
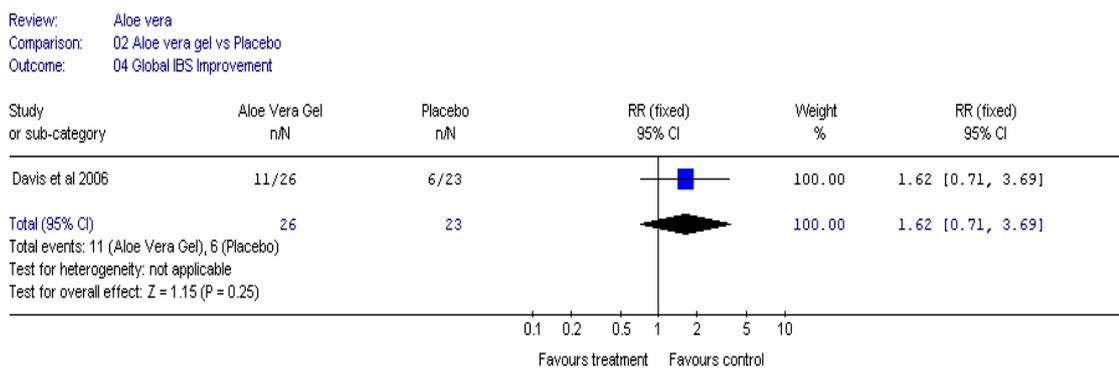
Figure 1b: Global improvement of symptoms at three months
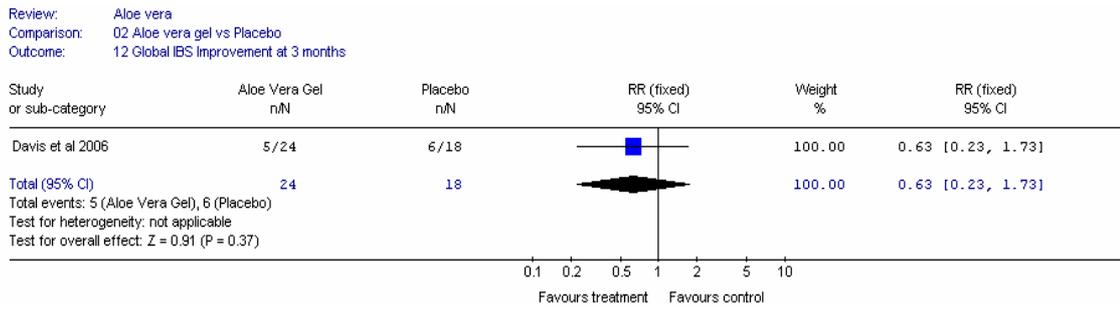
Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 12 Global IBS Improvement at 3 months

Study or sub-category	Aloe Vera Gel n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Davis et al 2006	5/24	6/18		100.00	0.63 [0.23, 1.73]
Total (95% CI)	24	18		100.00	0.63 [0.23, 1.73]

Total events: 5 (Aloe Vera Gel), 6 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.91 (P = 0.37)

0.1 0.2 0.5 1 2 5 10
Favours treatment   Favours control

2. Individual symptoms

a) Pain

There was no significant difference, either at 1 month or at 3 months, in the change in pain score, on a scale of 0 to 100%, for which a positive change represented an improvement over baseline.

Figure 2a: Pain at one month

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 01 Mean change in Pain score at 1 month

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	26	6.15 (23.96)	23	-4.90 (25.89)		100.00	11.05 [-2.98, 25.08]
Total (95% CI)	26		23			100.00	11.05 [-2.98, 25.08]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.54 (P = 0.12)

-100 -50 0 50 100
Favours Control   Favours Treatment

Figure 2b: Pain at three months

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 05 Mean change in Pain score at 3 months

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	24	11.96 (21.75)	18	-3.11 (32.98)		100.00	15.07 [-2.48, 32.62]
Total (95% CI)	24		18			100.00	15.07 [-2.48, 32.62]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.68 (P = 0.09)

-100 -50 0 50 100
Favours Control   Favours Treatment

b) Bloating

There was no significant difference either at 1 month or at 3 months, in the change in distension score, on a scale of 0 to 100%, for which a positive change represented an improvement over baseline.

Figure 3a: Distension at one month

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 06 Mean change in distension score at 1 month

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	26	-0.65(29.01)	23	8.13(33.66)		100.00	-8.78 [-26.49, 8.93]
Total (95% CI)	26		23			100.00	-8.78 [-26.49, 8.93]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.97 (P = 0.33)

-100 -50 0 50 100
Favours Control   Favours Treatment
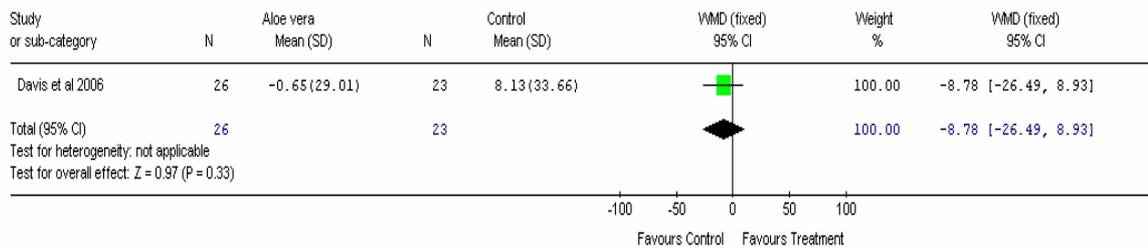
Figure 3b: Distension at three months

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 07 Mean change in distension score at 3 months

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	24	-3.33(26.89)	18	11.94(37.03)		100.00	-15.27 [-35.48, 4.94]
Total (95% CI)	24		18			100.00	-15.27 [-35.48, 4.94]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.48 (P = 0.14)

-100 -50 0 50 100
Favours Control   Favours Treatment
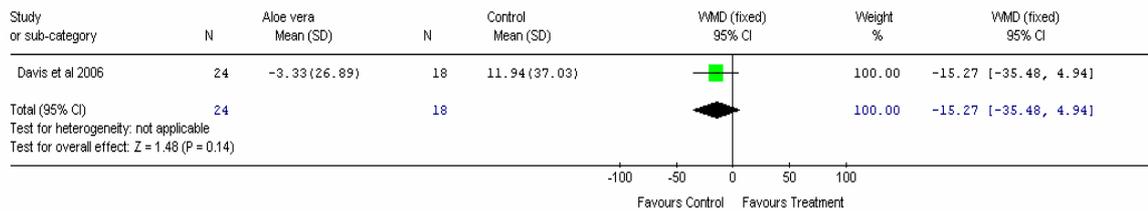
c) Bowel Habit

There was no significant difference, either at 1 month or at 3 months, in the change in bowel score, on a scale of 0 to 100%, for which a positive change represented an improvement over baseline.

Figure 4a: Change in bowel score at one month

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 08 Mean change in bowel habit score at 1 month

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	26	7.65(17.13)	23	3.78(28.48)		100.00	3.87 [-9.50, 17.24]
Total (95% CI)	26		23			100.00	3.87 [-9.50, 17.24]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.57 (P = 0.57)

-100 -50 0 50 100
Favours Control   Favours Treatment
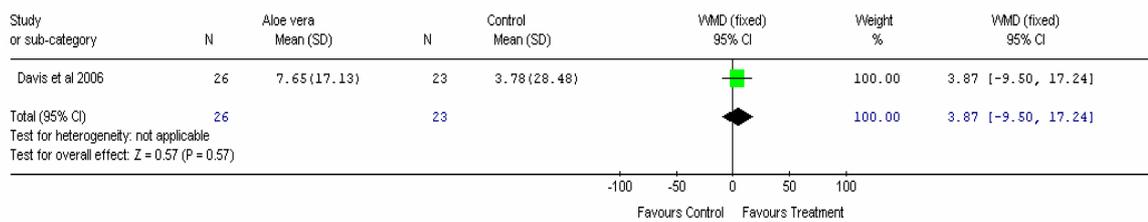
Figure 4b: Change in bowel score at three months

Review: Aloe vera
Comparison: 02 Aloe vera gel vs Placebo
Outcome: 09 Mean change in bowel habit score at 3 months

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Davis et al 2006	24	-1.17(29.78)	18	-0.17(29.78)		100.00	-1.00 [-19.20, 17.20]
Total (95% CI)	24		18			100.00	-1.00 [-19.20, 17.20]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.11 (P = 0.91)

-100 -50 0 50 100
Favours Control   Favours Treatment
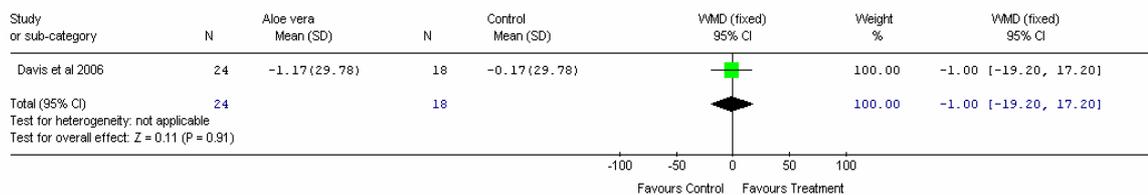
3. Quality of Life

There was no effect on quality of life at one month or at three months. The scale used was not stated.

Figure 5a: Change in quality of life at one month
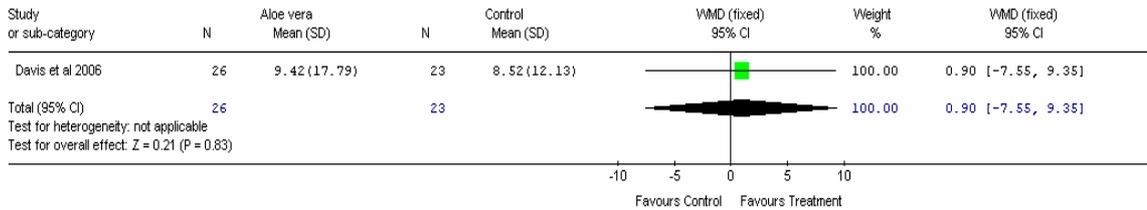
Figure 5b: Change in quality of life at three months
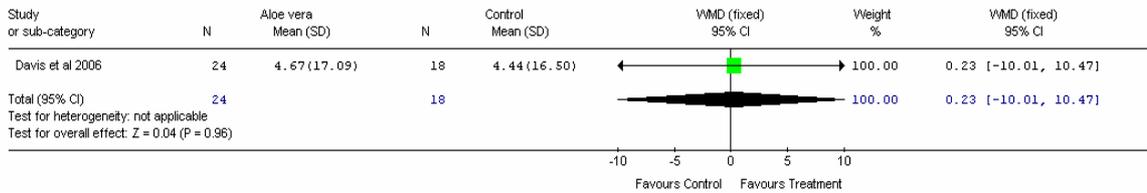
4. Adverse effects

2/31 people withdrew from the active group and 4/27 from the placebo group because of nausea and vomiting. The confidence interval was too wide to determine if there was a difference between groups.

Figure 6: Adverse effects
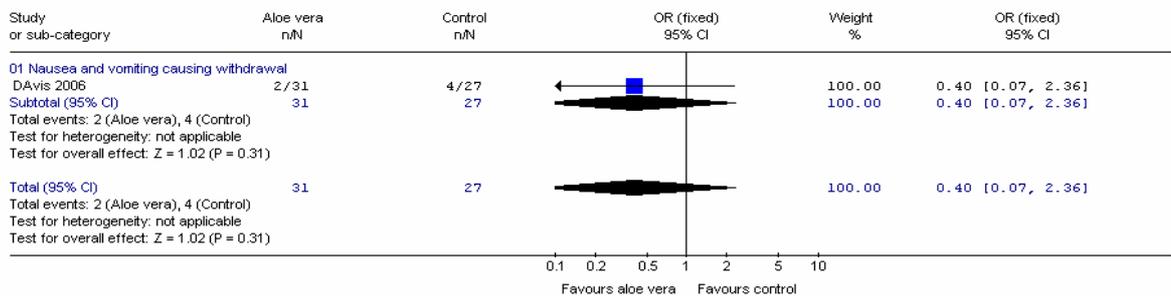
B. Combined capsule of celandin, aloe vera and psyllium versus placebo

1. Global improvement of symptoms

Odes and Madar (1991) did not report global symptoms.

2. Individual symptoms

One study with 32 participants reported differences in bowel habits (including frequency and consistency) and pain scores for the final two weeks of treatment compared with those in the 14 day pre-intervention run in period (Odes and Madar 1991). Eleven participants (34%) were identified as having IBS-C; the rest had simple constipation.

a) Pain

There was no significant difference in pain scores (number of episodes of pain per week) between groups.

Figure 7: Number of episodes of pain per week

Review:	Aloe vera						
Comparison:	01 Combination Aloe vera Capsule vs Placebo						
Outcome:	01 Pain score						

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Odes & Madar 1991	19	3.20(3.40)	13	1.80(1.90)		100.00	1.40 [-0.44, 3.24]
Total (95% CI)	19		13			100.00	1.40 [-0.44, 3.24]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.49 (P = 0.14)							

-10   -5   0   5   10
Favours Control   Favours Treatment

b) Bowel Habits

Compared to the placebo group, people in the intervention arm of the trial experienced a significant increase of 3.6 (95%CI 1.51, 5.69) in the mean number of bowel movements per week.

Figure 8: Stool frequency

Review:	Aloe vera						
Comparison:	01 Combination Aloe vera Capsule vs Placebo						
Outcome:	02 Stool Frequency						

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Odes & Madar 1991	19	7.90(3.90)	13	4.30(2.10)		100.00	3.60 [1.51, 5.69]
Total (95% CI)	19		13			100.00	3.60 [1.51, 5.69]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.37 (P = 0.0007)							

-10   -5   0   5   10
Favours Control   Favours Treatment

The consistency of the stools also improved. There was a statistically significant decrease in laxative use in the intervention group of -0.8 (95%CI -1.12, -0.47) on a scale of 1 to 3 and 16/19 people considered their bowel symptoms improved compared to 4/13 of the control group.

Figure 8: Stool consistency

Review:	Aloe vera						
Comparison:	01 Combination Aloe vera Capsule vs Placebo						
Outcome:	03 Constipation (low score=better)						

Study or sub-category	N	Aloe vera Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Odes & Madar 1991	19	1.60(0.40)	13	2.40(0.50)		100.00	-0.80 [-1.13, -0.47]
Total (95% CI)	19		13			100.00	-0.80 [-1.13, -0.47]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 4.81 (P < 0.00001)							

-10   -5   0   5   10
Favours treatment   Favours control

No adverse effects were reported and all participants reported that the capsules were easy and convenient to use.

Comment

Whilst this study shows a significant positive effect for bowel habit, we noted that this was a small study (35 patients) in a population of which only one-third had IBS. In addition, the intervention used a combination of aloe vera (30%), celandin and psyllium (soluble fibre). Therefore, it was not possible to attribute the effect to aloe vera alone, and this study was not included in the evidence statements.

SAFETY DATA

The following safety data is based on a systematic review of the scientific literature (case reports and systematic reviews) edited and peer reviewed by contributors to the US Natural Standard Research Collaboration (2006).

Adverse effects

The use of aloe by mouth for laxative effects can cause abdominal cramps and diarrhoea. Adverse effects, reported in a small number of studies, include low blood sugar levels and electrolyte imbalance, particularly lowered potassium levels.

Drug interaction

Use of aloe with laxative drugs may increase the risk of dehydration, electrolyte imbalance, potassium depletion and changes in blood pH.

Oral preparations of aloe have been reported to lower potassium levels, which may impact on the effectiveness of drugs used to manage heart rhythm disturbances, heart disease and renal disease.

Oral preparations of aloe may lower blood sugar, so have the potential to interact with drugs used in the management of diabetes.

Aloe vera should not be used by individuals who may be at increased risk from the aforementioned adverse effects, particularly people with heart disease, kidney disease, diabetes and blood disorders.

GDG DISCUSSION

The GDG expressed concerns that people with IBS purchase aloe vera products at considerable expense without evidence of effectiveness. The GDG also expressed concerns about the adverse effects of oral preparations of aloe vera, about which there was little awareness.

HEALTH ECONOMIC EVIDENCE

The cost-effectiveness of aloe vera was not estimated as it is not prescribed, but purchased by patients as a food supplement.

EVIDENCE STATEMENTS

1. There is fair evidence to show no significant effect of aloe vera, in comparison with placebo, in global improvement of symptoms, pain, bloating, bowel score or quality of life.

2. There is limited evidence of potentially serious adverse effects associated with oral aloe preparations.

EVIDENCE TO RECOMMENDATION

There is only one trial of aloe vera in people with IBS, but this gave fair evidence to show a lack of effectiveness. The GDG took this into account, together with aloe vera's potentially serious adverse effects, especially for people with comorbidities. Since aloe vera is a commercially available product that people with IBS pay for at considerable expense, the GDG wished to highlight these points by discouraging its use. Clinicians and people with IBS should be made aware of the lack of effectiveness and potential adverse effects.

RECOMMENDATION

Healthcare professionals should discourage the use of aloe vera in the treatment of IBS.

7.6 Exclusion Diets

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to this review and are reported below.

Types of studies

For intervention studies, randomised trials (RCTs) and quasi-randomised studies, examining the use of dietary manipulation/exclusion for the treatment of IBS were preferred. Crossover trials with a washout period of less than 2 weeks were included but treated with caution. Double-blind placebo controlled studies are technically difficult and most elimination diet studies use a non-randomised open dietary elimination and re-challenge design. This has the potential to introduce bias due to the large placebo effect identified in IBS patients. For this review non-randomised studies were also permitted. Studies were restricted to the English language, but the date was not restricted.

Types of intervention

Interventions were to be included if they referred to an exclusion diet (excluding certain foods) or an elimination diet (only allowing certain foods):

- Lactose restricted diet
- Elimination diet based on foods with IgG4 titres >250µg/l
- Elimination diet based on food challenge and re-challenge
- Elimination diet based on patient-reported intolerance and re-challenge
- Elimination diet using lamb, rice and pears
- Fasting therapy.

Types of comparisons

The following types of comparisons were to be included:

- True diet versus sham diet
- Elimination diet and food challenge with foods that had been identified as potential causes of intolerance
- Fasting therapy versus usual treatment.

Sensitivity analyses

The following sensitivity analyses may be considered:

- Setting (primary/secondary care)
- Blinding of patients.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

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. The search strategies are listed in Appendix B.

The search strategy identified 957 studies. The titles and abstracts of these studies were assessed. Of these, 33 that were potentially relevant to the review were retrieved in full. The reference lists of the retrieved studies were inspected for potential papers for inclusion in the review but none were identified. Sixteen studies were included in the review. The excluded studies are listed in Appendix E, along with reasons for exclusion.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

There were two randomised trials included (Atkinson 2004; Symons 1992). There were three reviews: one of which examined the evidence for the role of food hypersensitivity in IBS (Zar 2001) and the others were systematic reviews of non-randomised evidence for the dietary treatment of IBS (Niec 1998; Burden 2001). The remaining fourteen included studies were non-randomised studies (Bentley 1983; Böhmer and Tuynman 1996; Drisco 2006; Hawthorn 1991; Hunter 1985; Jones 1982; Kanazawa and Fukudo 2006; McKee 1987; Nanda 1989; Parker 1995; Petitpierre 1985; Smith 1985; Zar 2005; Zwetchkenbaum and Burakoff 1988).

Study Design

One study had a crossover design: Symons (1992) stated that the patients were randomised to interventions on two days, following a 12 hour fast, but the washout period was not clear.

Setting: all the studies included patients in secondary care, many of whom had not responded to previous treatment for IBS symptom management.

The majority of studies took place in the UK (Atkinson 2004; Bentley 1983; Hawthorn 1991; Hunter 1985; Jones 1982; McKee 1987; Nanda 1989; Parker 1995; Smith 1985; Zar 2005; Zwetchkenbaum and Burakoff 1988). One took place in Japan (Kanazawa and Fukudo 2006), two in Europe (Böhmer and Tuynman 1996; Petitpierre 1985), two in the US (Drisko 2006; Zwetchkenbaum and Burakoff 1988) and one in Australia (Symons 1992).

Population

All studies included patients with a diagnosis of IBS, although the definition varied between studies. Four studies used Rome criteria: Rome I (Hawthorne 1991); Rome II (Atkinson 2004, mean duration of IBS over 10 years; Drisko 2006; Zar 2006); Zar (2006) also predefined the IBS type. Two studies used the Manning Criteria (Kanazawa and Fukudo 2006; Symons 1992). Two studies used a definition described by the author (Böhmer and Tuynman 1996; Parker 1995). Seven studies simply said the patients 'had IBS' (Jones 1982; Bentley 1983; McKee 1987; Petitpierre 1985; Smith 1985; Nanda 1989; Zwetchkenbaum and Burakoff 1988). None of the studies stated that any patients had IBS as result of gastrointestinal infection.

Atkinson (2004) did not state whether patients had bloating and/or pain but described the symptom severity to be severe. Another study described the duration and frequency of symptom episodes (Nanda 1989), and one study reported duration of symptoms and the percentage of patients with pain, bloating and urgency (Hawthorne 1991). The remaining studies did not state the severity of symptoms. The age range of patients was 18 to 80 years with the average mean age being approximately 28 to 44 years. None of the studies particularly identified elderly patients. All studies had more women than men.

Interventions

Fructose-Sorbitol Dose 1 versus Dose 2

The RCT, Symons (1992), compared the difference in symptom provocation in IBS patients using two different doses of fructose–sorbitol solution. The lower dose solution was made up of 20g fructose and 3.5g sorbitol in 200 ml water; the higher dose contained 25g fructose and 5g sorbitol in 250ml water, i.e. a comparison of 17.5 and 20g/litre of sorbitol, for a constant concentration of fructose 100g/litre. Thirty-nine patients (15 IBS patients and 24 healthy controls) were randomised to receive the higher or lower dose on different days, and results were reported separately for the two population groups.

Exclusion diets

The other RCT, Atkinson (2004), tested patients' blood for IgG antibodies against 29 foods. A true and a sham diet sheet were then prepared for each patient. The true intervention diet excluded those foods to which the patient had antibodies; the sham diet excluded an equal number of foods but not those to which the patients had antibodies. The sham diet also included an equally difficult-to-exclude staple food as the true diet (for example, cow's milk was replaced by potato, wheat with rice, yeast with whole egg, etc.).

Exclusion diets (non-randomised studies)

The majority of studies used a low allergenic diet, and initially excluded a range of foods, including dairy products, wheat, corn, yeast, eggs, rye, potatoes, onions, cocoa, citrus, coffee, tea spices, alcohol, peas, banana, additives, preservatives and tomatoes. Then an open or single-blinded food challenge re-introduced foods 2 to 7 days apart.

- One study used a diet of one meat, one fruit and distilled water (Jones 1982)
- One study used only lamb, rice and pears (Bentley 1983)
- One study used lamb, white fish, cabbage, carrots, peas, 'Ryvita', weak black tea and dairy free margarine (Smith 1985)
- Two studies used IgG4 antibody and mould guided exclusion diets (Drisko 2006; Zar 2006)
- One study used a lactose restricted diet, but gave no further details. Low lactose consumption was defined as less than 9 g per day (Böhmer and Tuynman 1996)
- One study used starvation followed by 5 days of re-feeding in hospitalised IBS patients (Kanazawa and Fukudo 2006).

Comparisons

One RCT compared true diet with sham diet (Atkinson 2004).

One RCT compared two different doses of fructose-sorbitol solution (Symons 1992).

The remaining studies used diet and food challenge in all patients. The duration of the exclusion diet ranged from seven days (Jones 1982) to six months (Zar 2006). The challenge tests used in the studies involved patients being placed on a diet excluding foods believed to provoke symptoms and then re-introducing the foods in a double-blind or controlled way.

OUTCOMES

I. RANDOMISED TRIALS

1. Global symptoms score

a) Global improvement of IBS score

Atkinson (2004) used a validated IBS symptom severity score with a range from 0 to 500. The scale took into consideration scores for pain, distension, bowel function and general well-being, with mild, moderate and severe cases indicated by scores of 75-175, 175-300 and >300 respectively. A reduction in score of 50 or more was regarded as a clinically significant improvement.

Atkinson (2004) also reported a global rating of IBS using the question: 'Compared with your IBS before you started the food elimination diet, are you now: terrible, worse, slightly worse, no change, slightly better, better or excellent?' Significant improvement was defined as 'better' or 'excellent'.

Symons (1992) used a symptom score composite: abdominal pain/discomfort, bloating, distension, belching, nausea, bowel frequency, flatulence and borborygmi were each scored on a scale of 0 to 3 with 0 = absent and 3 = severe.

2. Quality of Life

Atkinson (2004) assessed the patients using a validated quality of life scale that is sensitive to change in IBS (range 0 to 500).

3. Mental health

Atkinson (2004) assessed the patients using the Hospital Anxiety and Depression (HAD). This instrument scores anxiety and depression up to a maximum score of 21 for each parameter, and a score above 9 indicates significant psychopathology.

Symons (1992) did not record other outcomes.

METHODOLOGICAL QUALITY OF RANDOMISED TRIALS

The quality assessment for included studies is shown in Appendix D.

The methods of randomisation and allocation concealment were reported in one randomised study, both of which were classified as adequate (computer generated and the sequence was retained by a central telephone centre: Atkinson 2004). Symons (1992) gave no details of the methods of randomisation or allocation concealment.

Atkinson (2004) described an *a-priori* power calculation and used an intention to treat analysis. The groups were mainly comparable, except that the baseline IBS symptom score was higher in the intervention group (331.9 (SD 70.8) versus 309.0 (SD 78.5), which is not a significant difference (p=0.06)). The number of patients who withdrew from the studies or were lost to follow-up was minimal. Atkinson (2004) reported data from 131 of the original 150 patients (65/75 true and 66/75 sham diet groups) at 12 weeks (87%).

Symons (1992) did not describe an *a-priori* power calculation. All patients completed the study. The study reported no baseline data so it was not possible to judge whether the groups were comparable.

RESULTS

A. True diet versus sham diet

1. Global symptoms

a) Number of patients with improvement in global symptoms

Atkinson (2004) randomised 150 patients to true and sham exclusion diets. They reported the number of patients with improvement in global symptoms. There was no significant difference between true and sham diets; however, the confidence interval was fairly wide.

Figure 1

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 34 Exclusion diet versus sham diet
Outcome: 01 Global improvement (number of patients)

Study or sub-category	True diet n/N	Sham diet n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Atkinson 2004	18/65	11/66		100.00	1.66 [0.85, 3.24]
Total (95% CI)	65	66		100.00	1.66 [0.85, 3.24]

Total events: 18 (True diet), 11 (Sham diet)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.49 (P = 0.14)

0.1 0.2 0.5 1 2 5 10
Favours sham diet   Favours true diet

b) Global improvement of symptoms score

Atkinson (2004) reported the final global symptom score on a scale from 0 to 500 (where lower scores are better). There was no significant difference between true and sham diets, although the true diet was favoured.

Figure 2

Review: IBS 31 August 2006
Comparison: 34 Exclusion diet versus sham diet
Outcome: 02 Global improvement score

Study or sub-category	N	True diet Mean (SD)	N	Sham diet Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Atkinson 2004	65	232.00(58.00)	66	248.00(58.00)		100.00	-16.00 [-35.86, 3.86]
Total (95% CI)	65		66			100.00	-16.00 [-35.86, 3.86]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.58 (P = 0.11)

-100 -50 0 50 100
Favours true diet   Favours sham diet

2. Quality of life

Atkinson (2004) reported a significant improvement in quality of life change-from-baseline scores for the true diet compared with sham diet, but the confidence interval was fairly wide. The mean difference was 38.0 (95%CI 2.36, 73.64), for a sham diet change from baseline of 50 points. The scale was 0 to 500.

Figure 3

Review: IBS 31 August 2006
Comparison: 34 Exclusion diet versus sham diet
Outcome: 03 Quality of life (change score)

Study or sub-category	N	True diet Mean (SD)	N	Sham diet Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Atkinson 2004	65	88.00(107.00)	66	50.00(101.00)		100.00	38.00 [2.36, 73.64]
Total (95% CI)	65		66			100.00	38.00 [2.36, 73.64]

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.09 (P = 0.04)

-100 -50 0 50 100
Favours sham diet   Favours true diet

3. Mental health

There was a small significant difference between the sham and true diet groups in the HAD anxiety scores (scale 0 to 21), but no significant difference in the depression scores.

Figure 4: HAD anxiety

Review: IBS 31 August 2006
Comparison: 34 Exclusion diet versus sham diet
Outcome: 04 HAD anxiety score (change)

Study or sub-category	N	True diet Mean (SD)	N	Sham diet Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Atkinson 2004	65	1.60(1.58)	66	0.87(1.57)		100.00	0.73 [0.19, 1.27]
Total (95% CI)	65		66			100.00	0.73 [0.19, 1.27]

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.65 (P = 0.008)

-10 -5 0 5 10
Favours sham diet   Favours true diet

Figure 5: HAD depression

Review: IBS 31 August 2006
Comparison: 34 Exclusion diet versus sham diet
Outcome: 05 HAD depression score (change)

Study or sub-category	N	True diet Mean (SD)	N	Sham diet Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Atkinson 2004	65	0.42(1.40)	66	0.27(1.33)		100.00	0.15 [-0.32, 0.62]
Total (95% CI)	65		66			100.00	0.15 [-0.32, 0.62]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.63 (P = 0.53)

-10 -5 0 5 10
Favours sham diet   Favours true diet

Symptoms on reintroduction of excluded foods at end of trial period

Of the 131 patients who gave data at the end of the trial period in Atkinson (2004), 93 (41 in the true diet group and 52 in the sham diet group) agreed to attempt reintroduction of eliminated foods. The mean IBS symptom score significantly increased (i.e. worsened) more in the true diet group (83.3 points) than in the sham diet group (31 points, p=0.003; standard deviations not given).

The change in global symptom score also showed that significantly more patients in the true diet group worsened on reintroduction of foods to which they had IgG antibodies (i.e. those that had been excluded during the diet): 41.5% of these patients worsened on reintroduction of these foods, versus 25% worsening in the sham diet group on reintroduction of similar foods (to which they had not been shown to have antibodies), p=0.047. We noted though that the self-selecting group taking part in this section of the trial may not have been representative of the randomised groups.

B. Fructose-sorbitol solution dose 1 versus fructose-sorbitol dose 2

Fructose and sorbitol, when ingested together, are thought to provoke symptoms of IBS. Sorbitol is found naturally in fruits, particularly peaches, pears, and plums. It is also added to soft drinks and diet products. One study (Symons 1992) compared two concentrations of sorbitol in a mixed solution of fructose and sorbitol. Concentrations compared were 17.5 and 20 g/litre sorbitol (the fructose concentration was kept constant). We noted that the duration of the study was very short – two days in each phase.

1. Global symptoms

Symons (1992) used a symptom score composite: abdominal pain/discomfort, bloating, distension, belching, nausea, bowel frequency, flatulence and borborygmi were each scored on a scale of 0 to 3 with 0 = absent and 3 = severe. It is unclear what the maximum score is, but it could be 21. The data were expressed as median (interquartile range). For the IBS patients the total symptom score was significantly greater (i.e. more severe) following consumption of the higher concentration solution, compared to the lower concentration solution (p=0.04).

Intervention	Median symptom score (range) (n=15)
Lower F-S dose	1.5 (0 to 4)
Higher F-S dose	3.5 (1 to 9) *

* p = 0.04

II. NON-RANDOMISED STUDIES

All thirteen studies identified food intolerance with varying degrees of response to exclusion
diets (Bentley 1983, Böhmer and Tuynman 1996, Drisco 2006, Hawthorn 1991, Hunter 1985,
Jones 1982, Kanazawa and Fukudo 2006, McKee 1987, Nanda 1989, Parker 1995, Petitpierre
1985, Zar 2005, Zwetchkenbaum and Burakoff 1988), illustrated in Table 1.

Table 1: Non-randomised studies; exclusion diets and results

Study (Drop out rate)	No. of Patients	Diet	Results
LAMB, PEARS AND RICE DIET			
Bentley 1983 8/27 (29.6%)	27	**Diet:** 2 weeks duration; initially only lamb, pears and rice, then other foods introduced individually. **Challenge**: identified foods reintroduced on 3 occasions, 3 days apart	14/21 remission after ED. This is just significant, but wide CI. Taking into account drop outs and assuming they are treatment failures makes the result non significant. 10/21identified specific food intolerance – 8 had double blind challenge and 3/8 confirmed food intolerance originally identified.
Parker 1995 53/253(21%) 33/129 (25%)	253 (phase 1) 129 (phase 2)	**Diet**: 2 weeks ED comprising of lamb, pears, white rice and spring water **Challenge**: single food re-introduction at daily intervals **Phase 2**: less restricted diet	100/200 improved on diet Phase 2: 39/96 improved on diet
1 MEAT, 1 FRUIT AND DISTILLED WATER			
Jones 1982 4/25 (20%)	25 (6 = food challenge)	**Diet**: 1 week of single meat, 1single fruit & distilled water **Challenge**: hospital double blind challenge	4/25 refused diet. 14/21 improved and identified foods that provoked symptoms – this is just significant, but wide CI. Food challenge: 10/12 test solutions identified correctly – majority of foods that patients had identified as provoking symptoms were confirmed by food challenge.

Study (Drop out rate)	No. of Patients	Diet	Results
LOW ALLERGENIC DIET AND SIMILAR			
McKee 1987 (not stated)	40	**Diet**: 1 week low allergenic diet, excluded all sources of salicylates, amines, glutamates, additives **Challenge**: Open, frequency not stated	6/40 remission during exclusion diet
Nanda 1989 11/200 (5.5%)	200	**Diet**: 3 week low allergenic, excluded dairy, cereals, citrus fruit, potato, tea, coffee, additives. **Challenge**: open challenge every 2 days	91/189 remission during ED 73/189 found specific foods by open food challenge Follow up approx 14 months 73/91 responders still compliant with ED
Petitpierre 1985 0% drop out	24	**Diet**: 3 weeks Low allergenic **Challenge**: open and single blind, Frequency not stated.	3/24 remission with ED but challenges negative 14/24 specific foods identified and confirmed by blind challenge 7/24 symptoms unchanged
Hawthorne 1991 5/38 (9.5%)	38	**Diet**: 2 weeks exclusion of dairy, cereals, yeast, eggs, citrus fruits, tea, coffee, alcohol, potato, onion, tomato, banana, peas. **Challenge**: foods re-introduced at 2 day intervals following set protocol	5/38 refused to try diet 18/33 improved: 16/18 identified foods which exacerbated symptoms, 2/18 did not. 15/33 had no improvement from diet Follow-up of 16 improvers at 3 to 45 months (results not reported).
Smith 1985 Not stated	28	**Diet**: 2 weeks diet allowed, lamb, white fish, cabbage, carrots, peas, Ryvita, dairy free margarine, black tea. **Challenge**: foods were reintroduced at 2 day intervals in responders	11/28 improved Follow-up at 1yr: 7/9 responders were still well and maintaining diet.

Study (Drop out rate)	No. of Patients	Diet	Results
FOOD EXCLUSION BASED ON IgG ANTIBODIES			
Drisko 2006 All patients completed study and follow up at 1 year.	20	**Diet**: 2-3 weeks duration; tailored food exclusion based on IgE and IgG food and mould panels. **Challenge**: food reintroduced over several months	Statistically significant reduction in stool frequency (diarrhoea) from 4.29 (2.49) stools per day to 3.43 (1.22) Pain score (1 to 5 scale) 3.65 (1.12) to 2.71(1.38) p>0.5 (not significant) Overall QoL scores (100 point scale, high = better) 46.51(21.08) to 67.22(20.92) p<0.001
Zar 2005	25	**Diet**: 6 months duration; IgG4 antibody titres to 16 common foods. These were excluded if titres >250µg/l – most common exclusions: milk, cheese, eggs, beef, lamb, wheat and tomato. On average patients excluded 8 (3 -13) foods	Symptom score (scale 1-100) 21/25 showed statistically significant improvement in pain severity p<0.001, pain frequency p=0.034, bloating severity p=0.001, improved bowel habit p=0.004, QOL p=0.008 Follow up at 6 months: 6/15 lost to follow-up, the remaining patients maintained improvement
FOOD EXCLUSION PARTLY BASED ON IgG ANTIBODIES			
Zwetchkenbaum and Burakoff 1988 1/10 (10%)	10	**Diet**: 2 week exclusion of foods identified from patient food diaries, skin testing and IgG testing. **Challenge**: open for 2 days, 2 days apart Double blind provocation for patients showing persistent exacerbation of symptoms on open food re-introduction.	3/9 remission of symptoms with ED; 6/9 had no change in symptoms. Challenges did not identify provoking food

Study (Drop out rate)	No. of Patients	Diet	Results
STARVATION DIET			
Kanazawa and Fukudo 2006 No drop out	58 hospitalised pts.	**Diet**: 10 days starvation diet followed by 5 days re-feeding (from 225 – 2100kca). Patients were allowed 2 litres of water + 500 ml xylitol solution. Patients also received brief psychotherapy for 12 weeks hospital stay.	Starvation significantly decreased the following symptoms: abdominal pain/discomfort, distension, diarrhoea, anxiety and QOL (p=0.001), nausea (p<0.01), anorexia p=0.02)
LACTOSE RESTRICTED DIET			
Böhmer and Tuynman 1996 No drop out.	105 (70 IBS patients, 35 healthy controls)	**Diet**: 6 week duration; lactose restricted diet (no details given)	17/70 IBS patients had positive hydrogen breath test and glucose blood test compared to 2/35 controls. There was no difference in symptom score between groups at baseline. After dietary therapy, statistically significant decrease in symptom score in lactose intolerant group p<0.001. The lactose tolerant group had no change in scores. The incidence of lactose malabsorption was 4 times higher in IBS group than in healthy controls.

The non-randomised studies have been grouped according to the type of diet, and the following general conclusions can be drawn:

o There is some evidence to suggest that a simple diet of lamb, pears and rice or one meat and one fruit may improve symptoms

o A low allergenic diet does not appear to give remission in IBS symptoms

o Food exclusion diets based on IgG antibody testing appear to be effective in improving symptoms

o A starvation diet significantly decreased symptoms, but this was in a group of hospitalised patients, and is not applicable to primary care

o A lactose restricted diet gave a significant decrease in symptom score for lactose intolerant patients, but not for the lactose tolerant group.

It should be remembered that this evidence is from non-randomised studies, so its overall quality is reduced.

GDG DISCUSSION

The GDG discussed this review at length. They noted that lifestyle change and adjustment of diet according to symptoms can offer relief to people with IBS. Further dietary manipulation in the form of avoidance of specific foods offers improvement for up to two-thirds of people suffering from IBS. However, the GDG was concerned that exclusion diets undertaken without the advice of a dietician could lead to malnourishment and deficiencies.

Diet and nutrition are fundamental in the management of IBS to avoid malnutrition and to achieve optimal symptom control. The gold standard diagnosis for intolerance to a food is by elimination and reintroduction. Intolerance is demonstrated if symptoms resolve on elimination and reappear on reintroduction. Importantly, dietary advice will vary depending on symptoms, e.g. diarrhoea and/or constipation, abdominal bloating and therefore needs to be tailored to the individual to manage symptoms. Consequently, the GDG did not wish to produce a list of possible suspect foods, or to encourage patients to adopt a trial-and-error approach.

The GDG also emphasised that the dietitian should be registered and therefore trained to work in clinical settings and able to advise on all aspects of diet. The GDG noted that, currently, anyone can call themselves a nutritionist, regardless of qualifications. The Nutrition Society is the professional organisation for nutritionists; registration can be checked at www.hpcheck.org.

The GDG commented that an implementable dietary assessment tool would be useful, but accepted that such a tool had to be validated before it could be recommended.

Finally, the GDG recommended that the term, 'balanced diet' should be avoided because it was not specific. They commented that the 5-a-day public health recommendation could be problematic, especially for IBS-D patients.

The consensus was as follows:
o Patients should have a dietary assessment at initial consultation, and this should include examining eating patterns and when patients are eating
o Regular eating patterns should be encouraged
o Exclusion diets should be reserved for severe cases of IBS and should be carried out only under the advice of a dietician
o Dietary referral would be a useful option for mild IBS.

Several of these consensus points have been included in recommendations in the general dietary lifestyle and advice section.

EVIDENCE STATEMENTS

1. There is fair evidence to show no significant difference in global symptoms, between true and sham exclusion diets (i.e. foods excluded for which the patient had or had not IgG antibodies).

2. There is fair evidence to show a significant difference in quality of life, favouring a true exclusion diet, in comparison with a sham diet.

3. There is weak evidence to show that reintroduction of excluded foods to patients previously given a true exclusion diet resulted in significant worsening of global symptoms in comparison with those given a sham diet.

4. There is weak evidence to suggest that food exclusion diets based on IgG antibody testing are effective in improving symptoms, but a low allergenic diet does not appear to be effective.

5. There is weak evidence to suggest that a simple diet of lamb, pears and rice or 1 meat and 1 fruit may improve symptoms.

6. There is weak evidence that a lactose restricted diet gave a significant decrease in symptom score for lactose intolerant patients, but not for lactose tolerant patients.

7. There is limited evidence to show significantly more severe symptoms following consumption of a solution containing 20 g/litre sorbitol, compared to one with 17.5 g/litre, in the presence of fructose.

EVIDENCE TO RECOMMENDATIONS

The GDG took into consideration the clinical effectiveness evidence. Although there was some evidence to support the use of exclusion diets, the GDG believed that such diets should only be undertaken with the specialist help of a dietitian to ensure the diet remains well balanced.

RECOMMENDATION

If diet continues to be considered a major factor in a person's symptoms and they are following general lifestyle/dietary advice, they should be referred to a dietitian for advice and treatment, including single food avoidance and exclusion diets. Such advice should only be given by a dietitian.

---

8 PHARMACOLOGICAL INTERVENTIONS

<div style="border:1px solid">

Clinical Questions

1. Are antispasmodics effective in managing IBS symptoms?

2. Are laxatives effective in the management of IBS?

3. Are anti-motility agents effective in symptom control in IBS?

4. Do tricyclics and SSRI's have a role in the management of IBS symptoms?

</div>

BACKGROUND

The pharmacological management of IBS can provide clinicians with a major therapeutic challenge. People with IBS may present with a multi-symptom profile and it is unlikely that all patients will respond in the same way to the same single agent. There have been no new drugs specifically developed for the treatment of IBS in the last twenty years and the quality of the trials in the majority of pharmacological agents currently available is variable and often conducted on secondary populations. The drug management strategy should be based on the nature and severity of the symptoms and individual or combinations of medication directed at the predominant symptom/s. Irritable bowel syndrome can present with pain, constipation, or diarrhoea. Antispasmodic, TCA and SSRI drugs may relieve IBS pain. Antimotility drugs may relieve diarrhoea. Opioids with a central action such as codeine are better avoided because of the risk of dependence. Laxatives may be needed to relieve constipation. It is important to be sure that the patient *is* constipated. People who complain of constipation need to understand that bowel habit can vary considerably in frequency without doing harm. Some people tend to consider themselves constipated if they do not have a bowel movement each day. Misconceptions about 'normal' bowel habits have led to excessive or inappropriate laxative use. Laxative abuse may lead to hypokalaemia.

In some people with IBS there may be important psychological aggravating factors which respond to reassurance and possibly specific treatment e.g. with an antidepressant.

Antispasmodics

The abdominal pain experienced by people with IBS may be a result of irregular and intermittent intestinal contractions along the length of the colon. This may lead to symptoms of abdominal pain, bloating and gas. Pain is most common after a meal and may last for several hours.

Antispasmodics can be separated into two main categories: antimuscarinics, and smooth muscle relaxants. Antimuscarinics reduce intestinal motility; smooth muscle relaxants directly

relax intestinal smooth muscle. The use of antispasmodics is primarily to relax the smooth muscles of the gut, helping to prevent or relieve the painful cramping spasms in the intestines. They are typically taken 30 to 45 minutes before meals.

Antimotility agents

Diarrhoea is associated with alterations of fluid and electrolyte movement in either the small intestine or the colon. This can be due to decreased intestinal absorption, altered intestinal motility, or increased intestinal secretions (e.g. due to bacterial enterotoxins or laxatives). Antimotility agents are used to manage acute or chronic diarrhoea or exacerbations of chronic diarrhoea and work by altering one or more of these mechanisms.

Antimotility agents for IBS can be separated into four main categories: codeine phosphate; co-phenotrope (mixture of diphenoxylate hydrochloride and atropine sulphate in the mass proportions 100 parts to 1 part); loperamide and morphine-containing preparations. Prolonged codeine use can lead to dependency. Loperamide is considered especially useful as it tends to increase anal sphincter tone.

Laxatives

Laxatives can be separated into four main categories: bulk forming laxatives; stimulant laxatives; faecal softeners and osmotic laxatives. Bulk-forming laxatives relieve constipation by increasing faecal mass, which stimulates peristalsis; adequate fluid intake should be maintained to avoid intestinal obstruction. Stimulant laxatives work by increasing intestinal motility, but they often cause abdominal cramps. Faecal softeners may lubricate the passage of stools and/or soften them. Osmotic laxatives increase the amount of water in the large bowel, either by drawing fluid from the body into the bowel or by retaining the fluid with which they were administered. The route of adminstration for laxatives may be oral or rectal. Laxatives can be used in two ways: as short-term rescue medication or as longer-term maintenance treatment. There is no evidence that long term laxative use damages the bowel.

Tricyclics and Antidepressants

Since their introduction approximately fifty years ago, antidepressants have been used in a variety of gastrointestinal (GI) conditions. In the last twenty years antidepressants have been increasingly used in the treatment of functional GI disorders such as IBS. The prevalence of anxiety and depressive disorders is high in patients with severe and/or intractable IBS and may be present to some degree in all IBS patients. Antidepressants appear have an analgesic effect separate to their antidepressant effect. Visceral pain syndromes including IBS may be effectively treated by a range of therapies, including antidepressants that modulate the interactions between the central and enteric nervous systems. Tricyclics also have a peripheral anticholinergic action in addition to their central analgesic and antidepressant actions.

Antidepressants can be divided into three major classes: tricyclics and related antidepressants; selective serotonin re-uptake inhibitors (SSRIs), and monoamine oxidase inhibitors (MAOIs). There are other antidepressants that do not fit easily into these categories: Duloxetine (Cymbalta); Flupentixol (Fluanxol); Mirtazapine (Zispin Soltab); Reboxetine (Edronax); Tryptophan (Optimax), and; Venlafaxine (Efexor).

8.1 Laxatives

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the laxatives review and are reported below.

Types of studies

For longer-term studies, the GDG decided that the washout period for this review should be at least two weeks. Trials with no washout were not included in the analysis and trials with one week washout were considered if there was no other information. Crossover studies are not appropriate for short-term (rescue) medication.

Types of participants

For this review, participants with IBS were included, but the review was extended to include people with simple constipation as well because the same drugs are used. Studies in these participants were regarded as indirect as far as population was concerned.

Types of intervention

Studies were to include the following interventions:

- **Bulk-forming laxatives**, which, as a class, include some dietary fibres (see fibres review), but here we consider only non-dietary bulk forming agents:
    o Ispaghula husk (trade names: Fibrelief®; Fybogel®; Isogel®; Ispagel Orange®; Regulan®)
    o Methylcellulose (trade name: Celevac®)
    o Sterculia (trade names: Normacol®; Normacol Plus®).

- **Stimulant laxatives:**
    o Biascodyl (trade name: *Dulco-lax®), given as either oral tablets or rectal suppositories
    o Docusate sodium (synonym: dioctyl sodium sulphosuccinate; trade names: Dioctyl® (oral); Docusol® (oral); Norgalax Micro-enema® (rectal))
    o Glycerol (synonym: glycerine), given as rectal suppositories
    o Senna (non proprietary tablets; trade names: Senokot® granules; Manevac® granules (senna fruit 12.4%, ispaghula 54.2%)), given as an oral preparation

- o Sodium picosulphate (Trade names: Laxoberal® (oral elixir); *Dulco-lax® Liquid (oral elixir); *Dulco-lax Perles® (oral capsules)), given as an oral preparation.

  * note that the trade name Dulco-lax is used for different drugs, but with different qualifiers.

- **Faecal softeners:**
  - o Arachis oil (Trade name: Fletchers' Arachis Oil Retention Enema®), given as a rectal preparation
  - o Liquid paraffin (Liquid Paraffin Oral Emulsion, BP), given as oral preparation.

- **Osmotic laxatives:**
  - o Lactulose (trade names include: *Duphalac®*; *Lactugal®*; *Regulose®*), given as oral solution
  - o Macrogols (synonyms: polyethylene glycols, PEG; trade names: Idrolax® (oral powder, PEG 4000); Movicol® (oral powder, PEG 3350); Movicol®-Half (oral powder, PEG 3350))
  - o Magnesium salts (Magnesium Hydroxide Mixture, BP, oral aqueous suspension; Liquid Paraffin and Magnesium Hydroxide Oral Emulsion, BP (oral aqueous suspension), Magnesium Sulphate (Epsom Salts); trade name: *Milpar®*)
  - o Sodium phosphate – this is not in the BNF but is in routine use and so was included by the GDG.

The following comparisons were included:
- Laxative versus placebo (or nothing)
- Laxative type 1 versus type 2
- Laxative dose 1 versus dose 2
- Laxative + another intervention versus the other intervention alone
- Laxative route of delivery 1 versus route 2
- Duration of treatment 1 versus duration 2.

NB: In spite of the large placebo effect associated with IBS, comparisons with no treatment are included.

The laxatives review was concerned with both longer-term maintenance treatment and short-term symptom relief.

The GDG decided that there should be a minimum duration of treatment of four weeks for maintenance in this review. Maintenance studies of shorter durations were not included in the analysis.

Subgroup analyses

We carried out subgroup analyses by type of laxative (bulk forming laxatives; stimulant laxatives; faecal softeners, and; osmotic laxatives); dose; route of delivery (oral, rectal), and; duration of intervention.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

The initial search identified a Cochrane Review (Quartero 2005, *Bulking agents, antispasmodic and antidepressant medication for the treatment of irritable bowel syndrome*). Searches were partly based on the terms in this review. 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. For this review, the search was extended to cover the population with simple constipation as well as IBS. The search strategies are listed in Appendix B.

The Cochrane review identified 11 studies, ten of which are included in the fibres review. The remaining study (Piai 1987) used a cellulose material, glucomannan, which is not used in the UK. The titles and abstracts identified by the NCC search strategy were assessed and fifty studies were retrieved in full. The reference lists for each of the retrieved studies were inspected for further potential papers, but none were identified. The 38 excluded studies are listed in Appendix E, along with reasons for exclusion. Searches were updated to June 2007 and a further two papers were identified and obtained from the authors with some further information (Wulkow 2007; Kienzle-Horn 2007).

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

Study Design

There were four crossover studies (Cleveland 2001; Connolly 1975; Quah 2006; Sobhani 1996) in which participants were allocated to receive both the intervention and control treatments during the course of the study, in a random order. Two of these studies (Cleveland 2001; Sobhani 1996) had either no washout period or it was not reported, in which case this was assumed to be none. The other two studies had a washout period of one week. The GDG had specified a washout period of two weeks minimum, but would consider the two one-week studies if there was no other data. No crossover studies reported first-period results only or individual patient data.

The remaining 15 studies had a parallel design (Attar 1999; Bouhnik 2004; Chaussade 2003; Corazziari 1996; Corazziari 2000; Dettmar 1998; DiPalma 2000; Hamilton 1988; Kienzle-Horn 2006; Kienzle-Horn 2007; Marlett 1987; Medoff 2004; Rouse 1991; Wang 2004; Wulkow 2007).

The GDG had specified a minimum treatment period of four weeks for each intervention in the maintenance studies. Two studies had a treatment duration of one week (Connolly 1975; Marlett 1987); one had ten days (Hamilton 1988); three had two weeks (Cleveland 2001; DiPalma 2000; Wang 2004); one had three weeks (Sobhani 1996). All these studies were transferred to the excluded studies table.

The remaining studies had durations of four weeks (Attar 1999; Bouhnik 2004; Chaussade 2003; Dettmar 1998; Kienzle-Horn 2007; Medoff 2004; Quah 2006 (1-week washout crossover); Rouse 1991); eight weeks (Corazziari 1996), and; 20 weeks (Corazziari 2000).

Two studies investigated the use of laxatives for acute constipation (Kienzle-Horn 2006; Wulkow 2007).

Twelve studies were therefore included in the analysis (11 parallel and one 1-week washout crossover trial: Attar 1999; Bouhnik 2004; Chaussade 2003; Corazziari 1996; Corazziari 2000; Dettmar 1998; Kienzle-Horn 2006; Kienzle-Horn 2007; Medoff 2004; Quah 2006; Rouse 1991; Wulkow 2007).

One of the remaining studies had more than two arms: Chaussade (2003) compared four PEG interventions; there were thus 14 comparisons in the laxatives review. The rest of the description of studies will focus on these studies/comparisons.

Two were conducted in the UK (Dettmar 1998 and Rouse 1991); one in sites in the UK and France (Attar 1999); seven in the rest of Europe, one in each of the USA and China.

Setting: Seven studies took place in primary care (Bouhnik 2004; Chaussade 2003; Dettmar 1998; Kienzle-Horn 2006; Medoff 2004; Rouse 1991; Wulkow 2007); and five were in secondary care (Attar 1999 (with 31% from geriatric institutions); Corazziari 1996; Corazziari 2000; Kienzle-Horn 2007; Quah 2006).

The majority of studies (7/12) had fewer than 100 patients, with two having 25 or fewer in the intervention arm (Corazziari 1996; Medoff 2004). Two studies had more than 200 patients in total (Chaussade 2003; Dettmar 1998).

Funding: Six studies had some industry sponsorship: Bouhnik (2004) was sponsored by Solvay Pharmaceuticals, manufacturers of lactulose; Chaussade (2003) was supported by a grant from by Hoffmann La Roche, the manufacturers of PEG 3350, and; the Dettmar (1998) authors were from Reckitt & Colman, manufacturers of Fybogel. Kienzle-Horn (2006), Kienzle-Horn (2007) and; Wulkow (2007) were all funded by Boehringer-Ingelheim GmbH, manufacturers of Bisacodyl, and Sodium Picosulphate.

Population

- Only one study (Medoff 2004) definitely included patients with IBS: 7/43 patients had a diagnosis of constipation predominant IBS, but separate results were not reported. All the other studies stated they had patients with simple constipation, usually defined as 3 or less, or 2 or less bowel movements per week. Most studies defined a minimum period of constipation symptoms, ranging from 3 weeks (Rouse 1991) to 12 months (Corazziari 1996; Corazziari 2000). The GDG suspected, however, that many of these studies may have had the classification of 'simple constipation' because this was the primary symptom being treated by laxatives, rather than the only symptom. Indeed, they thought it was quite likely that the patients had IBS. This was further investigated. Communication with the authors of one study (Kienzle-Horn 2007) revealed that abdominal pain was not excluded and the author believed that some patients entering the studies would have had IBS if they had been checked for it by their physician. Seven of these studies (Attar 1999; Bouhnik 2004; Corazziari 1996; Corazziari 2000; Chaussade 2003; Dettmar 1998; Rouse 1991) reported that some patients had pain and/or bloating. However, Dettmar (1998) and Rouse (1991) did not report sufficient duration to be defined as IBS. Two studies did not mention the incidence of pain or bloating before treatment (Kienzle-Horn 2006; Quah 2006) but the patients had had constipation for at least 3 months.

- In Attar (1999), 20% and 35% of the control group had pain and bloating respectively during the trial. The patients were stated to have had chronic idiopathic constipation for at least 3 months.

- In Bouhnik (2004), in the washout period, 45 and 53% had bloating at washout and 30 and 45% had pain for lactulose and PEG respectively. The patients were stated to have had chronic idiopathic constipation for at least 6 months.

- In Chaussade (2003), at baseline, the bloating score was ~ 3 points on a scale of 1 to 4 (considerable) and pain was 2.6. There was an implied use of the Rome II criteria for chronic idiopathic constipation, which the patients were to have had for at least 3 months

- In Corazziari (1996), in the run-in period 52-60% pts had pain and 84-91% had bloating. The patients were stated to have chronic non-organic constipation and had had this for at least 12 months.

- In Corazziari (2000), the pain and bloating scores were non zero, even after the patients had received PEG for 4 weeks. Chronic constipation was defined using Rome criteria and the patients had had chronic constipation for at least 12 months.

- Dettmar (1998) reported that the majority of the patients experienced abdominal symptoms, including pain, distension or flatulence. The patients were said to have 'simple constipation' and there were no details about the duration of constipation.

- In Rouse (1991) 53-54% patients in both groups had abdominal pain after seven days. Bloating was not mentioned. The patients were treated for chronic constipation, which they had had for at least 3 weeks.

The GDG concluded that all of these studies, with the exception of Dettmar (1998) and Rouse (1991) were likely to have some, if not all, patients with IBS. It was unclear if Kienzle-Horn (2006) and Quah (2006) included patients with IBS. Further details are given in the included studies table (Appendix C).

One study (Corazziari 2000) gave PEG electrolyte to patients for 4 weeks and then randomised only the responders (at least two bowel movements per week with no other defaecatory disturbances or more than three bowel movements per week) to PEG electrolyte or placebo.

The age range of participants across the studies was 18 to 89 years, with the mean age (where given) ranging from 42 to 58 years. All the studies had more women than men.

Interventions

The studies varied in the type of laxatives used:

- Two used stimulant laxatives:
    - One bisacodyl (Kienzle-Horn 2006)
    - One sodium picosulphate (Wulkow 2007).
- Eight had osmotic laxatives:
    - Four lactulose (Attar 1999; Bouhnik 2004; Quah 2006; Rouse 1991)
    - Five polyethylene glycol (Attar 1999; Chaussade 2003; Corazziari 1996; Corazziari 2000; Bouhnik 2004)
    - One sodium phosphate (Medoff 2004).
- One study (Dettmar 1998) allowed the patient any laxative (which was mainly lactulose) and also reported the lactulose patients as a subgroup.

Comparisons

The included studies covered the following comparisons:

- Four comparisons of laxatives versus placebo:
    - Two gave stimulant laxatives (Kienzle-Horn 2006; Wulkow 2007 treatment for acute episodes)
    - Two gave osmotic laxatives (Corazziari 1996; Corazziari 2000).
- Three studies compared a laxative with fibre:
    - Two compared an osmotic laxative (lactulose) with fibre (ispaghula), (Quah 2006; Rouse 1991)
    - One compared usual laxatives (mainly lactulose) with fibre (ispaghula) (Dettmar 1998).
- Seven comparisons of different types of laxative in the same class:
    - Two studies (Attar 1999; Bouhnik 2004) compared lactulose with PEG electrolyte (Osmotic Laxatives)

- Four comparisons of PEG 3350 (Transipeg) plus electrolytes versus PEG 4000 without electrolytes (Forlax) (Chaussade 2003 x4) (Osmotic Laxatives)
  - One comparison of Bisacodyl versus Sodium Picosulphate (Kienzle–Horn 2007).
- Three comparisons of different doses of an osmotic laxative:
  - One comparison of 11.8g versus 5.9g PEG 3350 (Transipeg) plus electrolytes (Chaussade 2003)
  - One comparison of 20g versus 10g PEG 4000 (Forlax) (Chaussade 2003)
  - One comparison of 10.56g (mean) versus 6.84g (mean) sodium phosphate.

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D.

An adequate method of randomisation was reported in four studies (Attar 1999; Medoff 2004; Quah 2006; Wulrow 2007), all of which used a computer generated method. The other studies did not state the method.

Allocation concealment was reported in three studies (Attar 1999; Bouhnik 2004; Quah 2006), both of which reported an adequate method, in which the statistician prepared the list and the investigators were unaware of the allocation (Attar 1999) or by telephoning a central office (Bouhnik 2004; Quah 2006).

Six studies reported that the patients were blinded to the interventions (Chaussade 2003; Corazziari 1996; Corazziari 2000; Kienzle-Horn 2006; Kienzle-Horn 2007; Wulrow 2007); these included all the placebo controlled studies. The remaining studies stated that the patients were not blinded, or could not have been because of differences between drugs in appearance and taste.

Five studies (Attar 1999; Bouhnik 2004; Kienzle-Horn 2006; Kienzle-Horn 2007; Wulrow 2007) described an *a-priori* power calculation. All studies included in the review demonstrated baseline comparability of the groups, apart from one study which was not comparable at baseline (Medoff 2004) for rectal irritation, which was greater in the group receiving four tablets.

There was loss to follow-up in the majority of studies, and all but one had less than 20% drop-outs. One study (Corazziari 2000) reported that more than 20% of patients in at least one arm (or overall) were not analysed or were lost to follow-up (attrition bias). In Corazziari (2000), for the first eight weeks 1/33 (3%) PMF and 4/37 (11%) placebo did not complete the period, but 10/33 (30%) PMF and 22/37 (59%) placebo did not complete the 20 weeks. Consequently, results at eight weeks only were taken for this study. In Quah (2006), 8/50 (22%) withdrew before receiving the interventions, then 3/21 (14%) withdrew from fibre group and 0% on lactulose. The GDG did not regard this level of missing data as significant.

Seven studies stated that they did not permit any concomitant medication that would change the GI motility (Bouhnik 2004; Chaussade 2003; Corazziari 1996; Kienzle-Horn 2006; Kienzle-Horn 2007; Quah 2006; Wulrow 2007). Five studies allowed the patients to have laxatives as relief medication: in two studies (Corazziari 1996; Corazziari 2000) there had to be five consecutive days without a bowel movement; in one study (Chaussade 2003) there had to be three consecutive days, after which the patients could have suppositories. In the other studies (Attar 1999; Medoff 2004) patients could use suppositories or microenemas for relief, apparently without restriction. In another study (Rouse 1991) 12/124 patients took other laxatives during the study and were considered to be protocol violators.

The risk of bias was assessed for each included study and no studies were excluded from the analysis (although the 20 week results for Corazziari 2000 were disregarded). The two studies in which laxatives could be taken apparently without restriction (Attar 1999; Medoff 2004) were regarded with caution.

RESULTS

I. Treatment for acute episodes of constipation

A. Laxatives versus placebo

Two studies (Kienzle-Horn 2006; Wulkow 2007) in 112 patients compared laxatives with placebo for the treatment of acute episodes of constipation. It was unclear if the patients had IBS. Stimulant laxatives 10mg bisacodyl (Kienzle-Horn 2006) or 7mg sodium picosulphate (Wulkow 2007) or placebo was given once-a-day for three days.

1. Global symptoms

Global symptoms (pain, bloating and bowel habit) were not reported.

2. Individual symptoms

a) Bowel habits

i. Number of patients with improvement in bowel habit assessed by investigators

The investigators assessed the improvement in bowel habit, based on diary recordings of the patients. Overall, the relative risk was 1.34 (95%CI 1.02, 1.76) (Figure 1), i.e. statistically significant difference between laxative and placebo (p=0.04). This corresponded to a number needed to treat (NNT) of 6 (95%CI 3, 50) for a control group risk of 52 to 61%. There was no heterogeneity.

Figure 1:

Review: IBS 31 August 2006 (Version 03)
Comparison: 11 Laxatives vs Placebo (simple constipation) - acute treatment
Outcome: 01 Improvement of bowel symptoms

Study or sub-category	Laxative n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 stimulant laxatives (bisacodyl, sodium picosulphate 3 days)					
Kienzle-Horn 2006	19/28	14/27		45.18	1.31 [0.84, 2.04]
Wulkow et al 2007	24/29	17/28		54.82	1.36 [0.97, 1.92]
Subtotal (95% CI)	57	55		100.00	1.34 [1.02, 1.76]
Total events: 43 (Laxative), 31 (Placebo)					
Test for heterogeneity: Chi² = 0.02, df = 1 (P = 0.89), I² = 0%					
Test for overall effect: Z = 2.09 (P = 0.04)					
02 Stool consistency (Sodium picosulphate)					
Wulkow et al 2007	25/29	16/28		100.00	1.51 [1.06, 2.15]
Subtotal (95% CI)	29	28		100.00	1.51 [1.06, 2.15]
Total events: 25 (Laxative), 16 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.29 (P = 0.02)					
04 Straining never/rarely (Sodium picosulphate)					
Wulkow et al 2007	24/29	14/28		100.00	1.66 [1.10, 2.48]
Subtotal (95% CI)	29	28		100.00	1.66 [1.10, 2.48]
Total events: 24 (Laxative), 14 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.43 (P = 0.01)					

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours Laxative

ii. Stool score - consistency

The consistency of stool was measured on the scale of 1 to 5, where 5=hard, 4=moderately hard, 3=well-formed, 2=soft, 1=liquid (Kienzle-Horn 2006) and on a 4 point scale where 4=hard, 3=well-formed, 2=pasty,1=liquid (Wulkow 2007). The Wulkow study reported the number of patients with soft and/or well formed stools. The relative risk was 1.51 (95%CI 1.06, 2.15) i.e. statistically significant difference between laxative and placebo (p=0.02) favouring laxative (Figure 1).

Kienzle-Horn (2006) reported baseline mean scores which were 5.0 for each group, so any decrease in score constituted an improvement. The study did not report the standard deviation for the placebo group, but gave the difference in change score between the two groups and the 95%CI (Figure 2). This was -1.4 (95%CI -2.0, -0.76), for a placebo group score of 4.2, i.e. a statistically significant difference between groups, such that the bisacodyl group had a value between soft and well-formed.

Figure 2:

Review: IBS 31 August 2006
Comparison: 11 Laxatives vs Placebo - acute treatment
Outcome: 03 Stool score (consistency) - liq=1, normal=3, 5=hard; lower=good

Study or sub-category	Laxative N	Control N	mean difference (SE)	mean difference (fixed) 95% CI	Weight %	mean difference (fixed) 95% CI
01 stimulant laxatives (bisacodyl 3 days)						
Kienzle-Horn 2006	27	27	-1.4000 (0.3265)		100.00	-1.40 [-2.04, -0.76]
Subtotal (95% CI)	27	27			100.00	-1.40 [-2.04, -0.76]
Test for heterogeneity: not applicable						
Test for overall effect: Z = 4.29 (P < 0.0001)						
Total (95% CI)	27	27			100.00	-1.40 [-2.04, -0.76]
Test for heterogeneity: not applicable						
Test for overall effect: Z = 4.29 (P < 0.0001)						

-4 -2 0 2 4
Favours laxative   Favours control

iii. Stool frequency

The stool frequency per day was statistically significantly higher for the bisacodyl group: mean difference 0.85 (95%CI 0.24, 1.46) for a placebo group mean of 0.95 stools/day (Figure 3).

Figure 3:

Review: IBS 31 August 2006
Comparison: 11 Laxatives vs Placebo - acute treatment
Outcome: 02 Stool score - frequency (Low=worse)

Study or sub-category	N	Laxative Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 stimulant laxatives (bisacodyl 3 days)							
Kienzle-Horn 2006	27	1.80(1.50)	27	0.95(0.60)		100.00	0.85 [0.24, 1.46]
Subtotal (95% CI)	27		27			100.00	0.85 [0.24, 1.46]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.73 (P = 0.006)							
Total (95% CI)	27		27			100.00	0.85 [0.24, 1.46]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.73 (P = 0.006)							

-4   -2   0   2   4
Favours Placebo   Favours Laxative

3. Adverse effects

The confidence interval was too wide to decide if there was a difference in the number of patients reporting adverse effects that could have been drug related (Figure 4).

Figure 4:

Review: IBS 31 August 2006
Comparison: 11 Laxatives vs Placebo - acute treatment
Outcome: 04 No of patients with adverse effects (drug related)

Study or sub-category	Laxatives n/N	Placebo n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 stimulant laxatives (bisacodyl 3 days)					
Kienzle-Horn 2006	4/28	5/27		100.00	0.73 [0.17, 3.08]
Subtotal (95% CI)	28	27		100.00	0.73 [0.17, 3.08]
Total events: 4 (Laxatives), 5 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.42 (P = 0.67)					
Total (95% CI)	28	27		100.00	0.73 [0.17, 3.08]
Total events: 4 (Laxatives), 5 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.42 (P = 0.67)					

0.1 0.2 0.5 1 2 5 10
Favours Laxatives   Favours Placebo

II. Laxatives for maintenance treatment

A. Laxatives versus Placebo

There were two studies included in the analysis that compared laxatives with placebo in patients with constipation (Corazziari 1996 and Corazziari 2000). Both studies gave the patients an isosmotic PEG electrolyte balanced solution (PMF-100) containing 14.6g PEG 4000, twice a day. However, in Corazziari (2000), all patients received 4 weeks of PMF-100 initially, with responders (more than 3 bowel movements per week) then randomised to PEG or placebo for a further 20 weeks. Thus the populations were different in the two trials, and Corazziari (2000) was regarded as an investigation of the effects of stopping the laxative. Their results are therefore reported separately. Both studies had patients who were outpatients in secondary care. The GDG considered it likely that both studies had at least some patients with IBS.

In both trials, patients were allowed other laxatives when they had no bowel movements for at least 5 consecutive days, and they were allowed to adjust the intervention dose downwards (but not upwards above 2 sachets per day).

Where outcomes were measured at different times during the study, we took the end-study results unless there were significant numbers of withdrawals or problems with compliance. Therefore, for the Corazziari (2000) study we took the values at eight weeks (i.e. from the start of randomisation).

1. Global symptoms
Neither study reported global symptoms.

2. Number of patients using additional laxatives / not using additional laxatives as rescue medication
The GDG considered this to be an important outcome for this review and gave it the status of primary outcome measure. We gave both the number of patients using additional laxatives (as reported in the papers), and the number not using additional laxatives (calculated).

The comparison of PEG and placebo over 8 weeks in 48 patients (Corazziari 1996) showed a statistically significant decrease in the number of patients using other laxatives as rescue medication; Figure 5; RR 0.33 (95%CI 0.12, 0.90), although the confidence interval was wide. This corresponded to a number needed to treat of 4 (95%CI 2, 15) for a placebo group risk of 48%.

For the outcome measure, the number of patients not using rescue medication was calculated for Corazziari (1996) (Figure 6). There was a statistically significant difference between PEG and placebo, favouring the former; RR 1.61 (95%CI 1.05, 2.47), which gave an NNT of 4 (95%CI 2, 15) for a control group risk of 52%.

Figure 5: Number of patients taking rescue laxatives
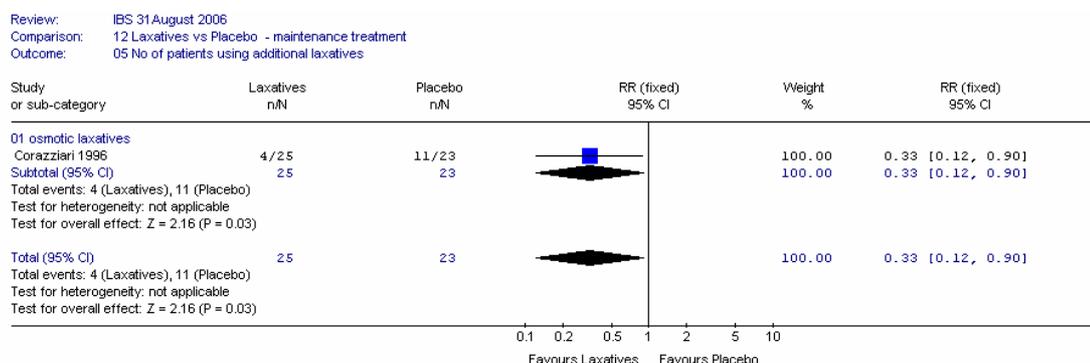
Figure 6: Number of patients not taking rescue laxatives

Review: IBS 31 August 2006
Comparison: 12 Laxatives vs Placebo - maintenance treatment
Outcome: 06 No of patients not using rescue laxatives

Study or sub-category	Laxatives n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 osmotic laxatives					
Corazziari 1996	21/25	12/23		100.00	1.61 [1.05, 2.47]
Subtotal (95% CI)	25	23		100.00	1.61 [1.05, 2.47]

Total events: 21 (Laxatives), 12 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.19 (P = 0.03)

| Total (95% CI) | 25 | 23 | | 100.00 | 1.61 [1.05, 2.47] |

Total events: 21 (Laxatives), 12 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.19 (P = 0.03)

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours Laxative

Corazziari (2000) (withdrawal of laxative following four weeks PEG electrolyte solution, in responders) reported that the use of other oral laxatives, rectal evacuants, suppositories and enemas was more frequent in the placebo group compared to the PEG group. At eight weeks the difference in number of other laxatives used per four weeks was statistically significant (Figure 6), but the confidence interval was fairly wide.

Figure 7:

Review: IBS 31 August 2006
Comparison: 13 Laxatives vs Placebo (simple constipation) - maintenance treatment, withdrawal following 4 weeks PEG
Outcome: 02 Use of other laxatives / 4 weeks at 8 weeks

Study or sub-category	Laxative N	Mean (SD)	Placebo N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 osmotic laxatives							
Corazziari 2000	32	0.70 (2.70)	33	2.20 (3.30)		100.00	-1.50 [-2.96, -0.04]
Subtotal (95% CI)	32		33			100.00	-1.50 [-2.96, -0.04]

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.01 (P = 0.04)

| Total (95% CI) | 32 | | 33 | | | 100.00 | -1.50 [-2.96, -0.04] |

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.01 (P = 0.04)

-4 -2 0 2 4
Favours Laxative   Favours Placebo

At the same time, statistically significantly more sachets of the intervention were used in the placebo group (even though the more severely constipated patients dropped out), compared with the PEG group (Figure 7).

In the PEG group, the authors reported that the use of other laxatives progressively decreased in the PEG group but increased in the placebo group.

Figure 8:

Review: IBS 31 August 2006
Comparison: 13 Laxatives vs Placebo - maintenance treatment, withdrawal following 4 weeks PEG
Outcome: 07 No of sachets of intervention used / 4 weeks at 8 weeks

Study or sub-category	Laxative N	Mean (SD)	Placebo N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 osmotic laxatives							
Corazziari 2000	32	33.00 (13.00)	33	43.00 (12.00)		100.00	-10.00 [-16.09, -3.91]
Subtotal (95% CI)	32		33			100.00	-10.00 [-16.09, -3.91]

Test for heterogeneity: not applicable
Test for overall effect: Z = 3.22 (P = 0.001)

| Total (95% CI) | 32 | | 33 | | | 100.00 | -10.00 [-16.09, -3.91] |

Test for heterogeneity: not applicable
Test for overall effect: Z = 3.22 (P = 0.001)

-100 -50 0 50 100
Favours Laxative   Favours Placebo

3. Individual symptoms

a) Pain

There was no significant difference between PEG and placebo for this outcome (Corazziari 1996) in 48 patients. The confidence intervals were fairly wide so there was some uncertainty over the results for the difference between groups in the number of patients with abdominal pain (Figure 8).

Figure 9:
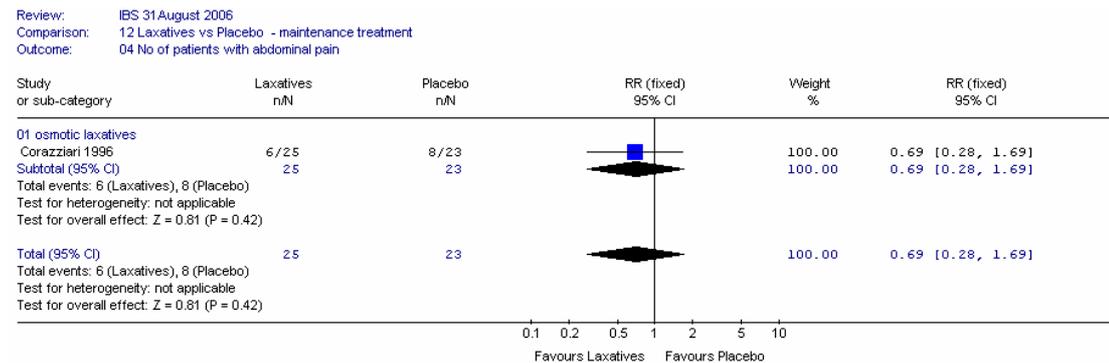
Corazziari (2000) (withdrawal of laxative following 4 weeks PEG electrolyte solution, in responders) reported that the abdominal pain score progressively decreased in the PEG group and increased in the placebo group. No data were given.

b) Bloating

The comparison of PEG and placebo (Corazziari 1996) in 48 patients showed no statistically significant difference between groups in the number of patients with bloating (Figure 9).

Figure 10:
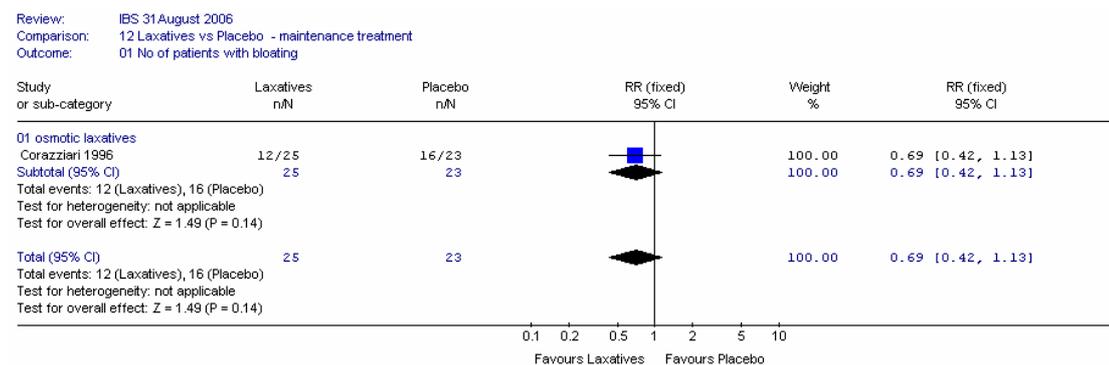
Corazziari (2000) (withdrawal of laxative following 4 weeks PEG electrolyte solution, in responders) reported that bloating was less severe in the PEG group compared to the placebo group throughout the study. At 8 weeks the difference was statistically significant: p<0.001. No other statistics were given.

c) Bowel habits

i. Stool frequency

The comparison of PEG and placebo over 8 weeks in 48 patients (Corazziari 1996) showed a statistically significant increase in stool frequency per week (figure 10) of 2.00 (95%CI 0.89, 3.11), for a placebo group value of 2.8 stools per week.

Figure 11:
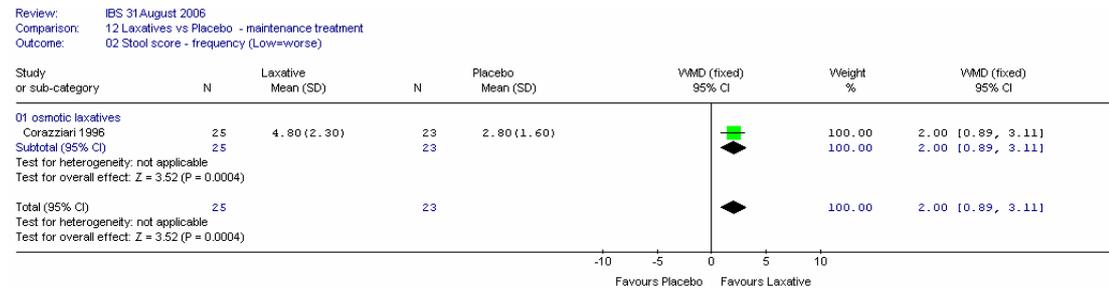
Corazziari (2000) (withdrawal of laxative following 4 weeks PEG electrolyte solution, in responders) found a statistically significant increase in stool frequency per week for the PEG group compared to the placebo group throughout the study. At 8 weeks (Figure 11) the difference was 3.13 (95%CI 1.35, 4.91) for a placebo group value of 4.39 stools per week.

Figure 12:
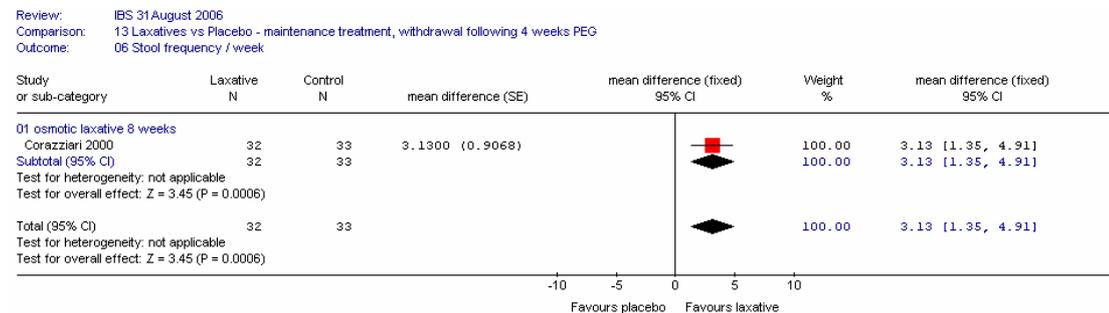
ii. Improvement in bowel habit

Corazziari (2000) (withdrawal of laxative following 4 weeks PEG electrolyte solution, in responders) found a statistically significantly greater number of patients with complete remission of constipation symptoms (more than three bowel movements per week, no use of other laxatives, no straining at defecation, no feeling of incomplete evacuation, no hard/pellety stools) for the PEG group compared to the placebo group throughout the study. At 8 weeks (Figure 12) the RR was 3.95 (95%CI 1.86, 8.42); this corresponds to an NNT of 2 (95%CI 2, 3) for a control group rate of 18%.

Figure 13:


```
Review:        IBS 31 August 2006
Comparison:    13 Laxatives vs Placebo - maintenance treatment, withdrawal following 4 weeks PEG
Outcome:       01 Improvement of bowel symptoms

Study              Laxative      Placebo         RR (fixed)        Weight        RR (fixed)
or sub-category    n/N           n/N             95% CI            %             95% CI

01 osmotic laxatives PEG 8 weeks
 Corazziari 2000    23/32         6/33                             100.00        3.95 [1.86, 8.42]
Subtotal (95% CI)   32            33                               100.00        3.95 [1.86, 8.42]
Total events: 23 (Laxative), 6 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.57 (P = 0.0004)

Total (95% CI)      32            33                               100.00        3.95 [1.86, 8.42]
Total events: 23 (Laxative), 6 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.57 (P = 0.0004)

                        0.1  0.2  0.5  1   2   5   10
                     Favours Placebo    Favours Laxative
```

iii. Number of patients withdrawing from study

Corazziari (2000) (withdrawal of laxative following 4 weeks PEG electrolyte solution, in responders) reported statistically significantly more patients withdrew from the study by 20 weeks because of non-response to treatment in the placebo group compared to the PEG group. At the end of the study the RR was 0.13 (95%CI 0.03, 0.53), i.e. statistically significantly in favour of the PEG group, although the confidence interval was very wide. This corresponded to an NNT of 3 (95%CI 2, 5) for a placebo group rate of 46%.

Figure 14:


```
Review:        IBS 31 August 2006
Comparison:    13 Laxatives vs Placebo - maintenance treatment, withdrawal following 4 weeks PEG
Outcome:       05 No of patients withdrawing from study because of non-response

Study              Laxatives     Placebo         RR (fixed)        Weight        RR (fixed)
or sub-category    n/N           n/N             95% CI            %             95% CI

01 osmotic laxatives
 Corazziari 2000    2/33          17/37                            100.00        0.13 [0.03, 0.53]
Subtotal (95% CI)   33            37                               100.00        0.13 [0.03, 0.53]
Total events: 2 (Laxatives), 17 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.86 (P = 0.004)

Total (95% CI)      33            37                               100.00        0.13 [0.03, 0.53]
Total events: 2 (Laxatives), 17 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.86 (P = 0.004)

                     0.01    0.1    1    10   100
                     Favours Laxatives    Favours Placebo
```

4. Adverse effects

In the comparison of PEG and placebo over 8 weeks in 48 patients (Corazziari 1996), there was too much uncertainty to determine if there was a difference in the number of patients reporting anorexia, headache or asthenia (Figure 15).

Figure 15:

Review: IBS 31 August 2006
Comparison: 12 Laxatives vs Placebo - maintenance treatment
Outcome: 03 No of patients with adverse effects

Study or sub-category	Laxatives n/N	Placebo n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 anorexia					
Corazziari 1996	1/25	3/23		100.00	0.28 [0.03, 2.88]
Subtotal (95% CI)	25	23		100.00	0.28 [0.03, 2.88]
Total events: 1 (Laxatives), 3 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.07 (P = 0.28)					
02 headache					
Corazziari 1996	1/25	3/23		100.00	0.28 [0.03, 2.88]
Subtotal (95% CI)	25	23		100.00	0.28 [0.03, 2.88]
Total events: 1 (Laxatives), 3 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.07 (P = 0.28)					
03 asthenia					
Corazziari 1996	1/25	1/23		100.00	0.92 [0.05, 15.56]
Subtotal (95% CI)	25	23		100.00	0.92 [0.05, 15.56]
Total events: 1 (Laxatives), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.06 (P = 0.95)					

0.1 0.2 0.5 1 2 5 10
Favours Laxatives   Favours Placebo

Corazziari (2000) reported that there were no significant differences between groups in the incidence of adverse effects.

B. Osmotic Laxative type 1 versus Osmotic laxative type 2

Three studies compared different types of osmotic laxatives: two (Attar 1999; Bouhnik 2004) compared lactulose with PEG, and one compared different types of PEG: PEG 3350 plus electrolytes versus PEG 4000 without electrolytes (Chaussade 2003). We noted that the Bouhnik (2004) was sponsored by Solvay Pharmaceuticals, manufacturers of lactulose, and Chaussade (2003) was supported by a grant from by Hoffmann La Roche, the manufacturers of PEG 3350. The Attar (1999) study was in secondary care (of which 31% were in geriatric institutions), and the Chaussade (2003) and Bouhnik (2004) studies were in primary care. The GDG thought it likely that all of these studies had some patients with IBS.

B1. Lactulose versus PEG

Attar (1999) compared PEG 3350 plus electrolytes (Movicol) versus lactulose, and Bouhnik (2004) compared PEG 4000 plus electrolytes (Forlax) versus lactulose. Both studies had a duration of four weeks. The doses of PEG differed in the two studies: the patients in Attar (1999) started with 26.24 g (2 sachets) for the first two weeks, but could change to 1 or 3 sachets for the second two weeks. Patients in Bouhnik (2004) started at a dose of 20g (2 sachets) for the first week and then this could be varied to 10 or 30g. The lactulose dose in both studies was 20g which could also be varied as above.

In Attar (1999) patients could take suppositories or microenemas for relief of constipation, apparently without restriction. However, in Bouhnik (2004) patients were asked to stop enema/suppositories 48 hours before the first stool collection.

1. Global outcomes

a) Global improvement in symptoms score

One study (Attar 1999) in 99 patients recorded a global improvement score at four weeks on a VAS of 0 to 10 (0=no change, 10=excellent); comprising pain, bloating and bowel habit. The global improvement score was statistically significantly in favour of PEG electrolyte; mean difference 2.20 (95%CI 1.05, 3.35) for a control group value of 5.20.

Figure 16

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 01 Global improvement score

Study or sub-category	N	PEG Mean (SD)	N	Lactulose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte vs lactulose							
Attar 1999	50	7.40(2.50)	49	5.20(3.30)		100.00	2.20 [1.05, 3.35]
Subtotal (95% CI)	50		49			100.00	2.20 [1.05, 3.35]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.73 (P = 0.0002)							
Total (95% CI)	50		49			100.00	2.20 [1.05, 3.35]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.73 (P = 0.0002)							

-10 -5 0 5 10
Favours lactulose   Favours PEG

2. Use of microenemas as rescue medication

Attar (1999) also reported the number of patients using microenemas as rescue medication after four weeks, and we also calculated the number of patients not using rescue medication. The study found that statistically significantly more patients used microenemas in the lactulose group than in the PEG group; RR 0.48 (95%CI 0.25, 0.95), which corresponded to an NNT of 6 (95%CI 3, 50) for a lactulose group risk of 35% (Figure 17a). The confidence interval was fairly wide.

There was a statistically significant difference, favouring PEG, for the number of patients *not* using microenemas; RR 1.27 (95%CI 1.02, 1.59). This corresponded to an NNT of 6 (95%CI 3, 50) for a lactulose risk of 65% (Figure 17b).

Figure 17a: Number of patients using microenemas

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 08 No of patients using microenemas

Study or sub-category	PEG n/N	Lactulose n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
03 PEG 3350 electrolyte 26g vs lactulose 20g					
Attar 1999	10/60	19/55		100.00	0.48 [0.25, 0.95]
Subtotal (95% CI)	60	55		100.00	0.48 [0.25, 0.95]
Total events: 10 (PEG), 19 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.12 (P = 0.03)					
Total (95% CI)	60	55		100.00	0.48 [0.25, 0.95]
Total events: 10 (PEG), 19 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.12 (P = 0.03)					

0.1 0.2 0.5 1 2 5 10
Favours PEG   Favours Lactulose

Figure 17b: Number of patients not using microenemas

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 10 No of patients not using microenemas

Study or sub-category	PEG n/N	Lactulose n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
03 PEG 3350 electrolyte 26g vs lactulose 20g					
Attar 1999	50/60	36/55		100.00	1.27 [1.02, 1.59]
Subtotal (95% CI)	60	55		100.00	1.27 [1.02, 1.59]
Total events: 50 (PEG), 36 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.12 (P = 0.03)					
Total (95% CI)	60	55		100.00	1.27 [1.02, 1.59]
Total events: 50 (PEG), 36 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.12 (P = 0.03)					

0.1   0.2   0.5   1   2   5   10
Favours lactulose   Favours PEG

3. Number of sachets of intervention used

In Attar (1999) the number of sachets of laxative used over four weeks was statistically significantly lower for the PEG group (Figure 18), but there was no difference between groups in Bouhnik (2004). This led to significant heterogeneity between studies (I^2=83%, p=0.02). It is unclear if this was an effect of dose; type of PEG; use of other laxatives, or; any other reason.

Figure 18:

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 09 Number of sachets of laxative / day after adaption period

Study or sub-category	N	PEG Mean (SD)	N	Lactulose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte 26g vs lactulose 20g							
Attar 1999	50	1.60(0.70)	49	2.10(0.70)		53.37	-0.50 [-0.78, -0.22]
Subtotal (95% CI)	50		49			53.37	-0.50 [-0.78, -0.22]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.55 (P = 0.0004)							
02 PEG 4000 electrolyte 20g vs lactulose 20g							
Bouhnik 2004	32	1.90(0.50)	33	1.90(0.70)		46.63	0.00 [-0.30, 0.30]
Subtotal (95% CI)	32		33			46.63	0.00 [-0.30, 0.30]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.00 (P = 1.00)							
Total (95% CI)	82		82			100.00	-0.27 [-0.47, -0.07]
Test for heterogeneity: Chi² = 5.89, df = 1 (P = 0.02), I² = 83.0%							
Test for overall effect: Z = 2.60 (P = 0.009)							

-4   -2   0   2   4
Favours PEG   Favours lactulose

4. Individual symptoms

a) Pain

Both studies reported the number of patients with abdominal pain at four weeks. Meta-analysis of 180 patients gave a wide confidence interval.

Figure 19:

```
Review:         IBS 31 August 2006
Comparison:     15 PEG vs Lactulose  - maintenance treatment
Outcome:        02 No of patients with abdominal pain

Study                            PEG          Lactulose          OR (fixed)        Weight        OR (fixed)
or sub-category                  n/N            n/N               95% CI             %            95% CI

01 PEG 3350 electrolyte 26g vs lactulose 20g
 Attar 1999                      6/60         10/55                                 56.60        0.50 [0.17, 1.48]
Subtotal (95% CI)                 60            55                                  56.60        0.50 [0.17, 1.48]
Total events: 6 (PEG), 10 (Lactulose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.25 (P = 0.21)

02 PEG 4000 electrolyte 20g vs lactulose 20g
 Bouhnik 2004                    6/32          9/33                                 43.40        0.62 [0.19, 1.99]
Subtotal (95% CI)                 32            33                                  43.40        0.62 [0.19, 1.99]
Total events: 6 (PEG), 9 (Lactulose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.81 (P = 0.42)

Total (95% CI)                    92            88                                 100.00        0.55 [0.25, 1.22]
Total events: 12 (PEG), 19 (Lactulose)
Test for heterogeneity: Chi² = 0.06, df = 1 (P = 0.80), I² = 0%
Test for overall effect: Z = 1.47 (P = 0.14)

                                            0.1  0.2   0.5   1    2    5    10
                                              Favours PEG    Favours Lactulose
```

One study (Attar 1999) recorded pain on a scale of 0 to 3 (severe). The difference was not statistically significant.

Figure 20:

```
Review:        IBS 31 August 2006
Comparison:    15 PEG vs Lactulose  - maintenance treatment
Outcome:       03 Pain score; scale 0-3 (severe)

Study                             PEG                 Lactulose            WMD (fixed)        Weight        WMD (fixed)
or sub-category          N      Mean (SD)      N      Mean (SD)             95% CI             %            95% CI

01 PEG 3350 electrolyte 26g vs lactulose 20g
 Attar 1999              50     0.40(0.80)     49     0.70(0.90)                               100.00       -0.30 [-0.64, 0.04]
Subtotal (95% CI)        50                    49                                             100.00       -0.30 [-0.64, 0.04]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.75 (P = 0.08)

Total (95% CI)           50                    49                                             100.00       -0.30 [-0.64, 0.04]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.75 (P = 0.08)

                                                           -4    -2    0    2    4
                                                          Favours PEG   Favours lactulose
```

b) Bloating

Both studies reported the number of patients with bloating at four weeks. Meta-analysis of 180 patients gave a fairly wide confidence interval and some heterogeneity (I^2=50%; p=0.16). This difference may be an effect of dose or type of PEG. It is also noted that Bouhnik (2004) was sponsored by the manufacturers of lactulose and Attar (1999) allowed the patients to use other laxatives *ad libitum*.

Figure 21:

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 06 No of patients with bloating

Study or sub-category	PEG n/N	Lactulose n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 PEG 3350 electrolyte 26g vs lactulose 20g					
Attar 1999	9/60	18/55		63.43	0.46 [0.22, 0.93]
Subtotal (95% CI)	60	55		63.43	0.46 [0.22, 0.93]
Total events: 9 (PEG), 18 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.15 (P = 0.03)					
02 PEG 4000 electrolyte 20g vs lactulose 20g					
Bouhnik 2004	10/32	11/33		36.57	0.94 [0.46, 1.90]
Subtotal (95% CI)	32	33		36.57	0.94 [0.46, 1.90]
Total events: 10 (PEG), 11 (Lactulose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.18 (P = 0.86)					
Total (95% CI)	92	88		100.00	0.63 [0.39, 1.04]
Total events: 19 (PEG), 29 (Lactulose)					
Test for heterogeneity: Chi² = 1.98, df = 1 (P = 0.16), I² = 49.6%					
Test for overall effect: Z = 1.80 (P = 0.07)					

0.1 0.2 0.5 1 2 5 10
Favours PEG   Favours Lactulose

c) Bowel habits

i. Stool frequency

Both studies reported the stool frequency per day at four weeks. Meta-analysis of 180 patients gave a statistically significant difference of 0.27 (95%CI 0.09, 0.45) stools/day, favouring PEG, but there was some heterogeneity (I^2=50%; p=0.16).

Figure 22:

Review: IBS 31 August 2006
Comparison: 15 PEG vs Lactulose - maintenance treatment
Outcome: 04 Stool frequency / day

Study or sub-category	N	PEG Mean (SD)	N	Lactulose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte 26g vs lactulose 20g							
Attar 1999	50	1.30(0.70)	49	0.90(0.60)		49.04	0.40 [0.14, 0.66]
Subtotal (95% CI)	50		49			49.04	0.40 [0.14, 0.66]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.05 (P = 0.002)							
02 PEG 4000 electrolyte 20g vs lactulose 20g							
Bouhnik 2004	32	1.26(0.65)	33	1.12(0.33)		50.96	0.14 [-0.11, 0.39]
Subtotal (95% CI)	32		33			50.96	0.14 [-0.11, 0.39]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.09 (P = 0.28)							
Total (95% CI)	82		82			100.00	0.27 [0.09, 0.45]
Test for heterogeneity: Chi² = 2.01, df = 1 (P = 0.16), I² = 50.2%							
Test for overall effect: Z = 2.92 (P = 0.004)							

-4 -2 0 2 4
Favours Lactulose   Favours PEG

5. Adverse effects

Both studies reported the number of patients with adverse effects at four weeks. Meta-analysis of 180 patients gave a wide confidence interval and no heterogeneity (Figure 22). One study (Attar 1999) reported the number of patients with liquid stools, but this also had a wide confidence interval (Figure 23).

Figure 23:


```
Review:        IBS 31 August 2006
Comparison:    15 PEG vs Lactulose  - maintenance treatment
Outcome:       05 No of patients with adverse effects

Study                PEG         Lactulose      OR (fixed)        Weight      OR (fixed)
or sub-category      n/N           n/N           95% CI             %          95% CI

01 overall; PEG 3350 electrolyte 26g vs lactulose 20g
 Attar 1999          4/60          3/55                            30.54     1.24 [0.26, 5.80]
Subtotal (95% CI)     60            55                             30.54     1.24 [0.26, 5.80]
Total events: 4 (PEG), 3 (Lactulose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.27 (P = 0.79)

02 overall; PEG 4000 electrolyte 20g vs lactulose 20g
 Bouhnik 2004        5/32          8/33                            69.46     0.58 [0.17, 2.01]
Subtotal (95% CI)     32            33                             69.46     0.58 [0.17, 2.01]
Total events: 5 (PEG), 8 (Lactulose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.86 (P = 0.39)

Total (95% CI)        92            88                            100.00     0.78 [0.30, 2.02]
Total events: 9 (PEG), 11 (Lactulose)
Test for heterogeneity: Chi² = 0.57, df = 1 (P = 0.45), I² = 0%
Test for overall effect: Z = 0.51 (P = 0.61)

              0.1  0.2   0.5   1    2    5   10
                 Favours PEG   Favours Lactulose
```

Figure 24:


```
Review:        IBS 31 August 2006
Comparison:    15 PEG vs Lactulose  - maintenance treatment
Outcome:       07 No of patients with specific adverse effects

Study                PEG         Lactulose      OR (fixed)        Weight      OR (fixed)
or sub-category      n/N           n/N           95% CI             %          95% CI

03 liquid stools; PEG 3350 electrolyte 26g vs lactulose 20g
 Attar 1999          6/60          5/55                           100.00     1.11 [0.32, 3.87]
Subtotal (95% CI)     60            55                            100.00     1.11 [0.32, 3.87]
Total events: 6 (PEG), 5 (Lactulose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.17 (P = 0.87)

              0.1  0.2   0.5   1    2    5   10
                 Favours PEG   Favours Lactulose
```

B2. Comparison of different PEG laxatives

One study (Chaussade 2003) compared two doses of each of two types of PEG solution for a duration of four weeks. The PEG species were PEG 4000 (Forlax) without electrolytes and PEG 3350 plus electrolytes. Doses used were the maximum and standard recommended by the manufacturers.

1. Global improvement of symptoms

The study measured the patients' global impression of efficacy on a VAS, but the results are not reported. The authors state that there was no significant difference between groups.

2. Individual symptoms

a) Abdominal pain

Pain scores at four weeks were recorded on a scale of 1 (none) to 4 (considerable).

There was no significant difference between the two types of PEG for this outcome at either dose, and there was no heterogeneity (I^2=0%).

Figure 25:
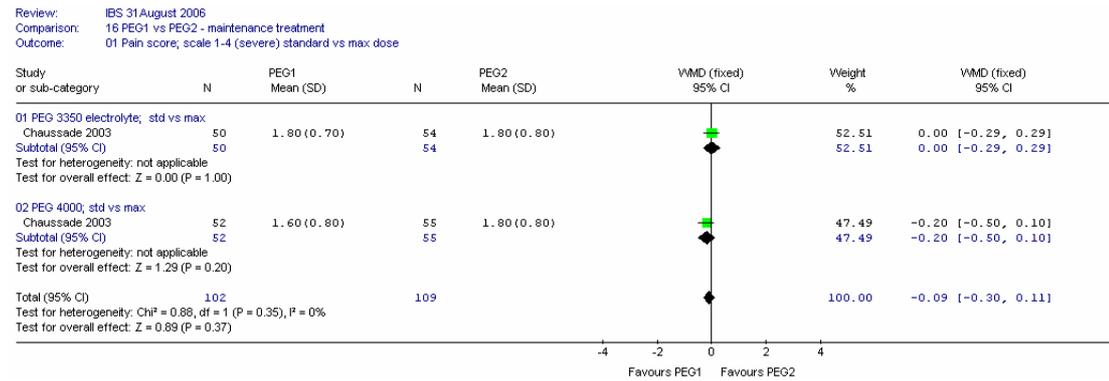
b) Bloating

Bloating scores at four weeks were recorded on a scale of 1 (none) to 4 (considerable). There was no significant difference between the two types of PEG for this outcome at either dose, and no heterogeneity (I^2=0%).

Figure 26:
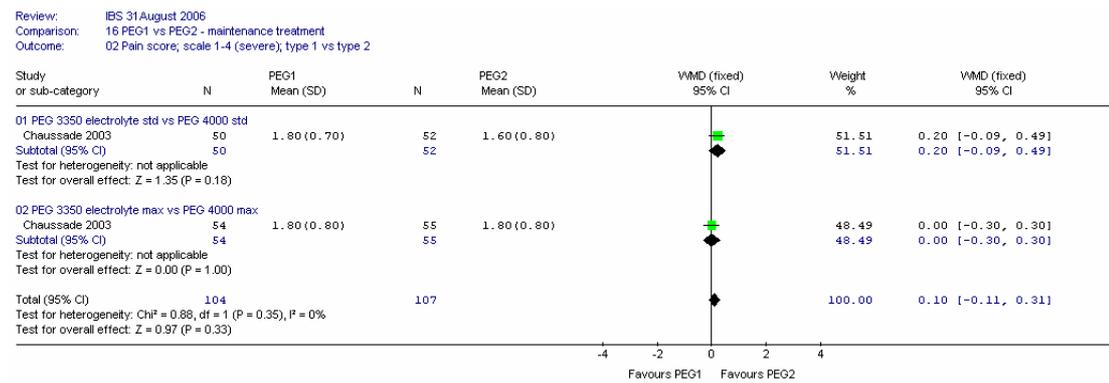
c) Bowel habits

i. Stool frequency per week

There was no significant difference between the two types of PEG for this outcome at four weeks at either dose, and no heterogeneity (I^2=0%).

Figure 27:
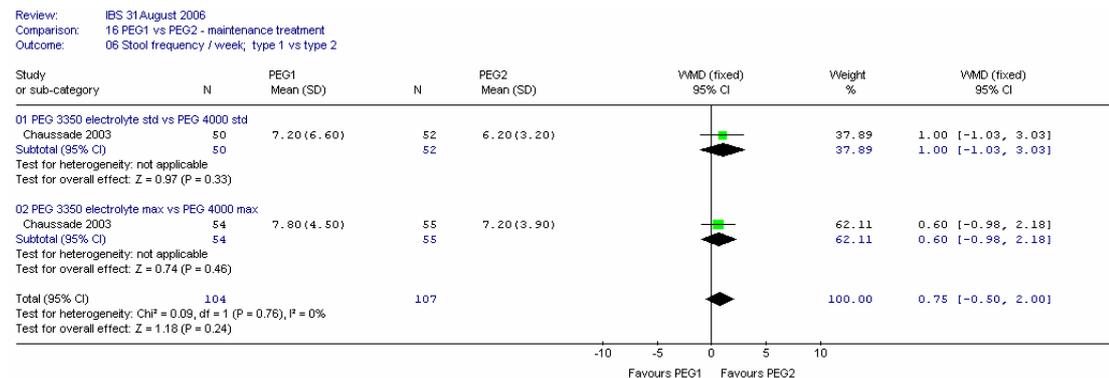
ii. Stool consistency

Consistency of stools at four weeks were recorded on a scale of 1 (liquid) to 6 (very hard). Meta-analysis of the two comparisons revealed heterogeneity (I^2=65%, p=0.09). For the standard dose, there was a statistically significant difference between the two PEG solutions, favouring PEG 4000; mean difference 0.30 (95%CI 0.01, 0.59). This was a fairly small change. There was no significant difference for the maximum dose.

Figure 28:
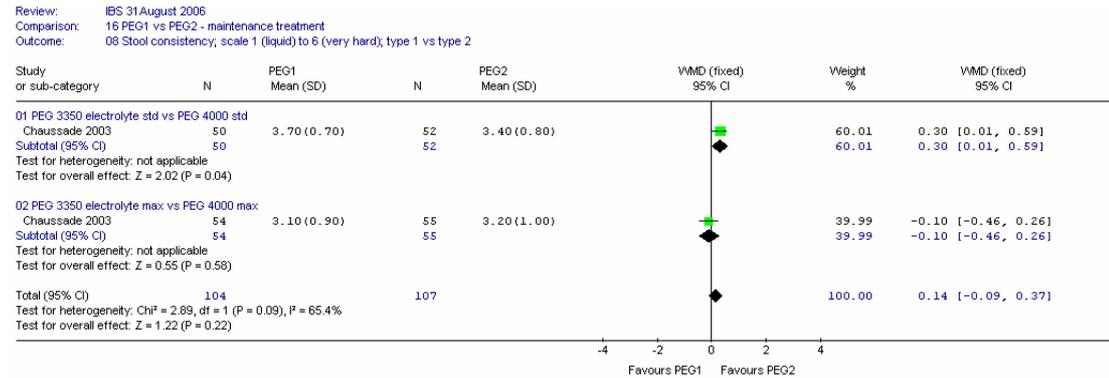
iii. Number of patients with normal stools

Meta-analysis showed no significant difference at four weeks between types of PEG.

Figure 29:
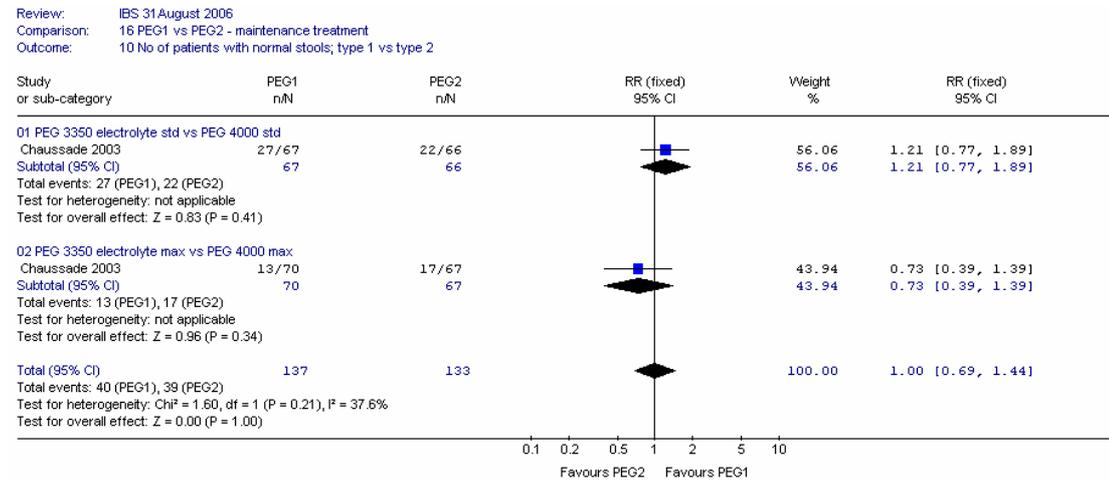
3. Quality of life

The study measured quality of life on a 100mm VAS at four weeks. Meta-analysis showed no significant difference between types of PEG and no heterogeneity (I^2=0%, p=0.93).

Figure 30:
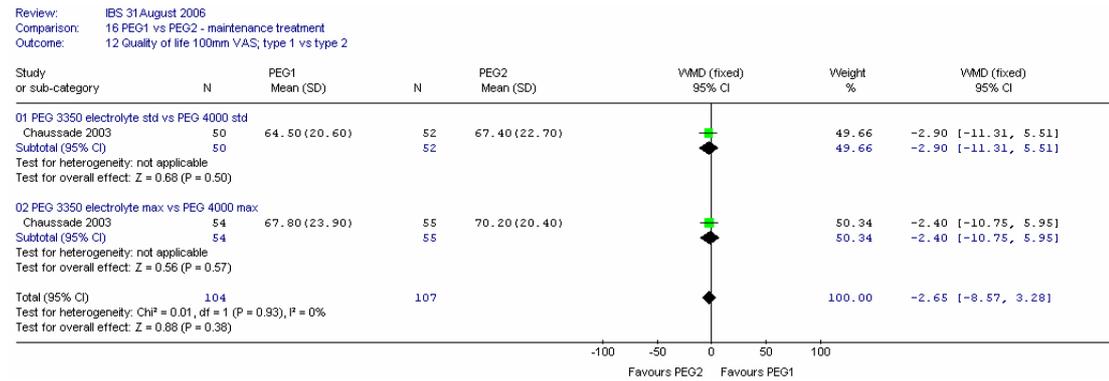
4. Adverse effects

For overall adverse effects at four weeks, the majority of which were gastrointestinal, meta-analysis showed no significant difference between types of PEG and no heterogeneity (I^2=0%, p=0.55).

Figure 31:
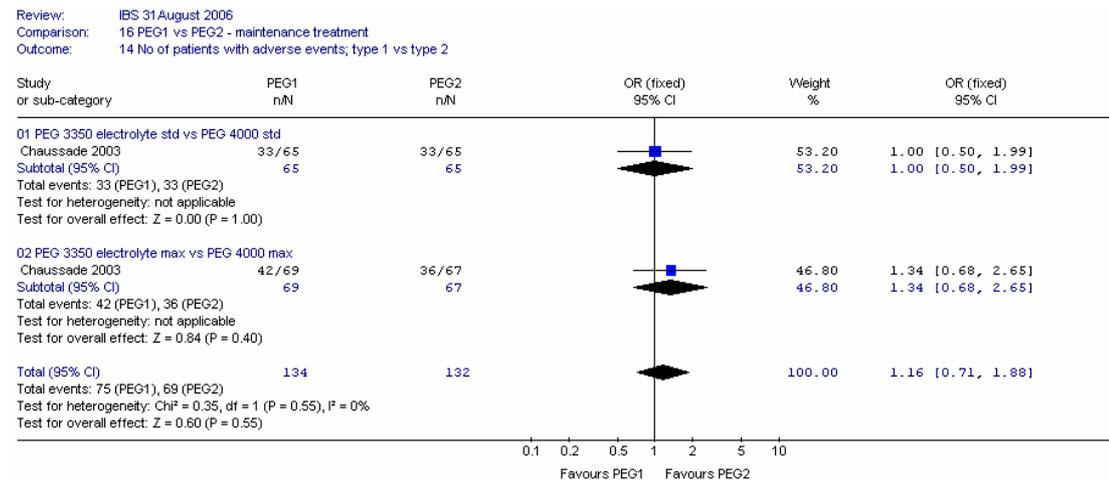
For the specific adverse effect of diarrhoea, there was no significant difference between types of PEG and no heterogeneity (I^2=0%, p=0.71).

Figure 32:
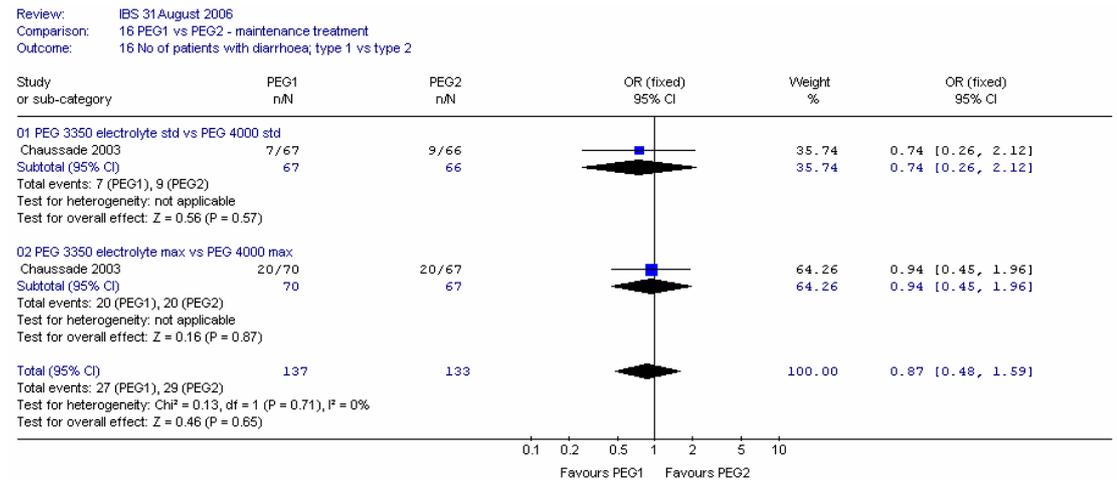
B3. Comparison of different doses of PEG laxatives

One study (Chaussade 2003) compared two doses of each of two types of PEG solution for a duration of four weeks. The PEG species were PEG 4000 (Forlax) without electrolytes and PEG 3350 plus electrolytes. Doses used were the maximum and standard recommended by the manufacturers.

1. Global improvement of symptoms

The study measured the patients' global impression of efficacy at four weeks on a VAS, but the results are not reported. The authors state that there was no significant difference between groups.

2. Individual symptoms

a) Abdominal pain

Pain scores at four weeks were recorded on a scale of 1 (none) to 4 (considerable). Meta-analysis showed no significant difference between the two doses for this outcome, and no heterogeneity (I^2=0%).

Figure 33:

b) Bloating

Bloating scores at four weeks were recorded on a scale of 1 (none) to 4 (considerable). There was no significant difference between the two doses for this outcome, and no heterogeneity (I^2=0%).

Figure 34:


```
Review:       IBS 31 August 2006
Comparison:   16 PEG1 vs PEG2 - maintenance treatment
Outcome:      03 Bloating score; scale 1-4 (severe); std vs max
```

Study or sub-category	N	std dose Mean (SD)	N	max dose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max							
Chaussade 2003	50	2.10(0.70)	54	2.10(0.80)		53.07	0.00 [-0.29, 0.29]
Subtotal (95% CI)	50		54			53.07	0.00 [-0.29, 0.29]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.00 (P = 1.00)							
02 PEG 4000; std vs max							
Chaussade 2003	52	1.90(0.90)	55	2.00(0.70)		46.93	-0.10 [-0.41, 0.21]
Subtotal (95% CI)	52		55			46.93	-0.10 [-0.41, 0.21]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.64 (P = 0.52)							
Total (95% CI)	102		109			100.00	-0.05 [-0.26, 0.16]
Test for heterogeneity: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%							
Test for overall effect: Z = 0.44 (P = 0.66)							

-4 -2 0 2 4
Favours std dose   Favours max dose

c) Bowel habits

i. Stool frequency

There was no significant difference between the two doses for this outcome at four weeks, and no heterogeneity (I^2=0%).

Figure 35:


```
Review:       IBS 31 August 2006
Comparison:   16 PEG1 vs PEG2 - maintenance treatment
Outcome:      05 Stool frequency / week; std vs max
```

Study or sub-category	N	std dose Mean (SD)	N	max dose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max							
Chaussade 2003	50	7.20(6.60)	54	7.80(4.50)		27.53	-0.60 [-2.79, 1.59]
Subtotal (95% CI)	50		54			27.53	-0.60 [-2.79, 1.59]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.54 (P = 0.59)							
02 PEG 4000; std vs max							
Chaussade 2003	52	6.20(3.20)	55	7.20(3.90)		72.47	-1.00 [-2.35, 0.35]
Subtotal (95% CI)	52		55			72.47	-1.00 [-2.35, 0.35]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.45 (P = 0.15)							
Total (95% CI)	102		109			100.00	-0.89 [-2.04, 0.26]
Test for heterogeneity: Chi² = 0.09, df = 1 (P = 0.76), I² = 0%							
Test for overall effect: Z = 1.52 (P = 0.13)							

-10 -5 0 5 10
Favours max dose   Favours std dose

ii. Stool consistency

Consistency of stools at four weeks was recorded on a scale of 1 (liquid) to 6 (very hard). Meta-analysis of the two comparisons revealed some heterogeneity (I^2=65%, p=0.09). For the PEG 3350 electrolyte dose, there was a statistically significant difference between the two doses, favouring the maximum dose; mean difference 0.60 (95%CI 0.29, 0.91). There was no significant difference for the PEG 4000.

Figure 36:

Review: IBS 31 August 2006
Comparison: 16 PEG1 vs PEG2 - maintenance treatment
Outcome: 07 Stool consistency; scale 1 (liquid) to 6 (very hard); std vs max

Study or sub-category	std dose N	Mean (SD)	max dose N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max							
Chaussade 2003	50	3.70(0.70)	54	3.10(0.90)		55.15	0.60 [0.29, 0.91]
Subtotal (95% CI)	50		54			55.15	0.60 [0.29, 0.91]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.81 (P = 0.0001)							
02 PEG 4000; std vs max							
Chaussade 2003	52	3.40(0.80)	55	3.20(1.00)		44.85	0.20 [-0.14, 0.54]
Subtotal (95% CI)	52		55			44.85	0.20 [-0.14, 0.54]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.15 (P = 0.25)							
Total (95% CI)	102		109			100.00	0.42 [0.19, 0.65]
Test for heterogeneity: Chi² = 2.89, df = 1 (P = 0.09), I² = 65.4%							
Test for overall effect: Z = 3.60 (P = 0.0003)							

-4 -2 0 2 4
Favours std dose   Favours max dose

iii. Number of patients with normal stools

Meta-analysis of the two comparisons reveals statistically significantly more patients with normal stools at four weeks for the standard dose groups. The RR was 1.68 (95%CI 1.14, 2.48), which corresponded to an NNH of 7 (95%CI 4, 25) for the higher dose rate of 19% or 25%.

Figure 37:

Review: IBS 31 August 2006
Comparison: 16 PEG1 vs PEG2 - maintenance treatment
Outcome: 09 No of patients with normal stools; std vs max

Study or sub-category	std dose n/N	max dose n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max					
Chaussade 2003	27/67	13/70		42.98	2.17 [1.23, 3.84]
Subtotal (95% CI)	67	70		42.98	2.17 [1.23, 3.84]
Total events: 27 (std dose), 13 (max dose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.66 (P = 0.008)					
02 PEG 4000; std vs max					
Chaussade 2003	22/66	17/67		57.02	1.31 [0.77, 2.24]
Subtotal (95% CI)	66	67		57.02	1.31 [0.77, 2.24]
Total events: 22 (std dose), 17 (max dose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.00 (P = 0.32)					
Total (95% CI)	133	137		100.00	1.68 [1.14, 2.48]
Total events: 49 (std dose), 30 (max dose)					
Test for heterogeneity: Chi² = 1.59, df = 1 (P = 0.21), I² = 37.0%					
Test for overall effect: Z = 2.63 (P = 0.009)					

0.1 0.2 0.5 1 2 5 10
Favours max dose   Favours std dose

3. Quality of life

The study measured quality of life at four weeks on a 100mm VAS. Meta-analysis showed no significant difference between doses and no heterogeneity (I^2=0%, p=0.93).

Figure 38:

Review: IBS 31 August 2006
Comparison: 16 PEG1 vs PEG2 - maintenance treatment
Outcome: 11 Quality of life 100mm VAS; std vs max

Study or sub-category	N	std dose Mean (SD)	N	max dose Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max							
Chaussade 2003	50	64.50(20.60)	54	67.80(23.90)		47.83	-3.30 [-11.86, 5.26]
Subtotal (95% CI)	50		54			47.83	-3.30 [-11.86, 5.26]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.76 (P = 0.45)							
02 PEG 4000; std vs max							
Chaussade 2003	52	67.40(22.70)	55	70.20(20.40)		52.17	-2.80 [-10.99, 5.39]
Subtotal (95% CI)	52		55			52.17	-2.80 [-10.99, 5.39]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.67 (P = 0.50)							
Total (95% CI)	102		109			100.00	-3.04 [-8.96, 2.88]
Test for heterogeneity: Chi² = 0.01, df = 1 (P = 0.93), I² = 0%							
Test for overall effect: Z = 1.01 (P = 0.31)							

-100 -50 0 50 100
Favours max dose   Favours std dose

4. Adverse effects

For overall adverse effects at four weeks, the majority of which were gastrointestinal, meta-analysis showed no significant difference between doses and no heterogeneity (I^2=0%, p=0.55).

Figure 39:

Review: IBS 31 August 2006
Comparison: 16 PEG1 vs PEG2 - maintenance treatment
Outcome: 13 No of patients with adverse events; std vs max

Study or sub-category	std dose n/N	max dose n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 PEG 3350 electrolyte; std vs max					
Chaussade 2003	33/65	42/69		53.47	0.83 [0.61, 1.13]
Subtotal (95% CI)	65	69		53.47	0.83 [0.61, 1.13]
Total events: 33 (std dose), 42 (max dose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.17 (P = 0.24)					
02 PEG 4000; std vs max					
Chaussade 2003	33/65	36/67		46.53	0.94 [0.68, 1.31]
Subtotal (95% CI)	65	67		46.53	0.94 [0.68, 1.31]
Total events: 33 (std dose), 36 (max dose)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.34 (P = 0.73)					
Total (95% CI)	130	136		100.00	0.89 [0.71, 1.11]
Total events: 66 (std dose), 78 (max dose)					
Test for heterogeneity: Chi² = 0.30, df = 1 (P = 0.58), I² = 0%					
Test for overall effect: Z = 1.07 (P = 0.29)					

0.1 0.2 0.5 1 2 5 10
Favours std dose   Favours max dose

For the specific adverse effect of diarrhoea, there was a statistically significant difference, favouring the standard dose, and no heterogeneity (I^2=0%, p=0.68). The RR was 0.41 (95%CI 0.24, 0.70); this corresponded to an NNT of 6 (95%CI 4, 13) for the higher dose rate of 30%.

Figure 40:


```
Review:        IBS 31 August 2006
Comparison:    16 PEG1 vs PEG2 - maintenance treatment
Outcome:       15 No of patients with diarrhoea; std vs max

Study                          std dose      max dose         RR (fixed)        Weight        RR (fixed)
or sub-category                 n/N            n/N             95% CI             %            95% CI

01 PEG 3350 electrolyte;  std vs max
 Chaussade 2003                 7/67          20/70                               49.64         0.37 [0.17, 0.81]
Subtotal (95% CI)               67            70                                  49.64         0.37 [0.17, 0.81]
Total events: 7 (std dose), 20 (max dose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.49 (P = 0.01)

02 PEG 4000; std vs max
 Chaussade 2003                 9/66          20/67                               50.36         0.46 [0.22, 0.93]
Subtotal (95% CI)               66            67                                  50.36         0.46 [0.22, 0.93]
Total events: 9 (std dose), 20 (max dose)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.16 (P = 0.03)

Total (95% CI)                  133           137                                 100.00        0.41 [0.24, 0.70]
Total events: 16 (std dose), 40 (max dose)
Test for heterogeneity: Chi² = 0.17, df = 1 (P = 0.68), I² = 0%
Test for overall effect: Z = 3.29 (P = 0.0010)

              0.1   0.2    0.5    1    2    5    10
              Favours std dose   Favours max dose
```

B4. Stimulant Laxative Type 1 versus Stimulant Laxative Type 2

One study compared two stimulant laxatives, bisacodyl versus sodium picosulphate (Kienzle-Horn 2007). Patients were treated daily for 4 weeks with 5 to10mg of either bisacodyl or sodium picosulphate. The primary outcome was the change in bowel habit recorded as the mean number of bowel movements per day and stool consistency measured on a 5 point scale where 5=hard, 4= moderately hard,3= well formed,2=soft, 1=liquid. Secondary outcomes included straining scored on a 4 point scale with 4 = severe and 0=absent. There was no statistically significant difference between the two laxatives for the number of bowel movements per day, WMD: -0,05 (95%CI -0.18, 0.08), and similarly for the stool consistency and straining score. Both were equally effective in treating constipation.

Figure 41:


```
Review:        IBS 31 August 2006 (Version 03)
Comparison:    22 Stimulant Laxative 1 vs Stimulant Laxative 2 (Maintenance Treatment) Simple constipation
Outcome:       01 Bisacodyl vs sodium picosulphate

Study                    Bisacodyl              sodium picosulphate      WMD (fixed)      Weight      WMD (fixed)
or sub-category      N    Mean (SD)          N    Mean (SD)             95% CI            %          95% CI

01 Number of Bowel movements per day (end of treatment)
 Kienzle -Horn 2007   67   1.06 (0.35)       72   1.11 (0.45)                            100.00      -0.05 [-0.18, 0.08]
Subtotal (95% CI)     67                     72                                          100.00      -0.05 [-0.18, 0.08]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.73 (P = 0.46)

02 Stool consistency
 Kienzle -Horn 2007   67   2.43 (0.54)       72   2.51 (0.50)                            100.00      -0.08 [-0.25, 0.09]
Subtotal (95% CI)     67                     72                                          100.00      -0.08 [-0.25, 0.09]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.90 (P = 0.37)

03 Straining score
 Kienzle -Horn 2007   67   1.20 (0.62)       72   1.17 (0.64)                            100.00       0.03 [-0.18, 0.24]
Subtotal (95% CI)     67                     72                                          100.00       0.03 [-0.18, 0.24]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.28 (P = 0.78)

              -4    -2    0    2    4
              Favours treatment   Favours control
```

C. Laxative versus fibre

Three studies compared a laxative with fibre: two compared an osmotic laxative (lactulose) with fibre (ispaghula), (Quah 2006; Rouse 1991) and one compared usual laxatives (mainly lactulose) with fibre (ispaghula) (Dettmar 1998). Quah (2006) had a crossover design, with a

washout period of 1 week, and this study was treated separately. Quah (2006) and Rouse (1991) compared respectively: 20 ml lactulose with 3.5g ispaghula husk, and 30 ml lactulose with 7g ispaghula husk. Dettmar (1998) did not record the dose of lactulose (other laxatives), but 7g ispaghula husk was given. Dettmar (1998) also reported results for the lactulose subgroup of 'other laxatives'. Rouse (1991) and Dettmar (1998) were in primary care and Quah (2006) in secondary care. The authors of Dettmar (1998) were from Reckitt and Colman, manufacturers of Fybogel. The GDG considered it unlikely that the patients in Dettmar (1998) and Rouse (1991) had IBS. It was unclear if the patients in Quah (2006) had IBS.

1. Global symptoms

Two studies reported an outcome of global symptoms (Rouse 1991; Dettmar 1998).

a) Global improvement of symptoms

One study (Rouse 1991) in 112 patients showed little difference in global improvement of symptoms at four weeks between patients given 30ml lactulose and 7g ispaghula husk (figure 40).

Figure 42:
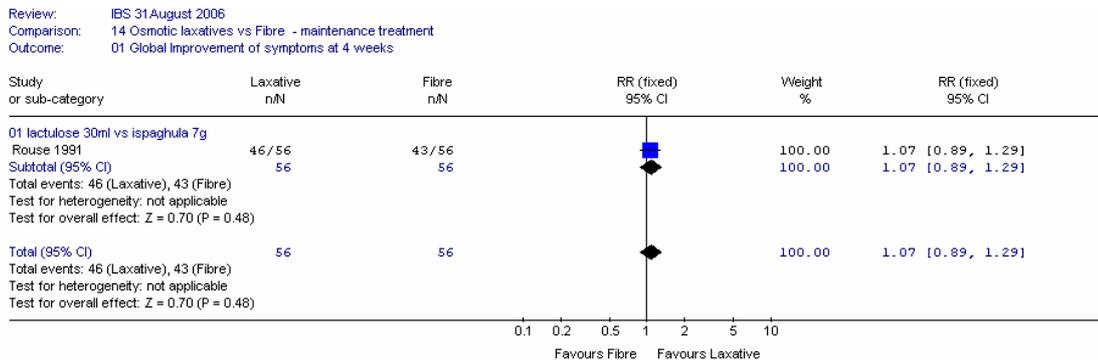
a) Global effectiveness

Dettmar (1998), in 315 patients, asked the patients to rate the effectiveness at four weeks of treatment with ispaghula 7g and other laxatives, mainly lactulose. Statistically significantly more patients given ispaghula rated the effectiveness as excellent, good or satisfactory. RR (all other laxatives versus fibre) was 0.87 (95%CI 0.80, 0.96) and for the subgroup with lactulose the RR was 0.90 (95%CI 0.85, 0.96). It was noted that authors of Dettmar (1998) were from the manufacturers of ispaghula (Reckitt and Colman).

Figure 43:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 02 Global effectiveness at 4 weeks

Study or sub-category	Laxative n/N	Fibre n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Subgroup of lactulose vs ispaghula 7g					
Dettmar 1998	77/91	217/224		100.00	0.87 [0.80, 0.96]
Subtotal (95% CI)	91	224		100.00	0.87 [0.80, 0.96]
Total events: 77 (Laxative), 217 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.92 (P = 0.003)					
02 Other laxatives mainly lactulose vs ispaghula 7g					
Dettmar 1998	149/170	217/224		100.00	0.90 [0.85, 0.96]
Subtotal (95% CI)	170	224		100.00	0.90 [0.85, 0.96]
Total events: 149 (Laxative), 217 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.21 (P = 0.001)					

0.1 0.2 0.5 1 2 5 10
Favours Fibre   Favours Laxative

When these two studies were combined in a meta-analysis, the RR was 0.92 (0.85, 1.00) and there was a significant heterogeneity I^2 = 73.5%, p=0.05. It is unclear what caused this, but the overall effect was small.

Figure 44:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 10 Global Improvement at 4 weeks

Study or sub-category	Laxative n/N	Fibre n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 lactulose 30ml vs ispaghula 7g					
Rouse 1991	46/56	43/56		25.54	1.07 [0.89, 1.29]
Subtotal (95% CI)	56	56		25.54	1.07 [0.89, 1.29]
Total events: 46 (Laxative), 43 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.70 (P = 0.48)					
02 Subgroup of lactulose vs ispaghula 7g					
Dettmar 1998	77/91	217/224		74.46	0.87 [0.80, 0.96]
Subtotal (95% CI)	91	224		74.46	0.87 [0.80, 0.96]
Total events: 77 (Laxative), 217 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.92 (P = 0.003)					
Total (95% CI)	147	280		100.00	0.92 [0.85, 1.00]
Total events: 123 (Laxative), 260 (Fibre)					
Test for heterogeneity: Chi² = 3.78, df = 1 (P = 0.05), I² = 73.5%					
Test for overall effect: Z = 1.86 (P = 0.06)					

0.1 0.2 0.5 1 2 5 10
Favours Fibre   Favours Laxative

2. Individual symptoms

a) Pain

Two studies (Rouse 1991; Quah 2006) in 93 and 78 patients respectively, recorded the number of patients with abdominal pain at four weeks. We did not combine these studies because one was a crossover study and the other parallel. We did not draw conclusions for the crossover study because the confidence interval was too wide; there was also only 1 week washout for this study. The confidence interval was fairly wide for Rouse (1991), but there was no significant difference between lactulose and ispaghula.

Figure 45:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 07 No of patients with abdominal pain

Study or sub-category	Laxatives n/N	Fibre n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	10/39	5/39		100.00	2.00 [0.75, 5.32]
Subtotal (95% CI)	39	39		100.00	2.00 [0.75, 5.32]
Total events: 10 (Laxatives), 5 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.39 (P = 0.16)					
02 lactulose 30ml vs ispaghula 7g					
Rouse 1991	14/48	14/45		100.00	0.94 [0.50, 1.74]
Subtotal (95% CI)	48	45		100.00	0.94 [0.50, 1.74]
Total events: 14 (Laxatives), 14 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.20 (P = 0.84)					

0.1 0.2 0.5 1 2 5 10
Favours Laxatives   Favours Fibre

b) Bloating

Two studies recorded the number of patients with bloating at four weeks: Quah (2006) in 76 patients, and; Dettmar (1998) in 394 patients. There was little difference in the numbers with bloating, although the confidence intervals were fairly wide.

Figure 46:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 06 No of patients with bloating

Study or sub-category	Laxatives n/N	Fibre n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	11/39	11/39		100.00	1.00 [0.49, 2.03]
Subtotal (95% CI)	39	39		100.00	1.00 [0.49, 2.03]
Total events: 11 (Laxatives), 11 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.00 (P = 1.00)					
02 overall; Other laxatives mainly lactulose vs ispaghula 7g					
Dettmar 1998	24/170	35/224		100.00	0.90 [0.56, 1.46]
Subtotal (95% CI)	170	224		100.00	0.90 [0.56, 1.46]
Total events: 24 (Laxatives), 35 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.41 (P = 0.68)					
03 Subgroup of lactulose vs ispaghula 7g					
Dettmar 1998	12/91	35/224		100.00	0.84 [0.46, 1.55]
Subtotal (95% CI)	91	224		100.00	0.84 [0.46, 1.55]
Total events: 12 (Laxatives), 35 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.55 (P = 0.58)					

0.1 0.2 0.5 1 2 5 10
Favours Laxatives   Favours Fibre

c) Bowel habits

i. Improvement in bowel score

One crossover study (Quah 2006) with 78 patients recorded improvement in bowel score at four weeks compared with baseline on a scale of 0 (no effect) to 10 (excellent) (Figure 45). There was a statistically significantly greater improvement with lactulose compared to ispaghula; mean difference 1.40 (95%CI 0.19, 2.61). It was noted that this crossover study had a washout period of only 1 week, so the results were treated with caution.

Figure 47:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 03 Improvement in bowel score

Study or sub-category	N	Laxative Mean (SD)	N	Fibre Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover (0-10 scale)							
Quah 2006	39	6.20(2.39)	39	4.80(3.03)		100.00	1.40 [0.19, 2.61]
Subtotal (95% CI)	39		39			100.00	1.40 [0.19, 2.61]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.27 (P = 0.02)							
Total (95% CI)	39		39			100.00	1.40 [0.19, 2.61]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.27 (P = 0.02)							

-10   -5   0   5   10
Favours Fibre   Favours Laxative

ii. Stool frequency

One study reported the stool frequency at four weeks (Quah 2006). There was a non-significant difference between lactulose and ispaghula, favouring the former; WMD 1.80 (95%CI -0.12, 3.72). It was noted that this crossover study had a washout period of only 1 week, so the results were treated with caution.

Figure 48:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 04 Stool frequency / week

Study or sub-category	N	Laxative Mean (SD)	N	Fibre Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover							
Quah 2006	39	7.30(5.10)	39	5.50(3.35)		100.00	1.80 [-0.12, 3.72]
Subtotal (95% CI)	39		39			100.00	1.80 [-0.12, 3.72]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.84 (P = 0.07)							
Total (95% CI)	39		39			100.00	1.80 [-0.12, 3.72]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.84 (P = 0.07)							

-10   -5   0   5   10
Favours Fibre   Favours Laxative

iii. Stool consistency

One study reported the stool consistency at four weeks on a scale of 0 (no bowel movement) to 3 (comfortable and solid) to 5 (loose) (Quah 2006). There was a borderline significant difference of 0.50 (95%CI 0.00, 1.00; p=0.05) between lactulose and ispaghula, favouring the former. However, since the normal rating is 3 and the fibre group is closer to this value (2.9) it could be argued that fibre is more favourable. It was noted that this crossover study had a washout period of only one week, so the results were treated with caution.

Figure 49:

Review: IBS 31 August 2006
Comparison: 14 Osmotic laxatives vs Fibre - maintenance treatment
Outcome: 05 Stool consistency / week (0=no BM, 1=hard, 3=comfortable, 5=loose)

Study or sub-category	N	Laxative Mean (SD)	N	Fibre Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover							
Quah 2006	39	3.40(0.96)	39	2.90(1.27)		100.00	0.50 [0.00, 1.00]
Subtotal (95% CI)	39		39			100.00	0.50 [0.00, 1.00]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.96 (P = 0.05)							
Total (95% CI)	39		39			100.00	0.50 [0.00, 1.00]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.96 (P = 0.05)							

-4   -2   0   2   4
Favours Fibre   Favours Laxative

3. Adverse effects

Two studies reported the number of patients with adverse effects at four weeks. In all cases there were wide confidence intervals (Figure 48).

Figure 50:

Study or sub-category	Laxatives n/N	Fibre n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 overall; Other laxatives mainly lactulose vs ispaghula 7g					
Dettmar 1998	12/170	10/224		100.00	1.63 [0.69, 3.86]
Subtotal (95% CI)	170	224		100.00	1.63 [0.69, 3.86]
Total events: 12 (Laxatives), 10 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.10 (P = 0.27)					
02 overall; Subgroup of lactulose vs ispaghula 7g					
Dettmar 1998	4/91	10/224		100.00	0.98 [0.30, 3.22]
Subtotal (95% CI)	91	224		100.00	0.98 [0.30, 3.22]
Total events: 4 (Laxatives), 10 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.03 (P = 0.98)					
03 nausea; lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	3/39	1/39		100.00	3.17 [0.31, 31.86]
Subtotal (95% CI)	39	39		100.00	3.17 [0.31, 31.86]
Total events: 3 (Laxatives), 1 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.98 (P = 0.33)					
04 nausea; Other laxatives mainly lactulose vs ispaghula 7g					
Dettmar 1998	12/170	18/224		100.00	0.87 [0.41, 1.86]
Subtotal (95% CI)	170	224		100.00	0.87 [0.41, 1.86]
Total events: 12 (Laxatives), 18 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.36 (P = 0.72)					
05 nausea; Subgroup of lactulose vs ispaghula 7g					
Dettmar 1998	7/91	18/224		100.00	0.95 [0.38, 2.37]
Subtotal (95% CI)	91	224		100.00	0.95 [0.38, 2.37]
Total events: 7 (Laxatives), 18 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.10 (P = 0.92)					
06 anorexia; lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	1/39	0/39		100.00	3.08 [0.12, 77.91]
Subtotal (95% CI)	39	39		100.00	3.08 [0.12, 77.91]
Total events: 1 (Laxatives), 0 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.68 (P = 0.50)					
07 headache; lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	2/39	2/39		100.00	1.00 [0.13, 7.48]
Subtotal (95% CI)	39	39		100.00	1.00 [0.13, 7.48]
Total events: 2 (Laxatives), 2 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.00 (P = 1.00)					

0.1 0.2 0.5 1 2 5 10

Favours Laxatives   Favours Fibre

4. Patient preference

The crossover study, Quah (2006), recorded patient preference at four weeks between lactulose and ispaghula. Statistically significantly more patients preferred lactulose; RR 1.71 (95%CI 1.05, 2.79). This gave an NNT of 4 (95%CI 3, 25). It was noted that this crossover study had a washout period of only one week, so the results were treated with caution.

Figure 51:

Study or sub-category	Laxatives n/N	Fibre n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 lactulose 20ml vs ispaghula 3.5g crossover					
Quah 2006	24/39	14/39		100.00	1.71 [1.05, 2.79]
Subtotal (95% CI)	39	39		100.00	1.71 [1.05, 2.79]
Total events: 24 (Laxatives), 14 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.17 (P = 0.03)					
Total (95% CI)	39	39		100.00	1.71 [1.05, 2.79]
Total events: 24 (Laxatives), 14 (Fibre)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.17 (P = 0.03)					

0.1 0.2 0.5 1 2 5 10

Favours Fibre   Favours Laxative

Adverse Effects

An adverse effects review has been carried out and is reported in section 8.5.1. The review included six RCTs (Quah 2006; Ferguson and Attar 1999; Bouhnik 2004; Corazziari 1996; Corazziari 2000; Chaussade 2003), and their results are reported in this effectiveness review. The RCTs were primarily aimed at assessing and reporting on the efficacy of the drug treatments. Evaluation of safety and reporting of adverse effects data was often cursory or non-existent. Even in instances where the methods sections had explicitly stated the intention of monitoring for adverse effects, trial reports did not follow a structured format (e.g. by WHO system organ class) of reporting adverse effects. The interventions and comparators were extremely varied, as was the reporting of adverse effects.

One non-randomised study in the US reported a series of adverse effects of laxatives, but did not distinguish laxative class, and used doses higher than in the UK.

Many of the adverse outcomes of interest are very similar to the symptoms of the IBS itself. For instance, laxatives are associated with flatulence, cramps and abdominal pain – all of which are commonly seen in untreated IBS patients and also form part of the efficacy assessment. It is not always possible to determine whether deterioration in these symptoms is due to lack of efficacy, or the natural history of the disease, or the adverse effect of the drug. Generally, though, the RCT data on lactulose was consistent with the findings of the non-randomised data with regards to increased risk of abdominal symptoms. The GDG's clinical experience of lactulose was that it caused bloating and other side effects.

ECONOMIC LITERATURE FOR LAXATIVES

One relevant health economic analysis was identified on the cost-effectiveness of laxatives in the treatment of IBS. Christie (2002) was a model based economic evaluation from a UK perspective which used efficacy data from a secondary care trial comparing two laxatives conducted in Scotland and France. The population included in the trial was patients with idiopathic constipation of greater than three months duration and the population included some elderly patients living in institutions. Whilst data from elderly residential patients is not directly

relevant to the IBS population, this paper was included as indirect evidence for the IBS population.

This study aimed to assess the economic impact of using low dose polyethylene glycol 3350 plus electrolytes (PEG+E) compared to lactulose in the treatment of idiopathic constipation using a decision analytic model. The economic analysis was carried out from an NHS perspective. As discussed earlier, this study was considered to be indirect evidence as the patient population was not restricted to patients with IBS but may have included some patients with IBS-C. The effectiveness inputs used in the model were obtained from a randomised controlled trial conducted in primary care. In this trial patients were randomised to treatment with either PEG+E or lactulose for one month. After this initial comparator controlled phase, patients aged over 65 continued on their allocated treatment for two months but those aged under 65 received lactulose for a further 2 months regardless of their initial allocation. The model considered the probability of various clinical outcomes over 2 weekly intervals for a 3 month period. The outcomes considered by the model were; successful treatment, discontinuation of treatment due to an adverse event, switching laxatives due to an adverse event, discontinuation of treatment due to lack of efficacy, switching laxatives due to lack of efficacy, not complying with either treatment and discontinuing treatment, not complying with either treatment and switching to another laxative. Resource use estimates were provided by a panel of experts.

PEG+E had a higher probability of achieving successful treatment at 3 months (53% versus 24%) but a higher acquisition cost (£25.42 versus £10.05 over 3 months). This was offset by a reduced number of GP appointments (2.9 visits versus 4.4 visits), resulting in an overall lower cost in patients initially treated with PEG+E (£85 versus £96). The sensitivity analyses showed that the overall costs were particularly sensitive to changes in the efficacy of first-line treatment with either treatment, the mean daily dose for PEG+E, the probability of senna being co-prescribed with lactulose, the probability of discontinuing treatment with lactulose and the number of GP appointments. Given that the costs were sensitive to dose it is important that this study is considered along-side evidence on the effective dose in patients with IBS. The model assumed that co-prescription of senna is more frequent in patients not experiencing successful resolution of symptoms following lactulose treatment (13%) than following PEG+E treatment (2%). In the trial patients were not allowed to take additional laxatives, so the effectiveness of adding senna to lactulose in this way would not be captured in the model but the cost of co-prescribing senna has been included in the model. Assuming no senna use in the lactulose arm reduced the cost to £87 which suggests that the cost-effectiveness is sensitive to the accuracy of this assumption.

The study was a partial economic evaluation as it did not assess the incremental cost of any benefit achieved in the form of a cost-effectiveness ratio. This may be appropriate given that the intervention was more effective than the comparator. However, the effectiveness was only

measured in terms of the probability of successful treatment rather than overall health impact. This may be misleading if adverse events have a higher impact on health than successful treatment. Adverse events were included in the analysis but from a cost perspective only. The evidence provided by this study was not directly relevant to the guideline as it considered a patient population that isn't fully representative of the population considered by this guideline. No potential areas of significant bias were identified, but the sensitivity analysis demonstrated that the magnitude of cost-saving estimated by the model was variable under the parameter ranges considered. Modelled direct health care costs were lower in the PEG+E arm despite a higher acquisition cost. As this study did not provide an estimate of the cost per QALY for PEG+E compared to lactulose, and did not consider the cost-effectiveness of either intervention compared to no laxative treatment, it was not particularly useful in determining whether recommending PEG+E or lactulose would result in the efficient use of NHS resources.

COST-EFFECTIVENESS ANALYSIS FOR LAXATIVES

This section describes the health economic analysis undertaken to inform recommendations on the use of laxatives as a long-term maintenance therapy in IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

The general approach was the same as for other maintenance therapies except:

- None of the trials provided an estimate of the relative risk of an improvement in global symptom score, which was the favoured outcome for determining a successful response to treatment for the economic model. An improvement in bowel habit was considered as a possible alternative definition for response, but this was also not available for any of the long-term maintenance studies. In the absence of this, a successful response was defined as no use of other laxatives.

- For sodium picosulfate and bisacodyl there was evidence for their effectiveness compared to placebo for short term use (3 days) but there was no evidence on their effectiveness compared to placebo for long-term use. In the absence of evidence on the effectiveness of long-term maintenance use, we have applied the effectiveness from the short-term trials and assumed that it would persist in the long-term. This is an extreme extrapolation beyond the available trial duration and should be considered with caution.

- PEG, sodium picosulfate and bisacodyl were included in the economic model as potential laxative treatments. We assumed that PEG is used first line as this was the only intervention with evidence of clinical effectiveness in long-term maintenance use. Sodium picosulfate and bisacodyl are assumed to be used second line in patients who do not respond to PEG.

- Lactulose was less effective than PEG 3350 (with electrolytes) (Attar 1999). GDG consensus was that people with IBS should be actively discouraged from taking Lactulose as it promotes gaseous bloating which can exacerbate IBS symptoms. It was therefore excluded from the cost-effectiveness analysis.

- The studies included in the clinical effectiveness review did not stratify results by IBS subtype, but all the studies were carried out in patients with chronic constipation. Therefore, the cost-effectiveness is estimated for patients with IBS-C. The applicability of these results to people with IBS-A was considered by the GDG as they may have intermittent periods of chronic constipation.

Modelled response rates

In the basecase scenario the response rate of 45% in the no treatment arm is taken from the Mearin (2004) cohort study. This represents the group of patients whose symptoms improve without any specific intervention. The RR of response for PEG versus placebo is 1.61; therefore the response rate in the PEG arm is 72% (=45% x 1.61), giving an absolute difference in response between the intervention and no treatment arms of 27% (=72%-45%) during the first month for PEG. The RR of bisacodyl and sodium picosulfate is 1.34 compared to placebo, so if these interventions are used first line then we would expect an absolute difference in response between intervention and no treatment of 15% (=1.34*45%-45%). The first line use of bisacodyl and sodium picosulfate has not been modelled due to a lack of longer-term data on their effectiveness compared to placebo.

In the basecase scenario the response rate for the subsequent interventions is assumed to be equal to the response rate to the first intervention. If bisacodyl and sodium picosulfate are used second line in patients who do not respond to PEG, we would expect an additional 4.2% (=15% x 28%) of the original cohort to respond to the second laxative, and an additional 3.6% (=15% x 24%) to respond to the third laxative, giving an overall response rate of 80% for laxatives. The response rate over time for the basecase is given in Figure 52. It is assumed that bisacodyl is tried first after a failure to respond to PEG as it has a lower cost than sodium picosulfate.

We have also considered an alternative scenario in which no patient in the comparator arm achieves an improvement in symptoms, but the absolute gain in response rates is maintained from the basecase (e.g. for first line PEG, we modelled a zero response to no treatment but a 27% response to PEG for this scenario).

Figure 52: Modelled response rates for laxatives (PEG followed by two switches to other laxatives for non responders) and no treatment
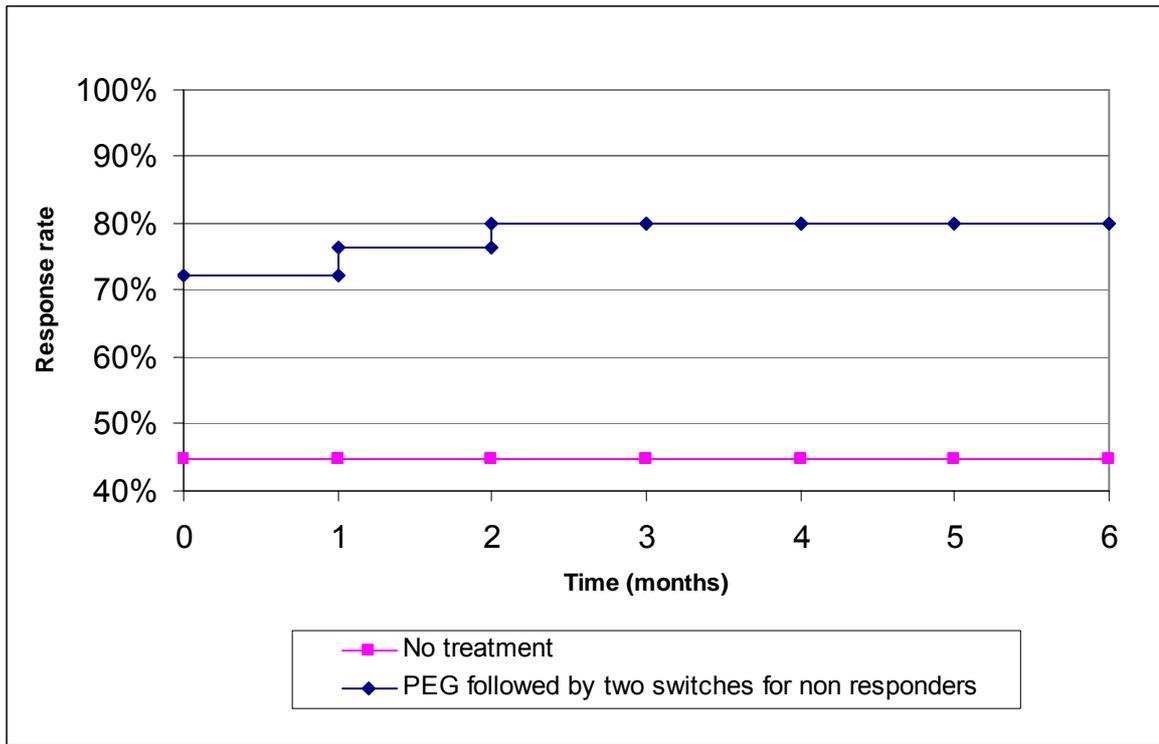
Table 1: Intervention specific parameters – laxatives

Description	Value		Evidence
RR of response for PEG vs placebo	1.61		Meta-analysis of RCT evidence for no use of other laxatives
RR of response for bisacodyl and sodium picosulfate compared to placebo	1.34		Meta-analysis of RCT evidence for improvement in bowel habit
Maximum number of switches considered	2		Limited by number of effective interventions
Drug costs			
Intervention	**Dose per day**	**Cost per month* (assuming lowest cost preparation)**	
PEG	23g (equiv to 1.8 sachets of Movicol or 2.3 sachets of Idrolax)	£12.54	
Bisacodyl	10mg	£1.43	
Sodium picosulfate	7mg	£3.94	

* British National Formulary (Joint Formulary Committee 2007)

Table 2: Incremental cost-effectiveness of allowing subsequent switches in laxative therapy

Strategy	Cost	QALY	Incremental Cost per QALY compared to previous row	Incremental cost per QALY compared to no treatment
No treatment	£0	1.60	N/A	N/A
Intervention, no switches	£7,575	2.57	£7,779	£7,779
Intervention with up to 1 switch	£8,141	2.70	£4,488	£7,401
Intervention with up to 2 switches	£8,703	2.78	£6,561	£7,341

Table 2 gives the incremental cost-effectiveness for several laxative treatment pathways in order of the benefits they achieve. It shows that whilst PEG provides additional benefit for a cost per QALY of £7,779, compared to no treatment, further benefit can be achieved by allowing non responders to PEG to switch to bisacodyl and if that is not effective to switch to sodium picosulfate. Each of these additional switches for non responders has a low cost per QALY compared to no further treatment for non-responders (£4,488 and £6,561 respectively). However, it should be noted that the cost-effectiveness of these second line laxatives is based on clinical effectiveness evidence for bisacodyl and sodium picosulfate from short term trials lasting only 3 days.

These results are an estimate of the cost-effectiveness over the first 6 months after the initiation of laxative therapy. The cost per QALY for continuing laxative therapy beyond 6 months is lower than the cost per QALY during the initial 6 months provided that treatment is reviewed every 6 months and discontinued in patients who no longer experience a therapeutic benefit, either due to lack of effectiveness or a change in their symptom profile.

The probabilistic sensitivity analysis provides an estimate of the uncertainty in the cost per QALY estimate due to uncertainty in the efficacy estimate, the probability of response in the no treatment arm and the utility gain. The CEAC in Figure 53 shows the uncertainty surrounding the cost-effectiveness of PEG, compared to no treatment and the incremental cost-effectiveness of allowing non-responders to switch to other laxatives. It shows that there is an 83% likelihood that the cost per QALY for PEG compared to no treatment is under £20K, suggesting that PEG provides health benefit at an acceptable cost in patients with IBS-C. Allowing non responders to PEG one treatment switch has an 81% probability of a being cost-effective when a £20K threshold is applied compared to no further treatment for non responders. Similarly, allowing a second treatment switch for non-responders has a 74% probability of being cost-effective.

Figure 53: CEAC for PEG with up to two switches for non responders compared to no treatment (NT)
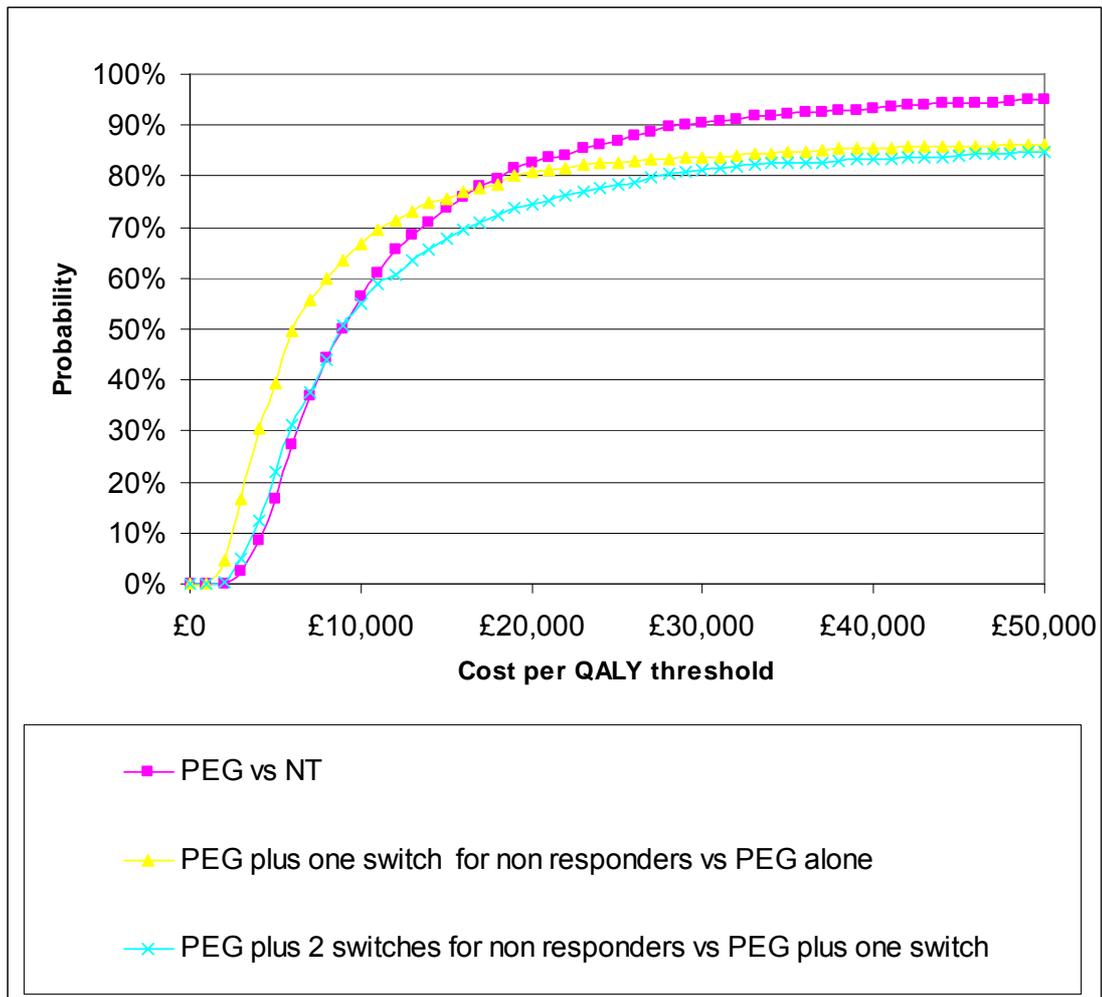
However, it should be noted that these estimates only consider the uncertainty in cost-effectiveness due to the accuracy of several input parameters and they do not reflect general uncertainty around the assumptions made in the model. The uncertainty from these assumptions was explored in the univariate sensitivity analysis.

Univariate sensitivity results for laxatives

The results of the univariate sensitivity analysis for PEG compared to no treatment are given in Table 3. Maintaining the 27% difference in response between the two arms but reducing the response rate in the no treatment arm from 45% to zero decreased the cost per QALY to £4,896. The use of higher cost formulations increased the cost per QALY to £9,980, whilst assuming that 50% of prescriptions were over the counter reduced the cost per QALY to £4,814.

The cost per QALY is less favourable for patients who only use the medication on 25% of days as the upfront costs of initiating therapy and establishing response are constant despite lower

benefit from less frequent use. PEG compared to no treatment has a cost per QALY of £13,325 when used on 25% of days.

We carried out a threshold analysis to determine whether laxative therapy would still be cost-effective for lower gains in health related quality of life. In the basecase it was assumed that patients who respond to therapy accumulate 0.071 QALYs more per annum than patients who do not respond. For comparison, a gain of 0.135 QALYs would represent a complete remission of IBS symptoms. If the QALY gain associated with a response to therapy was reduced to 0.027 QALYs, then the cost per QALY of providing PEG compared to no treatment would be above £20,000 per QALY. The methods used in this analysis vary from the methods used for other pharmacological interventions as we were unable to estimate the response rate in terms of an improvement in global symptoms from the trial data available. Instead we assumed that patients who did not use another laxative during the trial had responded to the trial intervention. This outcome was considered a less reliable indicator of whether there had been an overall improvement in HRQoL than an improvement in global symptoms. However, the threshold analysis shows that laxatives are cost-effective even if the utility gain associated with a therapeutic response is small.

We carried out a similar univariate sensitivity analysis on the incremental cost-effectiveness of allowing non responders to switch to an alternative laxative. The incremental cost per QALY for the first and second switches was increased to £8,784 and £12,624 when assuming that patients who demonstrate no response to PEG would be half as likely to respond to another laxative. The incremental cost per QALY for each subsequent treatment switch in non-responders was higher for patients using treatments intermittently with a cost per QALY of £8,468 for the first switch and a cost per QALY of £11,536 for the second switch in patients who use laxatives on only 50% of days. When laxatives are used on only 25% of days, the incremental cost per QALY estimates for the first and second switches are £16,428 and £21,485 respectively.

If a patient also takes another medication (an antispasmodic), then this medication can be reviewed at the same time, so it may be cost-effective to provide both therapies. For example, if laxatives are prescribed with the antispasmodic and both used on 25% of days then allowing up to 2 switches of both treatments was estimated to be cost-effective with a cost per QALY of £10,107 compared to no treatment a cost per QALY of £17,393 compared to 1 switch.

Table 3: Sensitivity results for PEG compared to no treatment for 100 patients with IBS-C

Scenario	No Treatment		Intervention		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	1.60	£7,575	2.57	£7,779
No response in no treatment arm	£0	0.00	£4,767	0.97	£4,896
Response rate in no treatment arm from RCTs	£0	1.69	£7,847	2.72	£7,601
Treatment used 75% of days	£0	1.60	£6,131	2.33	£8,395
Treatment used 50% of days	£0	1.60	£4,687	2.08	£9,628
Treatment used on 25% of days	£0	1.60	£3,244	1.84	£13,325
Half of treatment obtained over the counter	£0	1.60	£4,687	2.57	£4,814
Higher cost formulations (same dose)	£0	1.60	£9,717	2.57	£9,980
High utility gain of 0.135	£0.00	3.02	£7,575	4.87	£4,109
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.028 and 0.027.				

GDG DISCUSSION

Many of the studies included in the laxative review may be considered to be indirect evidence as the participants were defined as having simple constipation. However the GDG considered that many of these participants may have had IBS, but the studies did not use any IBS assessment criteria and the trials were designed to treat the symptoms of constipation. General consensus is that IBS is very different from simple constipation. People with IBS cannot cope with gas and some laxatives increase gas and exacerbate IBS symptoms, lactulose in particular. IBS patients should be actively discouraged from taking lactulose. The GDG also referred to best practice of titrating the dose of laxative to optimise symptoms, using the Bristol Stool Form Scale.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. In studies for short-term symptom relief of constipation there is a moderate amount of good evidence to show a significant improvement in bowel habit for stimulant laxatives (bisacodyl

and sodium picosulphate) compared to placebo. The study population included patients with IBS.

2. There were no studies identified that used global improvement of symptoms as an outcome in longer-term maintenance treatment of constipation for PEG.

3. In studies for longer-term maintenance treatment with PEG versus placebo in people with constipation (including people with IBS) there is a:
- Limited amount of good evidence that those taking PEG required significantly less rescue medication than those taking placebo.
- Limited amount of good evidence showing no significant difference in pain.
- Limited amount of good evidence showing significant reduction in bloating
- Moderate amount of good evidence showing a large significant improvement in bowel habit.

4. There are no trials of longer term treatment that compared:
- Lactulose versus placebo
- Bisacodyl versus placebo
- Sodium picosulphate versus placebo.

5. In studies for longer-term maintenance of PEG versus lactulose in people with constipation (including people with IBS) there is a:
- Fair amount of evidence showing significant improvement in global symptoms
- Moderate amount of good evidence that those taking PEG required significantly less rescue medication than those taking lactulose
- Moderate amount of good evidence showing a significant improvement in stool frequency.

6. In studies for longer-term maintenance of bisacodyl versus sodium picosulphate in patients with constipation (including participants with IBS) there is a moderate amount of good evidence that there is no significant difference in stool frequency.

7. In studies for longer term maintenance of PEG + Electrolyte versus PEG-Electrolyte in people with constipation (including participants with IBS) there is a:
- Moderate amount of good evidence to show that both are equally effective with no significant difference in pain, bloating, stool frequency, the number of people with normal stools, quality of life and adverse effects.

8. In studies for longer-term maintenance treatment with standard and maximum dose PEG in people with constipation (including people with IBS) there is a:

- Moderate amount of good evidence to show that both are equally effective with no significant difference in pain, bloating, quality of life and adverse effects
- Moderate amount of good evidence showing a significant increase in the number of people with normal stools (standard dose)
- Fair amount of good evidence showing a significant increase in the incidence of people with diarrhoea (maximum dose).

ADVERSE EFFECTS EVIDENCE STATEMENTS

1. There is limited evidence that laxatives are significantly associated with GI adverse effects (Abdominal cramps, abdominal discomfort, bloating, diarrhoea, abdominal pain, nausea).

2. There is consistent evidence that lactulose increases the risk of abdominal symptoms in people with IBS.

3. There is moderate evidence that low dose PEG is associated with fewer adverse effects compared to high dose PEG.

HEALTH ECONOMIC EVIDENCE STATEMENT

Evidence from a published model based economic evaluation comparing PEG with lactulose showed that PEG dominates lactulose by achieving a higher response to treatment rate at lower overall cost. The study was a partial economic evaluation as it did not assess the overall impact on health or provide the incremental cost of any benefit achieved in the form of a cost-effectiveness ratio. It is also considered to be indirect evidence as the population was not fully representative of the IBS population.

Evidence from a decision analytic model showed that laxatives (polyethylene glycol (PEG), bisacodyl and sodium picosulfate) are cost-effective for long-term maintenance use in individuals with IBS. The cost-effectiveness estimate is based on a clinical pathway in which response is assessed after one month and non-responders are switched to an alternative laxative with PEG used first line followed by bisacodyl and then sodium picosulfate. The cost-effectiveness analysis assumes that treatment is reviewed every 6 months to establish whether it is still relevant to the individual's symptom profile.

EVIDENCE TO RECOMMENDATIONS

The evidence from the review suggests that laxatives are clinically and cost effective in the management of constipation. However the GDG clinical opinion is that IBS is more complex than simple constipation. Some laxatives exacerbate IBS symptoms and should therefore be avoided by people with IBS. The GDG recommended the continuation of current best practice of titrating the dose of laxative to optimise symptoms, based on the Bristol Stool Form Scale.

RECOMMENDATION

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

RECOMMENDATION

People with IBS should be advised how to adjust their doses of laxative or antimotility agent according to the clinical response. The dose should be titrated according to stool consistency, with the aim of achieving a soft, well-formed stool (corresponding to Bristol Stool Form Scale type 4).

8.2 Antimotility agents

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the antimotility agents review and are reported below.

Types of participants

For this review, participants were required to have IBS and not to have inflammatory bowel disease or diarrhoea subsequent to surgery. This inclusion criterion was adhered to for the longer term maintenance studies, but, for short term relief of symptoms investigations, there were insufficient data for IBS patients. Therefore, for this section of the review only, the GDG extended the population, post-hoc, to include studies in patients with acute diarrhoea of any cause (including those with diarrhoea caused by infection or virus). Such studies were regarded as indirect as far as the population was concerned.

Types of studies

The GDG decided that the washout period for this review should be at least one week. Trials with shorter washout periods were not included in the analysis.

Types of intervention

Studies included the following interventions:

- Codeine phosphate
- Co-phenotrope (diphenoxylate and atropine mixture; Trade name: Lomotil®)
- Loperamide
    - Single drug: loperamide hydrochloride (Trade names: Norimode®, Imodium®)
    - Compound preparation: loperamide hydrochloride and simeticone (Trade name: Imodium® Plus)
- Morphine
    - Kaolin and Morphine mixture BP
    - Morphine preparations on sale to the public.

The following comparisons were included:

- Antimotility agent versus placebo (or nothing)
- Antimotility agent type 1 versus type 2
- Antimotility agent dose 1 versus dose 2
- Antimotility agent + another intervention versus the other intervention alone
- Antimotility agent delivery mode 1 versus delivery mode 2
- Duration of treatment 1 versus duration 2.

NB: In spite of the large placebo effect associated with IBS, comparisons with no treatment are included.

The antimotility agents review was concerned with both longer term maintenance treatment and short-term symptom relief.

For maintenance studies, the GDG had decided that there should be a minimum duration of treatment of four weeks, but on further reflection agreed to include studies of two weeks or more. Studies of shorter durations were excluded. Short-term /symptom relief studies had duration of less than one week.

Subgroup analyses

We planned to carry out subgroup analyses by type of antimotility agent, dose, mode of delivery (modified release/conventional) and duration of intervention.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

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. The search strategies are listed in Appendix B.

The search strategy identified 2869 possible studies. The titles and abstracts of these studies were assessed. Forty were identified to be potentially relevant to the review and these papers were retrieved in full. The reference lists for each of the retrieved studies were inspected for further potential papers but none were identified. The 18 excluded studies are listed in Appendix E, along with reasons for exclusion.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

There were 22 included studies, six of which had at least some patients with IBS and 16 were in an indirect population (Allison 1988; Amery 1975; Barbezat 1979; Cann 1984; Corbett 1980; Cornett 1977; Dettmer 1994; Dom 1974; Dreverman 1995; Efskind 1995; Ericsson 1990; Harford 1980; Hovdenak 1987; Jaffe 1977; Lavö 1987; Lee 1968; Lustman 1987; Palmer 1980; Pelemans and Vantrappen 1976; Taneja 2004; Tijtgat 1975; Verhaegen 1974). Seven were conducted in the UK (Allison 1988; Cann 1984; Corbett 1980; Jaffe 1977; Lee 1968; Lustman 1987 and Palmer 1980); ten in the rest of Europe, three in the USA; one in India and one in South Africa.

All studies but six had fewer than 100 patients, with seven having 20 or fewer in the intervention arm (Allison 1988; Harford 1980; Hovdenak 1987; Lavö 1987; Taneja 2004; Tijtgat 1975; Verhaegen 1974). One study included 227 patients in total but only 46 in the loperamide group and 45 in the placebo group (Ericsson 1990). The remaining studies had between 152 and 614 patients (Dom 1974). Some of the studies were of crossover design, so fewer patients are required to achieve adequate power.

Study Design

There were seven crossover studies (Allison 1988; Cann 1984; Corbett 1980; Harford 1980; Palmer 1980; Pelemans and Vantrappen 1976; Verhaegen 1974) in which participants were allocated to receive both the intervention and control treatments during the course of the study, in a random order. Four of these studies (Allison 1988; Cann 1984; Corbett 1980; Palmer 1980) had either no washout period or it was not reported, in which case this was assumed to be none. One acute study had a washout period of 12 to 24 hours (Harford 1980). One longer term study had a washout period of 2 to 7 days, and the drugs were discontinued until severe diarrhoea returned (Verhaegen 1974). The other longer term study (Pelemans and Vantrappen 1976) had a washout period of 3 to 20 days (median 7 days). As the GDG had specified a washout period of one week minimum for maintenance studies, the latter three studies were included on the basis of washout period, but those with no washout were excluded from the analysis and transferred to the excluded studies table. The remaining studies had a parallel design.

The GDG had specified a minimum treatment period of four weeks for each intervention in the maintenance studies. One study had a treatment duration of one week (Barbezat 1979); one had a duration of at least 10 days or until relapse (Tijtgat 1975); one had at least 12 days (Verhaegen 1974), one had from 14 to 49 days, median 25 days (Pelemans and Vantrappen 1976); one had three weeks (Hovdenak 1987). The GDG decided to exclude Barbezat (1979); Tijtgat (1975), and; Verhaegen (1974) on this basis, and to accept Hovdenak (1987), and; Pelemans and Vantrappen (1976). The remaining studies had durations of seven weeks (Efskind 1995), 2 months (Taneja 2004) and 13 weeks (Lavö 1987).

Thus, 15 studies were included in the analysis: thirteen parallel (Amery 1975; Cornett 1977; Dettmer 1994; Dom 1974; Dreverman 1995; Efskind 1995; Ericsson 1990; Hovdenak 1987; Jaffe 1977; Lavö 1987; Lee 1968; Lustman 1987; Taneja 2004), and two crossover (Harford 1980; Pelemans and Vantrappen 1976).

Ten studies investigated the use of anti-motility agents for acute diarrhoea (Amery 1975; Cornett 1977; Dettmer 1994; Dom 1974; Dreverman 1995; Ericsson 1990; Harford 1980; Jaffe 1977; Lee 1968; Lustman 1987). Only one of these had some patients with IBS (Harford 1980). Five studies investigated the effectiveness of anti-motility agents for the treatment chronic diarrhoea in IBS.

Three studies had more than two arms. Amery (1975) compared diphenoxylate with loperamide and placebo. Dettmer (1994) and Dreverman (1995) compared two dose of loperamide with placebo, giving a total of 21 comparisons.

Setting: seven studies were in primary care (Amery 1975; Dom 1974; Dreverman 1995; Efskind 1995; Jaffe 1977; Lee 1968; Lustman 1987); three were in secondary care (Dettmer 1994; Lavö 1987; Taneja 2004) and the others did not report the setting.

Funding: four studies (Amery 1975; Cornett 1977; Dettmer 1994; Dreverman 1995) were from Janssen Pharmaceutica (manufacturers of Imodium, i.e., loperamide) and Efskind (1995) stated that Janssen Pharmaceutica provided the drug, monitored the study and gave statistical support. Lustman (1987) did not specify funding, but the corresponding author was employed by Gold Cross Pharmaceuticals, a division of GD Searle & Co Ltd (manufacturers of Lomotil). Three studies were funded by non-industry sources (Ericsson 1990; Harford 1980; Taneja 2004) and the others did not state their funding.

Population

Four studies were in patients with IBS (Efskind 1995; Hovdenak 1987; Lavö 1987; Taneja 2004), and two studies had some patients with IBS. The acute study, Harford (1980), had 4/15 patients with IBS; and the maintenance study, Pelemans and Vantrappen (1976), had 4/23; 18 of the remaining patients in the latter study had inflammatory bowel disease. In both studies, individual patient data were reported for the IBS subgroup, but these were not stratified before randomisation and the small numbers give uncertainty and likely potential for bias. The definition of IBS varied between studies: one included patients meeting the Rome II criteria (Taneja 2004), one met criteria defined by the authors that were similar to the Rome criteria (Efskind 1995), and in the other studies, the authors stated that the patients had IBS, with no further explanation.

Most studies included patients with diarrhoea predominance, but one study (Hovednak 1987) had a mixture of types: IBS-D (16); IBS-A with pain (21); IBS-A without pain (12); IBS-C (9). None of the studies stated that any participants had IBS as result of gastrointestinal infection. The majority of studies did not state the number of participants with bloating. One study had some patients with bloating (Efskind 1995).

Most of the studies did not describe symptom severity. Two studies stated that participants had symptoms of mixed severity (Efskind 1995; Hovednak 1987), one of which excluded patients with mild symptoms (Hovednak 1987).

The remaining studies were in patients who did not have IBS and these were treated as indirect evidence. Further details are given in the included studies table.

The age range of participants across the IBS studies was 18 to 70 years, with the mean age, where given, ranging from 31 to 43 years. No study particularly identified elderly participants. The indirect studies included patients aged 9 to 95 years, with four of the studies definitely including children: Amery (1975) included patients aged 9 to 82 years, with a median age of 31

years; Cornett (1977) had an age range of 11 to 84 years, with 8% patients in the age group 10 to 19 years; Dom (1974) had a range of 14 to 95 years with a median of 35 years; and Drevermann (1995) had patients aged 16 to 75 years.

Five studies had more women than men (Efskind 1995; Harford 1980; Lavö 1987; Lee 1968; Lustman 1987); two studies had about the same number of men and women (Dettmer 1994; Pelemans and Vantrappen 1976); one study examined only men (Taneja 2004) and the other indirect studies had more men than women. One study did not report the numbers of men and women (Ericsson 1990).

Interventions

The studies varied in the type of antimotility agent used:

- No studies examined codeine phosphate
- Seven acute studies gave the patients co-phenotrope (Amery 1975; Cornett 1977; Dom 1974; Harford 1980; Jaffe 1977; Lee 1968; Lustman 1987). The Amery study stated that they gave the patients 2.5mg diphenoxylate and that the contents of the capsules were identical to Lomotil (co-phenotrope)
- Twelve studies (seven acute) gave the patients loperamide (Amery 1975; Cornett 1977; Dettmer 1994; Dom 1974; Dreverman 1995; Efskind 1995; Ericsson 1990; Hovdenak 1987; Jaffe 1977; Lavö 1987; Pelemans and Vantrappen 1976; Taneja 2004)
- One acute study examined morphine (Lee 1968 used a kaolin and morphine mixture).

Ten of the included studies used antimotility agents for acute diarrhoea (Amery 1975; Cornett 1977; Dettmer 1994; Dom 1974; Dreverman 1995; Ericsson 1990; Harford 1980; Jaffe 1977; Lee 1968; Lustman 1987).

In the maintenance studies, no study allowed rescue medication, but patients were allowed to vary the dose of study drug in three studies (Efskind 1995; Lavö 1987; Pelemans and Vantrappen 1976). A fixed dose was used in the remaining studies.

Comparisons

The included studies covered the following comparisons:

- Ten comparisons of antimotility agent versus placebo:
    - Three gave co-phenotrope for acute episodes (Amery 1975; Harford 1980; Lustman 1987)
    - Four gave loperamide for acute episodes (Amery 1975; Dettmer 1994, in a dose of 1mg or 2mg; Dreverman 1995 in a dose of 1mg or 0.5mg; Ericsson 1990 2mg)
    - Three gave loperamide for maintenance treatment (Efskind 1995; Hovdenak 1987; Lavö 1987).
- Three comparisons of different types of antimotility agent:

- o Four studies compared loperamide and diphenoxylate for acute episodes (Amery 1975; Cornett 1977; Dom 1974; Jaffe 1977)
  - o One study compared loperamide and co-phenotrope for maintenance treatment (Pelemans and Vantrappen 1976)
  - o One study compared co-phenotrope with kaolin-and-morphine for acute episodes (Lee 1968).
- Two acute studies compared two doses of loperamide (Dettmer 1994: 1mg versus 2mg; Dreverman 1995 1mg versus 0.5mg)
- One maintenance study compared loperamide with yoga (Taneja 2004).

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D. The method of randomisation was reported in none of the studies. Allocation concealment was reported in one study (Pelemans and Vantrappen 1976), which reported a partially adequate method in which bottles of the drug were marked with the patient's code number and contained identical capsules.

All the studies reported that the patients were blinded to the interventions except for three. Jaffe (1977) reported that the drugs were presented in their normal marketed form and as they are dissimilar in appearance and had different dose regimens, no attempt was made to blind participants or investigators. Lee (1968) did not report blinding and the treatments used were dissimilar: Lomotil-with-neomycin was used at the recommended dose of 4 tablets at the start of therapy then 2 tablets every 6 hours, while the kaolin-and-morphine mixture was used at 2 tablespoons at the start of treatment then 1 tablespoon every 6 hours. Taneja (2004) compared loperamide with yoga and this was not blinded. Only two studies reported a sample size calculation (Dettmer 1994; Ericsson 1990).

Most studies included in the review demonstrated baseline comparability of the groups, but two studies were not comparable at baseline for the age of the patients (Amery 1975 had significantly older patients in the loperamide group; Lavö 1987 had significantly younger patients in the loperamide group). The GDG did not regard these differences to be important.

Five acute diarrhoea studies reported no withdrawals (Amery 1975; Cornett 1977; Dom 1974; Jaffe 1977; Lee 1968). One of the 13 patients allocated to loperamide in the Taneja (2004) study could not attend for the final assessment but the nine treated with yoga all attended. Dettmer (1994) reported missing data for 7% (13/230) of patients. Dreverman (1995) reported missing data for 3% (8/242) of patients. Lustman (1987) reported missing data for 4% (7 of 152) of patients.

In the Ericsson (1990) acute diarrhoea study, 46 patients were allocated to the loperamide group and 45 to the placebo group. Of these, 6 and 4 patients respectively dropped out and were excluded from the efficacy analysis. In addition, those who were non-compliant with medication (13 and 14 patients respectively) were also excluded from the analysis. This led to the final numbers being 27 for each group (i.e. missing data for 41% in each group). This raises the potential for bias in this study.

Two maintenance studies reported that more than 20% of patients in at least one arm (or overall) were not analysed or were lost to follow-up (attrition bias). Efskind (1995) reported 23% (21/90) missing overall, but 8 of these did not arrive at the start of the trial (thus drop-outs related to treatment were 12/90, i.e. 13%). Pelemans and Vantrappen (1976) had 26% (6/23) missing overall for the stool frequency outcome measure, but none of the IBS patients were missing.

Two of the studies (Harford 1980; Pelemans and Vantrappen 1976) each gave individual patient data for 4 IBS patients. This was a within-trial subgroup analysis and stratification had not taken place. The GDG decided to consider the results for all patients in the Harford (1980) acute study, but decided that the Pelemans and Vantrappen (1976) maintenance study had too few IBS patients and would be misleading, so this study was not considered further.

The risk of bias was assessed for each included study and two studies were excluded from the analysis: Pelemans and Vantrappen (1976) was excluded on the basis of wrong population, and; Ericsson (1990) was considered to have too high a drop-out rate. None of the other studies were considered to be at risk of bias, although the inclusion of children in some studies was noted.

RESULTS
I. TREATMENT FOR ACUTE EPISODES OF DIARRHOEA

For the treatment of acute diarrhoea the GDG simply wished to know whether anti-motility agents were effective in stopping diarrhoea, regardless of its cause. Therefore only outcomes relating to bowel habit are reported.

A. Anti-motility agent versus placebo

Three studies gave co-phenotrope for acute episodes (Amery 1975; Harford 1980; Lustman 1987).

- Amery (1975) gave 5 mg diphenoxylate per day (1 tablet twice a day) for 24 hours
- Harford (1980) gave 10 or 20 mg per day (1 or 2 tablets 4 times a day) for 3 days.
- Lustmann (1987) gave 10 mg initially (4 tablets) then 20mg per day (2 tablets four times per day) for 3 days.

The recommended initial dose for co-phenotrope is initially 10 mg (4 tablets) then 20mg per day (2 tablets four times a day), thus the Amery (1975) study can be considered to have a low dose of diphenoxylate.

Three studies gave loperamide for acute episodes compared with placebo (Amery 1975; Dettmer 1994, 2 doses; Dreverman 1995, 2 doses).

- Amery (1975) gave 4 mg loperamide per day (1 tablet twice a day) for 24 hours
- Dreverman (1995a) gave 1mg loperamide (2 tablets) initially then up to 3.5mg (7 tablets) per day for 3 days
- Dreverman (1995b) gave 2mg loperamide (2 tablets) initially then up to 7mg (7 tablets) per day for 3 days
- Dettmer (1994a) gave loperamide in a slow release formulation, 2 mg initially (2 tablets) then up to 8 mg (8 tablets) per day for 3 days
- Dettmer (1994b) gave loperamide in a slow release formulation, 4 mg initially (2 tablets) then up to 16 mg (8 tablets) per day for 3 days.

The recommended dose is 4 mg initially (2 tablets) followed by 2 mg (1 tablet) after each loose stool for up to 5 days.

We noted that all of the loperamide studies were funded by Janssen Pharmaceutica, manufacturers of Imodium (i.e. loperamide).

A1. Co-phenotrope versus placebo

i. Number of patients with improvement in bowel habit

One study (Amery 1975) reported the number of patients without recurrence (unformed stools) after 1, 2, 4 and 24 hours. There was no significant difference between diphenoxylate and placebo at any duration.

Figure 1: Acute diarrhoea – co-phenotrope
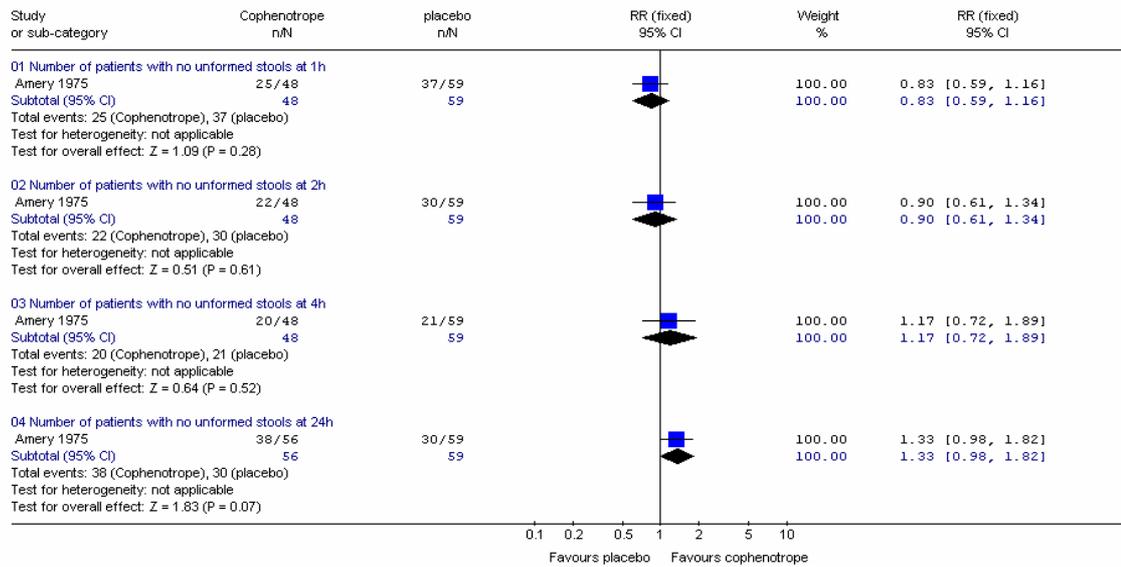
ii. **Stool frequency**

Harford (1980) recorded stool frequency averaged over 3 days of co-phenotrope or placebo in a crossover study in 15 patients; individual patient data were given. Results are reported for all patients and for the IBS subgroup separately. For all patients, there was a statistically significant difference in stool frequency of -2.29 stools per day (95%CI -4.47, -0.11), favouring co-phenotrope. For the IBS subgroup, the effect was not statistically significant, but the confidence interval was wide.

Figure 2: Stool frequency
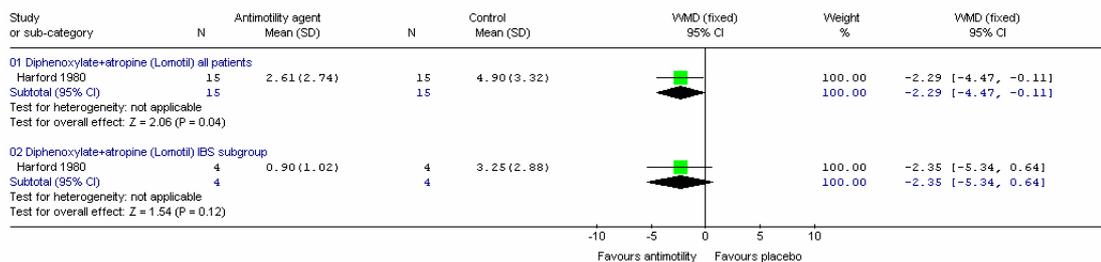
Lustman (1987) reported the change in the median number of bowel actions (24 hours prior to start of therapy minus first 24 hours of treatment). This fell from 5 in each group to 3 on diphenoxylate, compared to 4 on placebo, which was a statistically significant difference (p=0.046).

iii. Time to recurrence of unformed stools

Amery (1975) also reported the median time to first recurrence of unformed stools. This was 2 hours for both diphenoxylate and placebo groups, i.e. no significant difference (p=0.48).

Lustman (1987) reported the median time to the last loose or watery stool, which was 25 hours on diphenoxylate compared with 30 hours for placebo (not statistically significant).

Lustman (1987) also reported the median time (hours) to the start of an interval of at least 12 hours between bowel actions: 14 hours for diphenoxylate vs. 24 hours for placebo, p=0.025.

iv. Stool weight

Harford (1980) recorded stool weight averaged over 3 days of co-phenotrope or placebo in a crossover study in 15 patients. The results are reported for all patients and for the IBS subgroup separately. The confidence interval was too wide to determine if there was a difference for all patients and there was no statistically significant difference for the IBS patients.

Figure 3: Stool weight
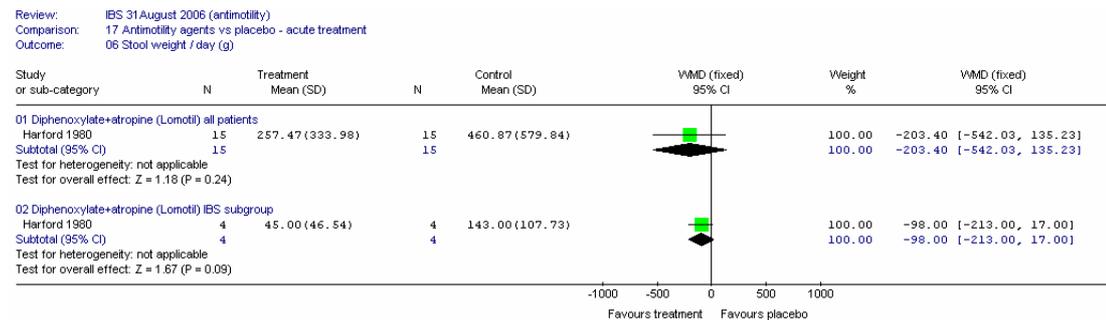
A2. Loperamide versus placebo

i. Number of patients with improvement in bowel habit

Amery (1975) reported the number of patients without recurrence (unformed stools) after 1, 2, 4 and 24 hours. Dettmer (1994) also recorded the number of patients without diarrhoea at 3 days for the two loperamide doses combined.

At 1 and 2 hours Amery (1975) showed no significant difference between loperamide and placebo. After 4 hours there were significantly fewer patients with unformed stools in the loperamide group compared to placebo, with a number needed to treat of 5 (95%CI 3, 17), for a control group risk of 36%. After 24 hours, loperamide showed fewer patients with unformed stools, with borderline significance, but the confidence interval was fairly wide. At 72 hours, Dettmar (1994) showed a small statistically significant difference between loperamide and placebo, with an NNT of 7 (95%CI 4, 34), for a control group risk of 73%.

Figure 4: Acute diarrhoea – loperamide

Review: IBS 31 August 2006 (antimotility)
Comparison: 17 Antimotility agents vs placebo - acute treatment
Outcome: 10 Number of patients with no unformed stools (acute diarrhoea) loperamide

Study or sub-category	Loperamide n/N	placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Number of patients with no unformed stools at 1h					
Amery 1975	44/56	37/59		100.00	1.25 [0.99, 1.59]
Subtotal (95% CI)	56	59		100.00	1.25 [0.99, 1.59]
Total events: 44 (Loperamide), 37 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.84 (P = 0.07)					
02 Number of patients with no unformed stools at 2h					
Amery 1975	38/56	30/59		100.00	1.33 [0.98, 1.82]
Subtotal (95% CI)	56	59		100.00	1.33 [0.98, 1.82]
Total events: 38 (Loperamide), 30 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.83 (P = 0.07)					
03 Number of patients with no unformed stools at 4h					
Amery 1975	33/56	21/59		100.00	1.66 [1.10, 2.49]
Subtotal (95% CI)	56	59		100.00	1.66 [1.10, 2.49]
Total events: 33 (Loperamide), 21 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.43 (P = 0.02)					
04 Number of patients with no unformed stools at 24h					
Amery 1975	23/56	14/59		100.00	1.73 [0.99, 3.01]
Subtotal (95% CI)	56	59		100.00	1.73 [0.99, 3.01]
Total events: 23 (Loperamide), 14 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.94 (P = 0.05)					
05 Number of patients with no diarrhoea at 72 hours					
Dettmar 1998	125/142	52/71		100.00	1.20 [1.03, 1.40]
Subtotal (95% CI)	142	71		100.00	1.20 [1.03, 1.40]
Total events: 125 (Loperamide), 52 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.35 (P = 0.02)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours loperamide

Dreverman (1995) reported the number of patients achieving first relief (the start of a 24 hour period during which no more than 1 pasty stool and no watery or loose stools were passed). There were statistically significantly more patients achieving relief for each dose of loperamide compared with placebo, but the confidence intervals are wide.

Figure 5

Review: IBS 31 August 2006 (Version 03)
Comparison: 17 Antimotility agents vs placebo - acute treatment
Outcome: 11 Number of patients achieving first relief

Study or sub-category	Loperamide n/N	Control n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 0.5 mg loperamide					
Dreverman 1995	11/76	3/78		100.00	4.23 [1.13, 15.82]
Subtotal (95% CI)	76	78		100.00	4.23 [1.13, 15.82]
Total events: 11 (Loperamide), 3 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.14 (P = 0.03)					
02 1 mg loperamide					
Dreverman 1995	16/80	3/78		100.00	6.25 [1.74, 22.42]
Subtotal (95% CI)	80	78		100.00	6.25 [1.74, 22.42]
Total events: 16 (Loperamide), 3 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.81 (P = 0.005)					

0.1 0.2 0.5 1 2 5 10
Favours control   Favours loperamide

ii. Time taken to first relief of symptoms

Two studies reported the median time to first relief of symptoms, but different doses were used.

- 0.5 mg versus placebo

  Dreverman (1995) in 156 patients: 20 hours 15 minutes for loperamide 0.5mg versus 24

hours 50 minutes for placebo (p=0.012), i.e. statistically significantly in favour of loperamide

- 1 mg versus placebo (2 studies)
  - ○ Dettmer (1994): 22.4 hours for loperamide 1mg versus 30 hours for placebo
  - ○ Dreverman (1995) in 158 patients: 15 hours 30 minutes for loperamide 1mg versus 24 hours 50 minutes for placebo (p=0.003), i.e. highly statistically significant, in favour of loperamide
- 2 mg versus placebo
  - ○ Dettmer (1994): 22.1 hours for loperamide 1mg versus 30 hours for placebo.

In Dreverman (1995), the median time to complete relief of symptoms was 26 hours 30 minutes for loperamide 1mg; 25 hours 40 minutes for loperamide 0.5mg; and 34 hours 15 minutes for placebo (p=0.041 for 0.5 mg and 0.044 for 1mg).

### iii. Time to first recurrence

Amery (1975) also reported the median time to first recurrence of unformed stools. For the loperamide and placebo groups, this was 24, and 2 hours respectively. The p value for the difference between loperamide and placebo was p=0.016, i.e. statistically significant.

B. Anti-motility agent dose 1 versus dose 2

i. Number of patients with improvement in bowel habit

Dreverman (1995) reported the number of patients achieving first relief. There was no significant difference between the doses and the confidence interval was fairly wide.

Figure 6

Review: IBS MW 080507 (Version 03)
Comparison: 23 Antimotility agent dose 1 vs. dose 2
Outcome: 03 First relief after first dose

Study or sub-category	Loperamide 0.5mg n/N	Loperamide 1mg n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
Dreverman 1 vs .5mg	11/76	16/80		100.00	0.68 [0.29, 1.57]
Total (95% CI)	76	80		100.00	0.68 [0.29, 1.57]

Total events: 11 (Loperamide 0.5mg), 16 (Loperamide 1mg)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.91 (P = 0.36)

0.1 0.2 0.5 1 2 5 10
Favours treatment   Favours control

ii. Median time to first relief of symptoms

- 0.5 mg versus 1 mg
  Dreverman (1995): 20 hours 15 minutes for loperamide 0.5mg versus 15 hours 30 minutes for loperamide 1mg
- Dettmer (1994): 22.4 hours for loperamide 1mg versus 22.1 hours for loperamide

In Dettmar (1994), the median time to complete relief of symptoms was 27.55 hours for loperamide 1mg and 25.00 hours for loperamide 2mg, with no significant difference between dose groups.

C. Anti-motility agent type 1 versus type 2

C1. Co-phenotrope versus loperamide

Four studies compared co-phenotrope and loperamide for acute episodes (indirect population) (Amery 1975; Cornett 1977; Dom 1974; Jaffe 1977). Although different doses were compared across studies, the proportions of drugs were the same.

- Amery (1975): 5 mg diphenoxylate per day (1 tablet twice a day) versus loperamide 4 mg (2 mg twice a day) for 24 hours.
- Cornett (1977): initially 5mg (2 capsules) co-phenotrope then up to 20 mg per day versus 4 mg (2 capsules) loperamide initially, then up to 16 mg per day
- Dom (1974): initially 5 mg (2 capsules) co-phenotrope then up to 25 mg per day versus 4 mg loperamide intitially (2 capsules) then up to 20 mg per day
- Jaffe (1977): initially 10 mg (4 capsules) co-phenotrope then 20 mg per day (2 capsules x 4) versus 4 mg (2 capsules) loperamide initially, then up to 16 mg per day.

We noted that Amery (1975) and Cornett (1977) were funded by Janssen Pharmaceutica, manufacturers of Imodium (i.e. loperamide).

i. Number of patients with improvement in bowel habit

Three studies reported the number of patients with no unformed stools at different durations: Amery (1975) reported the number of patients without recurrence (no unformed stools) after 1, 2, 4 and 24 hours; Dom (1974) recorded the same outcome after 24 hours and 2 and 3 days; Cornett (1977) recorded the number of patients with unformed stools, from which we calculated the number with no unformed stools after 24 hours and 2 and 3 days. Jaffe (1977) reported the number of patients not reaching a 'cure' after 4 days (figure 10). At 1 and 2 hours there were statistically significantly more patients who were diarrhoea free for the loperamide group compared to the diphenoxylate. NNTs are 4 (95%CI 3, 12) and 5 (9%CI 3, 34). There was no significant difference at 4 hours. At 24 hours, meta-analysis of three studies showed statistically significantly more patients had no unformed stools for the loperamide group; RR 0.78 (95%CI 0.62, 0.98); with some heterogeneity (I^2=47%, p=0.15). The NNT was 20 (95%CI 10, 100) for a control group rate of 21 to 41%. At 2 and 3 days there was still a statistically significant effect favouring loperamide, but there was significant heterogeneity for the 3 day meta-analysis. We note, however, that some of these trials were industry sponsored.

Figure 7: Number of patients with no unformed stools

Study or sub-category	Diphenoxylate n/N	loperamide n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 1h					
Amery 1975	25/48	44/56		100.00	0.66 [0.49, 0.90]
Subtotal (95% CI)	48	56		100.00	0.66 [0.49, 0.90]
Total events: 25 (Diphenoxylate), 44 (loperamide)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.65 (P = 0.008)					
02 2h					
Amery 1975	22/48	38/56		100.00	0.68 [0.47, 0.96]
Subtotal (95% CI)	48	56		100.00	0.68 [0.47, 0.96]
Total events: 22 (Diphenoxylate), 38 (loperamide)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.16 (P = 0.03)					
03 4h					
Amery 1975	20/48	33/56		100.00	0.71 [0.47, 1.05]
Subtotal (95% CI)	48	56		100.00	0.71 [0.47, 1.05]
Total events: 20 (Diphenoxylate), 33 (loperamide)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.70 (P = 0.09)					
04 24h					
Amery 1975	11/56	23/56		17.49	0.48 [0.26, 0.89]
Cornett 1977	46/181	42/159		34.00	0.96 [0.67, 1.38]
Dom 1974	49/311	63/303		48.52	0.76 [0.54, 1.06]
Subtotal (95% CI)	548	518		100.00	0.78 [0.62, 0.98]
Total events: 106 (Diphenoxylate), 128 (loperamide)					
Test for heterogeneity: Chi² = 3.76, df = 2 (P = 0.15), I² = 46.8%					
Test for overall effect: Z = 2.16 (P = 0.03)					
05 2 days					
Cornett 1977	97/181	105/159		34.13	0.81 [0.68, 0.97]
Dom 1974	176/311	213/303		65.87	0.81 [0.71, 0.91]
Subtotal (95% CI)	492	462		100.00	0.81 [0.73, 0.89]
Total events: 273 (Diphenoxylate), 318 (loperamide)					
Test for heterogeneity: Chi² = 0.01, df = 1 (P = 0.94), I² = 0%					
Test for overall effect: Z = 4.19 (P < 0.0001)					
06 3 days					
Cornett 1977	126/181	134/159		33.31	0.83 [0.73, 0.93]
Dom 1974	267/311	282/303		66.69	0.92 [0.87, 0.97]
Subtotal (95% CI)	492	462		100.00	0.89 [0.84, 0.94]
Total events: 393 (Diphenoxylate), 416 (loperamide)					
Test for heterogeneity: Chi² = 3.19, df = 1 (P = 0.07), I² = 68.6%					
Test for overall effect: Z = 4.31 (P < 0.0001)					
07 4 days					
Jaffe 1977	38/42	37/41		100.00	1.00 [0.87, 1.15]
Subtotal (95% CI)	42	41		100.00	1.00 [0.87, 1.15]
Total events: 38 (Diphenoxylate), 37 (loperamide)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.04 (P = 0.97)					

0.1 0.2 0.5 1 2 5 10

Favours loperamide   Favours diphenoxylat

ii. Time taken to first stools

Jaffe (1977) reported the mean time to the 1st, 2nd, and 3rd motion. In all cases the confidence intervals were wide.

Figure 8

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	7.90 (20.50)	41	5.20 (11.50)		100.00	2.70 [-4.43, 9.83]
Total (95% CI)	42		41			100.00	2.70 [-4.43, 9.83]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.74 (P = 0.46)							

-10 -5 0 5 10

Favours loperamide   Favours diphenoxylat

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	17.90(28.20)	41	16.20(27.40)		100.00	1.70 [-10.26, 13.66]
Total (95% CI)	42		41			100.00	1.70 [-10.26, 13.66]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.28 (P = 0.78)

-10 -5 0 5 10
Favours loperamide   Favours diphenoxylat

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	29.00(32.10)	41	24.40(28.50)		100.00	4.60 [-8.45, 17.65]
Total (95% CI)	42		41			100.00	4.60 [-8.45, 17.65]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.69 (P = 0.49)

-10 -5 0 5 10
Favours loperamide   Favours diphenoxylat

Amery (1975) reported the median time to first recurrence of unformed stools. For the loperamide and diphenoxylate groups, this was 24 and 2 hours respectively. This was a statistically significant difference (p=0.024), favouring loperamide. We note, however, that this was an industry funded study.

iii. Stool frequency

Jaffe (1977) reported the stool frequency in the first 6 hours; first 12 hours; first 24 hours; and first 48 hours following the first dose of the medication. In all cases the confidence intervals were wide, so conclusions were not drawn.

Figure 9

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	1.70(1.50)	41	1.90(1.20)		100.00	-0.20 [-0.78, 0.38]
Total (95% CI)	42		41			100.00	-0.20 [-0.78, 0.38]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.67 (P = 0.50)

-1 -0.5 0 0.5 1
Favours treatment   Favours control

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	2.40(1.60)	41	2.60(1.60)		100.00	-0.20 [-0.89, 0.49]
Total (95% CI)	42		41			100.00	-0.20 [-0.89, 0.49]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.57 (P = 0.57)

-1 -0.5 0 0.5 1
Favours treatment   Favours control

Study or sub-category	N	Diphenoxylate Mean (SD)	N	Loperamide Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Jaffe 1977	42	3.70(2.10)	41	4.00(2.20)		100.00	-0.30 [-1.23, 0.63]
Total (95% CI)	42		41			100.00	-0.30 [-1.23, 0.63]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.64 (P = 0.53)

-1 -0.5 0 0.5 1
Favours treatment   Favours control

```
Review:          IBS MW 080507 (Version 03)
Comparison:      18 Co-phenotrope vs loperamide - acute treatment ( (not necessarily IBS)
Outcome:         16 Mean number of motions in first 48 hours

Study                          Diphenoxylate              Loperamide            WMD (fixed)          Weight        WMD (fixed)
or sub-category        N        Mean (SD)         N        Mean (SD)             95% CI               %             95% CI

 Jaffe 1977           42       5.50(3.40)        41       6.30(3.50)              [square]           100.00      -0.80 [-2.28, 0.68]

Total (95% CI)        42                         41                             [diamond]           100.00      -0.80 [-2.28, 0.68]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.06 (P = 0.29)
                                                                    -4   -2    0    2    4
                                                                 Favours treatment  Favours control
```

Cornett (1977) also reported the frequency of unformed stools over 72 hours. From a mean of 9.14 unformed stools before treatment, patients on diphenoxylate reduced to 5.47 (a fall of 3.67) while those on loperamide fell from 8.75 to 4.38 (a fall of 4.37). No standard deviations were given.

Dom (1974) reported the change in the mean number of unformed stools in a 72 hour period (from 8.06 before treatment with diphenoxylate and 8.02 in the loperamide group to 3.68 and 2.65 respectively, i.e. a fall of 4.38 and 5.37 respectively, this was statistically significantly in favour of loperamide (p=0.011).

iv. Time to cure

Jaffe (1977) reported the mean time to cure. There was a wide confidence interval so conclusions were not drawn.

Figure 10


```
Review:          IBS MW 080507 (Version 03)
Comparison:      18 Co-phenotrope vs loperamide - acute treatment ( (not necessarily IBS)
Outcome:         17 Mean time to cure

Study                          Diphenoxylate              Loperamide            WMD (fixed)          Weight        WMD (fixed)
or sub-category        N        Mean (SD)         N        Mean (SD)             95% CI               %             95% CI

 Jaffe 1977           42       21.60(20.40)      41       26.30(26.00)        [square]             100.00      -4.70 [-14.77, 5.37]

Total (95% CI)        42                         41                          [diamond]            100.00      -4.70 [-14.77, 5.37]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.91 (P = 0.36)
                                                                 -10   -5    0    5    10
                                                                 Favours treatment  Favours control
```

C2. Co-phenotrope versus morphine

i. Number of patients with normal stool frequency

One study (Lee 1968) compared cophenotrope-with-neomycin to kaolin-and-morphine. This study reported the number of patients with abnormal stool frequency after 12 hours, 24 hours, 2 days, 3 days and 4 days. We have reported the number of patients with normal stool frequency. There was a statistically significant difference favouring co-phenotrope at durations up to 48 hours. At 12 hours the NNT was 4 (95%CI 3, 6), for a control group risk of 14%.

Figure 11
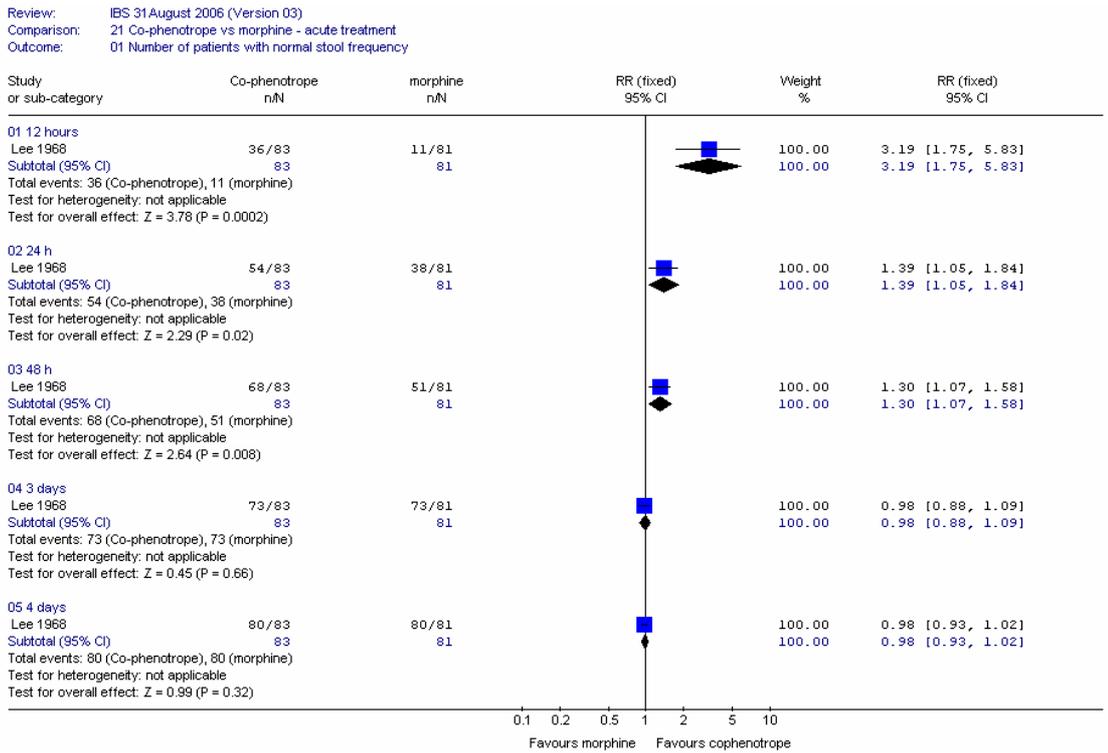
ii. Number of patients with normal stool consistency

One study (Lee 1968) compared cophenotrope-with-neomycin to kaolin-and-morphine. This study reported the number of patients with abnormal stool consistency after 12 hours, 24 hours, 2 days, 3 days and 4 days. We have reported this as the number of patients with normal stool consistency. There was a statistically significant effect, favouring co-phenotrope at 12 hours (RR 3.49 (95%CI 1.60, 7.60); NNT 5 (95%CI 4, 10), for a control group risk of 9%. Otherwise there was no significant difference between interventions.

Figure 12

Study or sub-category	Co-phenotrope n/N	morphine n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 12 hours					
Lee 1968	25/83	7/81		100.00	3.49 [1.60, 7.60]
Subtotal (95% CI)	83	81		100.00	3.49 [1.60, 7.60]
Total events: 25 (Co-phenotrope), 7 (morphine)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.14 (P = 0.002)					
02 24 h					
Lee 1968	42/83	31/81		100.00	1.32 [0.93, 1.87]
Subtotal (95% CI)	83	81		100.00	1.32 [0.93, 1.87]
Total events: 42 (Co-phenotrope), 31 (morphine)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.57 (P = 0.12)					
03 48 h					
Lee 1968	56/83	45/81		100.00	1.21 [0.95, 1.55]
Subtotal (95% CI)	83	81		100.00	1.21 [0.95, 1.55]
Total events: 56 (Co-phenotrope), 45 (morphine)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.55 (P = 0.12)					
04 3 days					
Lee 1968	64/83	61/81		100.00	1.02 [0.86, 1.22]
Subtotal (95% CI)	83	81		100.00	1.02 [0.86, 1.22]
Total events: 64 (Co-phenotrope), 61 (morphine)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.27 (P = 0.79)					
05 4 days					
Lee 1968	79/83	71/81		100.00	1.09 [0.99, 1.19]
Subtotal (95% CI)	83	81		100.00	1.09 [0.99, 1.19]
Total events: 79 (Co-phenotrope), 71 (morphine)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.70 (P = 0.09)					

0.1   0.2   0.5   1   2   5   10

Favours morphine   Favours cophenotrope

II. MAINTENANCE TREATMENT FOR CHRONIC DIARRHOEA

A. Antimotility agent versus placebo

Three studies gave loperamide for maintenance treatment (Efskind 1995; Hovdenak 1987; Lavö 1987). All were in patients with IBS, although Hovdenak (1987) stated that the patients had different types of IBS, including IBS-C. Subgroup analyses were presented, but these did not constitute stratification before randomisation. This study also had a duration of 3 weeks, but the GDG agreed that this was acceptable. We noted that Efskind (1995) was industry supported (by Janssen Pharmaceutica, the manufacturers of loperamide).

The dose in Hovdenak (1987) was fixed at 4 mg at night, whilst the patients were able to adjust the dose in the other two studies. Both started the patients on 2 mg at night, which was increased to 6 mg (Efskind 1995) or 8 mg (Lavö 1987) as required.

1. Global symptoms

a) Number of patients with improvement in global symptoms

One study (Hovdenak 1987) reported the number of patients with improvement in global symptoms after 3 weeks. The study reported results for three of the different IBS subgroups (IBS-C results were not reported). The exclusion of the IBS-C results breaks the randomisation, so these are post-hoc subgroups, but still gives some information. We have grouped together the results for IBS-D, IBS-A (with pain), IBS-A (without pain), and the separate subgroups. There was a statistically significant improvement in symptom score for

the 3-type IBS group and for the IBD-D group alone, although the confidence interval for the latter was wide. The relative risks were 2.00 (95%CI 1.15, 3.48) and 4.00 (1.20, 13.28) respectively. These corresponded to a number needed to treat of 3 (95%CI 2, 8) and 2 (95%CI 1, 3), for control group rates of 39% and 25% respectively.

Figure 13

Review: IBS
Comparison: 19 Antimotility agents vs placebo - maintenance treatment
Outcome: 01 Number of patients with global improvement in symptoms

Study or sub-category	Treatment n/N	placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Loperamide (IBS-D + IBS-A (+pain) + IBS-A (no pain))					
Hovdenak	18/23	9/23		100.00	2.00 [1.15, 3.48]
Subtotal (95% CI)	23	23		100.00	2.00 [1.15, 3.48]
Total events: 18 (Treatment), 9 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.46 (P = 0.01)					
02 Loperamide (IBS-D subgroup)					
Hovdenak	8/8	2/8		100.00	4.00 [1.20, 13.28]
Subtotal (95% CI)	8	8		100.00	4.00 [1.20, 13.28]
Total events: 8 (Treatment), 2 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.26 (P = 0.02)					
03 Loperamide (IBS-A subgroup)					
Hovdenak	10/14	7/18		100.00	1.84 [0.94, 3.58]
Subtotal (95% CI)	14	18		100.00	1.84 [0.94, 3.58]
Total events: 10 (Treatment), 7 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.79 (P = 0.07)					
04 Loperamide (IBS-A without pain subgroup)					
Hovdenak	4/5	2/7		100.00	2.80 [0.80, 9.78]
Subtotal (95% CI)	5	7		100.00	2.80 [0.80, 9.78]
Total events: 4 (Treatment), 2 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.61 (P = 0.11)					
05 Loperamide (IBS-A pain subgroup)					
Hovdenak	6/10	5/11		100.00	1.32 [0.58, 3.00]
Subtotal (95% CI)	10	11		100.00	1.32 [0.58, 3.00]
Total events: 6 (Treatment), 5 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.66 (P = 0.51)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours treatment

b) Global symptom improvement score

One study (Lavö 1987) reported a subjective overall response for the whole 13 week study period, which was said to statistically significantly favour loperamide (p<0.03).

2. Individual symptoms

a) Pain

i. Number of patients with less pain

Two studies reported the number of patients with less pain, one (Hovdenak 1987) after 3 weeks in a range of different IBS subgroups; and the other (Lavö 1987) after 13 weeks in patients with IBS-D only. For the former study we grouped together the results for IBS-D, IBS-A (with pain), IBS-A (without pain) and also reported these separately.

The confidence intervals were wide for each study. Meta-analysis of the 3-type IBS group with the Lavö study, in 70 patients gave a statistically significant difference between interventions, favouring loperamide; RR 2.60 (95%CI 1.02, 6.61) but there was some heterogeneity (I^2=48%, p=0.17). The NNT was 5 (95%CI 3, 25), for a control group rate of 8 to 17%.

Figure 14: Number of patients with less pain

Review: IBS
Comparison: 19 Antimotility agents vs placebo - maintenance treatment
Outcome: 02 Number of patients with less pain

Study or sub-category	Treatment n/N	placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Loperamide (IBS-D + IBS-A (+pain) + IBS-A (no pain))					
Hovdenak	6/23	4/23		80.00	1.50 [0.49, 4.62]
Lavo	7/12	1/12		20.00	7.00 [1.01, 48.54]
Subtotal (95% CI)	35	35		100.00	2.60 [1.02, 6.61]
Total events: 13 (Treatment), 5 (placebo)					
Test for heterogeneity: Chi² = 1.92, df = 1 (P = 0.17), I² = 48.0%					
Test for overall effect: Z = 2.01 (P = 0.04)					
02 Loperamide (IBS-D subgroup)					
Hovdenak	0/8	0/8			Not estimable
Lavo	7/12	1/12		100.00	7.00 [1.01, 48.54]
Subtotal (95% CI)	20	20		100.00	7.00 [1.01, 48.54]
Total events: 7 (Treatment), 1 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.97 (P = 0.05)					
03 Loperamide (IBS-A subgroup)					
Hovdenak	6/14	4/18		100.00	1.93 [0.67, 5.54]
Subtotal (95% CI)	14	18		100.00	1.93 [0.67, 5.54]
Total events: 6 (Treatment), 4 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.22 (P = 0.22)					
04 Loperamide (IBS-A without pain subgroup)					
Hovdenak	0/5	0/7			Not estimable
Subtotal (95% CI)	5	7			Not estimable
Total events: 0 (Treatment), 0 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: not applicable					
05 Loperamide (IBS-A pain subgroup)					
Hovdenak	6/10	4/11		100.00	1.65 [0.65, 4.19]
Subtotal (95% CI)	10	11		100.00	1.65 [0.65, 4.19]
Total events: 6 (Treatment), 4 (placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.05 (P = 0.29)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours treatment

ii. Number of patients with increased pain

Two studies reported the number of patients with more pain, one (Hovdenak 1987) after 3 weeks in a range of different IBS subgroups; and the other (Lavö 1987) after 13 weeks in patients with IBS-D only. For the former study we grouped together the results for IBS-D, IBS-A (with pain), IBS-A (without pain) and also reported these separately.

The confidence intervals were wide for each study. Meta-analysis of the 3-type IBS group with the Lavö (1987) study, in 70 patients, gave a statistically significant difference between interventions, favouring loperamide; RR 0.38 (95%CI 0.15, 0.96) and there was no heterogeneity (I²=0%), although the confidence interval was wide. The NNT was 5 (95%CI 3, 25), for a control group risk of 22 to 67%. For the meta-analysis of 40 patients with IBS-D, the difference was still statistically significant with no heterogeneity; RR 0.36 (95%CI 0.14, 0.96), i.e. 3 times more risk of pain for the placebo group (Figure 14). The NNT was 3 (95%CI 2, 13) for a control group risk of 38 to 67%, but the confidence interval was wide.

Figure 15: Number of patients with more pain


```
Review:         IBS
Comparison:     19 Antimotility agents vs placebo - maintenance treatment
Outcome:        03 Number of patients with more pain

Study              Treatment      placebo         RR (fixed)        Weight        RR (fixed)
or sub-category      n/N            n/N             95% CI            %             95% CI

01 Loperamide (IBS-D + IBS-A (+pain) + IBS-A (no pain))
Hovdenak            3/23           5/23                            38.46       0.60 [0.16, 2.22]
Lavo                2/12           8/12                            61.54       0.25 [0.07, 0.94]
Subtotal (95% CI)    35             35                            100.00      0.38 [0.15, 0.96]
Total events: 5 (Treatment), 13 (placebo)
Test for heterogeneity: Chi² = 0.85, df = 1 (P = 0.36), I² = 0%
Test for overall effect: Z = 2.06 (P = 0.04)

02 Loperamide (IBS-D subgroup)
Hovdenak            2/8            3/8                             27.27       0.67 [0.15, 2.98]
Lavo                2/12           8/12                            72.73       0.25 [0.07, 0.94]
Subtotal (95% CI)    20             20                            100.00      0.36 [0.14, 0.96]
Total events: 4 (Treatment), 11 (placebo)
Test for heterogeneity: Chi² = 0.94, df = 1 (P = 0.33), I² = 0%
Test for overall effect: Z = 2.04 (P = 0.04)

03 Loperamide (IBS-A subgroup)
Hovdenak            1/14           2/18                            100.00      0.64 [0.06, 6.39]
Subtotal (95% CI)    14             18                            100.00      0.64 [0.06, 6.39]
Total events: 1 (Treatment), 2 (placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.38 (P = 0.71)

04 Loperamide (IBS-A without pain subgroup)
Hovdenak            0/5            2/7                             100.00      0.27 [0.02, 4.59]
Subtotal (95% CI)     5             7                             100.00      0.27 [0.02, 4.59]
Total events: 0 (Treatment), 2 (placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.91 (P = 0.36)

05 Loperamide (IBS-A pain subgroup)
Hovdenak            1/10           0/11                            100.00      3.27 [0.15, 72.23]
Subtotal (95% CI)    10             11                            100.00      3.27 [0.15, 72.23]
Total events: 1 (Treatment), 0 (placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.75 (P = 0.45)

                                    0.1  0.2  0.5  1  2   5  10
                                  Favours treatment  Favours placebo
```

iii. Number of days with pain

One study (Hovdenak 1987) recorded the number of days with pain over 3 weeks. In the subgroup of IBS-A with pain, there was a statistically significant mean difference of 6.1 days between loperamide and placebo ($p<0.01$), favouring the former.

iv) Pain score

One study (Efskind 1996) in 69 patients reported no significant difference between interventions in the change from baseline in abdominal pain measured on a visual analogue scale. This was the case for both their 2 week dose adjustment period and for 5 weeks of fixed dose loperamide or placebo. However, the study found significantly more pain at night for the loperamide group ($p<0.05$ for both periods). In contrast, another study in 21 patients (Lavö 1987) found a statistically significant difference in pain score, favouring loperamide ($p<0.05$).

b) Bowel habit

i. Number of patients with improved bowel habit (frequency)

Two studies reported the number of patients with improved stool frequency, one (Hovdenak 1987) after 3 weeks in different IBS subgroups; and the other (Lavö 1987) after 13 weeks in patients with IBS-D only. For the former study we grouped together the results for IBS-D, IBS-A (with pain), IBS-A (without pain) and also reported these separately.

Meta-analysis of the 3-type IBS group with the Lavö study, in 70 patients gave a statistically significant difference between interventions, favouring loperamide; RR 2.38 (95%CI 1.53, 3.70) and there was no heterogeneity (I^2=0%). This corresponded to an NNT of 2 (95%CI 2, 4) for a control group risk of 25-43%. For the meta-analysis of 40 patients with IBS-D, the difference was still statistically significant with no heterogeneity; RR 2.83 (95%CI 1.43, 5.63). This corresponded to a number needed to treat of 2 (95%CI 2, 4) for a control group risk of 25-38% (figure 15). Statistically significant improvements were also achieved for the IBS-A group, which had an NNT of 2 (95%CI 2, 4), for a control group risk of 45%. Confidence intervals for these meta-analyses were fairly wide.

Figure 16: Improvement in stool frequency
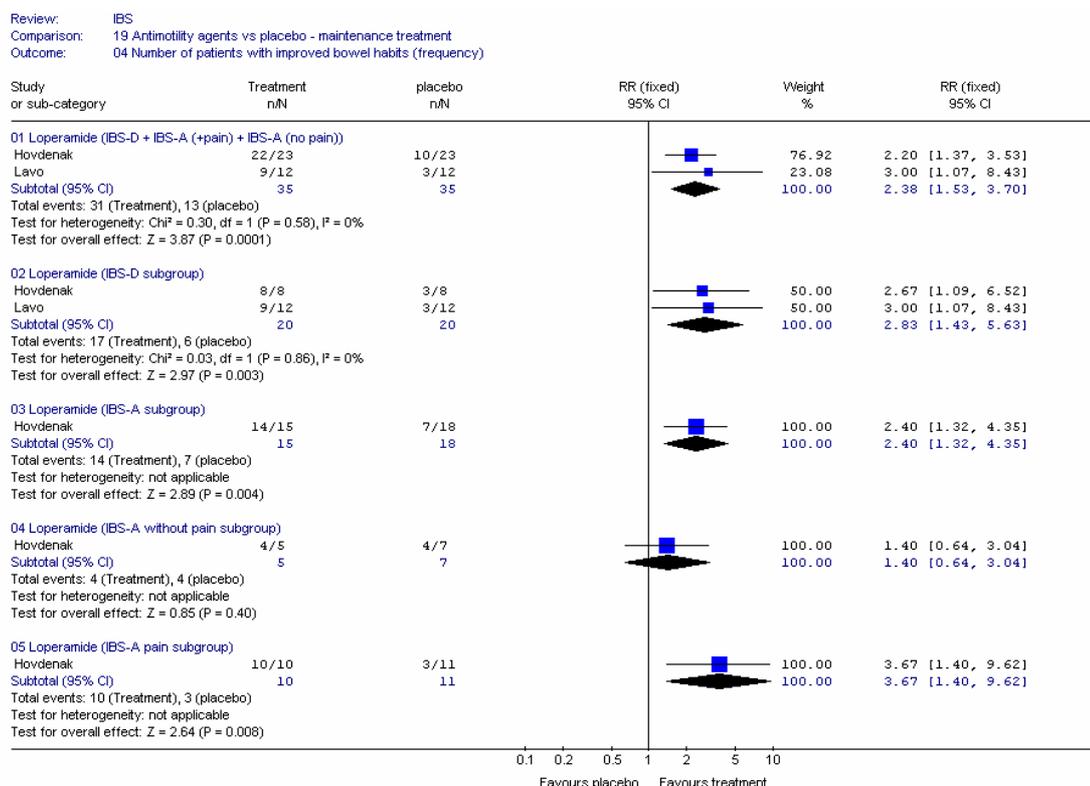
ii. Number of patients with improved bowel habit (consistency)

The same two studies reported the number of patients with improved stool consistency (Hovdenak 1987; Lavö 1987). Meta-analysis of the 3-type IBS group with the Lavö (1987) study, in 70 patients gave a statistically significant difference between interventions, favouring loperamide; but there was significant heterogeneity (I^2=69%; p=0.07). This heterogeneity may possibly be explained by differences in the study duration or the intervention dose, both of which were higher for the Lavö (1987) study. For the meta-analysis of 40 patients with IBS-D, the difference was still statistically significant with some heterogeneity (I^2=59%; p=0.12). Random effects analysis showed wide confidence intervals and neither analysis was statistically significant. Statistically significant improvements were also achieved for the

combined IBS-A group; RR: 2.10 (95%CI 1.23, 3.58), which corresponded to an NNT of 3 (95%CI 2, 5) for a control group risk of 44%.

Figure 17: Improvement in stool consistency

Review: IBS 31 August 2006 (antimotility)
Comparison: 19 Antimotility agents vs placebo - maintenance treatment
Outcome: 05 Number of patients with improved bowel habits (consistency)
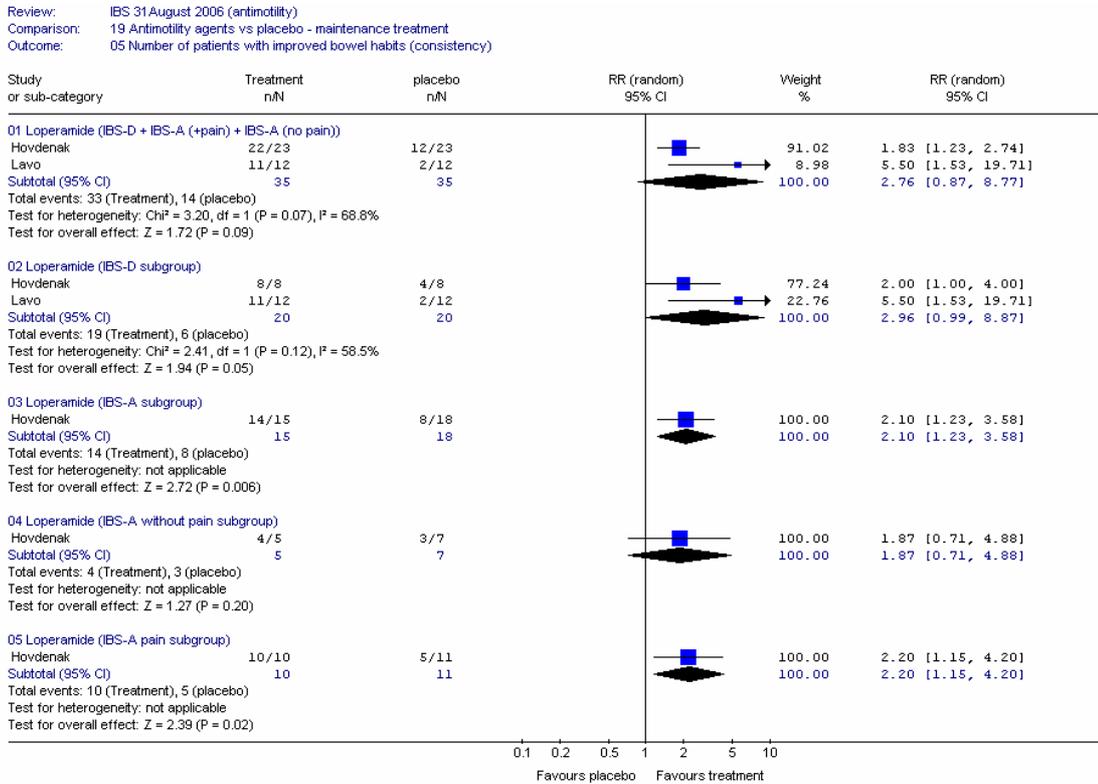
iii. Stool scores

Two studies (Efskind 1996 (n=69) and Lavö 1987 (n=21)) reported a stool score measure for both consistency and frequency. Each found a statistically significant improvement in stool consistency (p<0.002 and p<0.001) respectively, favouring loperamide. Efskind (1996) also found a statistically significant difference over 7 weeks for frequency (p<0.05) but Lavö (1987) found no significant difference. Lavö (1987) found a highly significant difference in the number of formed stools, favouring the loperamide group.

Figure 18: Number of formed stools

Review: IBS
Comparison: 19 Antimotility agents vs placebo - maintenance treatment
Outcome: 07 Number of formed stools
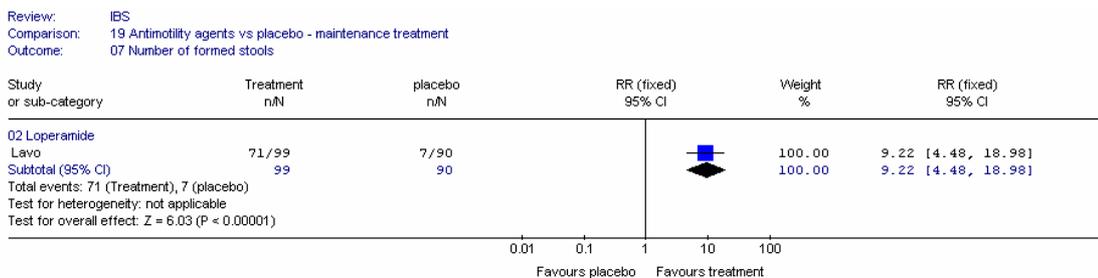
iv. Number of patients with improvement in urgency

One study (Lavö 1987) in 21 patients reported the number of patients with less urgency and found this to be significantly greater for the loperamide group, although the confidence interval

was wide; RR 3.00 (95%CI 1.07, 8.43). This corresponds to an NNT of 2 (95%CI 2, 7), for a control group rate of 25%.

Figure 19: Number of patients with less urgency
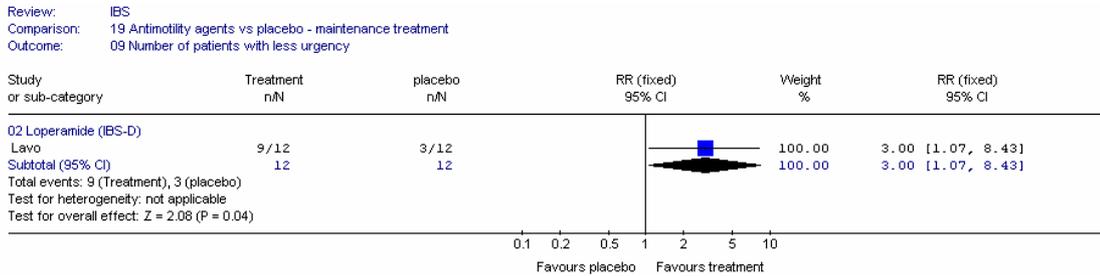
3. Adverse effects

Only one small study in 24 patients reported the number of patients with adverse effects (Lavö 1987). In all cases there are very wide confidence intervals and conclusions cannot be drawn.

Figure 20: Adverse effects
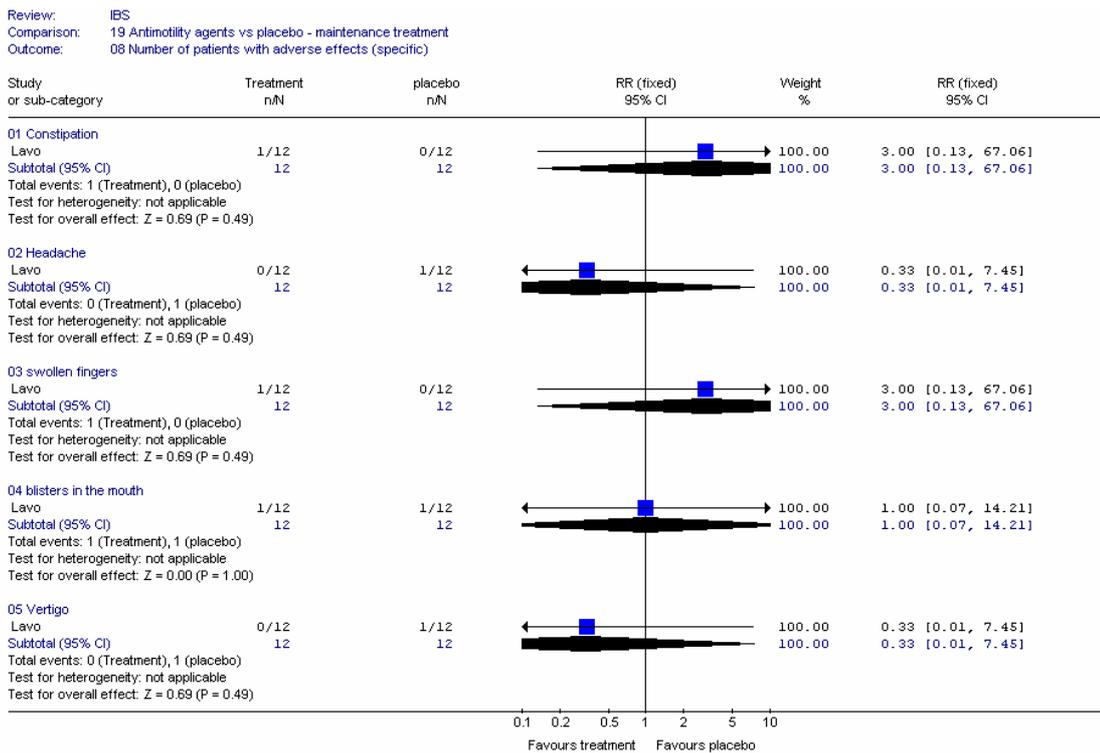
B. Anti-motility agent type 1 versus Anti-motility agent type 2

One study compared different types of anti-motility agent: Pelemans and Vantrappen (1976) compared co-phenotrope and loperamide for 14 to 49 days in a crossover trial in 23 patients. Only 4 of the patients had IBS, 18 of the rest had inflammatory bowel disease and the GDG considered this population to be too indirect to be useful, even though this was the only maintenance study considering co-phenotrope.

C. Antimotility agent versus yoga

1. Bowel symptom score

Taneja (2004) reported bowel symptom scores 1 and 2 months after starting either loperamide or yoga. Patients were asked the following questions: Is the frequency of stools >3 times per day? Is the frequency of stools <3 times per week? Are the stools lumpy or hard? Is there straining? Are the stools loose or watery? Is there urgency to pass the stools? Are the stools mucoid? For each symptom, if present >25% of the time, a positive answer scored 1 and a negative answer 0, so the range is 0-7. Both groups reduced their scores from baseline (loperamide 4.08 (SD 0.9); yoga 3.77 (SD 1.2)), but there was no significant difference between the groups.

Figure 21: Bowel symptom score
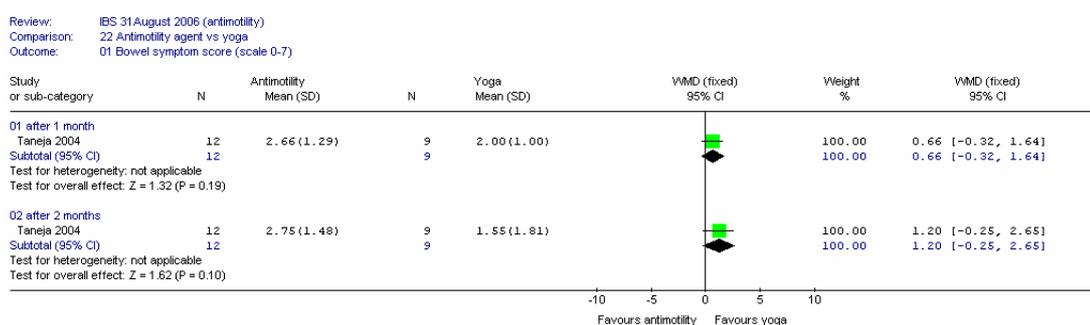
Adverse Effects

Two RCTs with adverse effects data were identified (Lavo 1987; Cann 1984). One was a crossover trial which should be treated with caution (Cann 1984). No clear trend could be identified from these two studies but abdominal cramps, dizziness, drowsiness, and skin reactions including urticaria; paralytic ileus and abdominal bloating have been reported (BNF 2007).

ECONOMIC LITERATURE FOR ANTIMOTILITY AGENTS

No relevant health economic analyses were identified on the cost-effectiveness of antimotility agents in the treatment of IBS.

COST-EFFECTIVENESS ANALYSIS FOR ANTIMOTILITY AGENTS

This section describes the health economic analysis undertaken to inform recommendations on the use of antimotility agents as a long-term maintenance therapy in IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

- Loperamide was the only intervention considered by the economic model as it was the only intervention which had evidence of clinical effectiveness for longer-term maintenance use.

- Separate efficacy estimates were available for subgroups of patients with IBS-D and patients with IBS-A with pain or IBS-A without pain. Loperamide was statistically significantly better than placebo in achieving an improvement in global symptoms for all three subgroups combined but was also statistically significantly better for the IBS-D subgroup but not for the IBS-A subgroups when considered separately. However, the confidence intervals for the subgroups were wide due to the small number of patients in each subgroup. The combined effectiveness estimate for IBS-D and IBS-A (with and without pain) was used in the basecase.

Modelled response rates

In the basecase scenario the response rate of 45% in the no treatment arm is taken from the Mearin (2004) cohort study. This represents the group of patients for whom we would expect an improvement in symptoms without any specific intervention. The RR of response for antimotility vs placebo is 2.00; therefore the response rate in the intervention arm is 90% (=45% x 2.00) giving an absolute difference in response between the intervention and no treatment arms of 45% (=90%-45%) during the 1st month.

We have also considered an alternative scenario in which no patient in the comparator arm achieves an improvement in symptoms but there is still a 45% chance of response for the intervention arm.

Table 1: Intervention specific parameters – anti-motility agents

Description	Value		Evidence
RR of response for intervention vs placebo	2.00		Meta-analysis of RCT evidence for improvement in global symptoms
Maximum number of switches considered	0		Limited by number of effective interventions
Drug costs			
Intervention	**Dose per day**	**Cost per month* (assuming lowest cost preparation)**	
Loperamide	4mg	£2.29	

* British National Formulary (Joint Formulary Committee 2007)

In the basecase analysis, treatment with loperamide for 100 patients with IBS-D or IBS-A results in an additional 1.60 QALYs but costs an additional £3,055 over a 6 month time frame, compared to no treatment, giving a cost per QALY of £1,914. The basecase analysis assumes that the intervention is used on a daily basis, meaning this estimate may be more relevant to patients with IBS-D. The cost-effectiveness of using loperamide on 25% to 75% of days is considered in the sensitivity analysis, and this estimate may be more relevant to patients with IBS-A whose primary symptom may vary from constipation to diarrhoea.

These results are an estimate of the cost-effectiveness over the first 6 months after initiating treatment with loperamide. The cost per QALY for continuing treatment with loperamide beyond 6 months is lower than the cost per QALY during the initial 6 months provided that treatment is reviewed every 6 months and discontinued in patients who no longer experience a therapeutic benefit, either due to lack of effectiveness or a change in their symptom profile.

The probabilistic sensitivity analysis considered the uncertainty in the basecase estimate of cost-effectiveness due to uncertainty in the parameters used to estimate the cost-effectiveness. The CEAC presented in Figure 22 shows the probability that the cost per QALY for loperamide compared to no treatment is under various cost per QALY thresholds. For example, it shows that there is an 83% probability that loperamide has a cost per QALY of under £5,000 compared to no treatment and a 98% probability that it has a cost per QALY under £20,000. However, it should be noted that these estimates only consider the uncertainty in cost-effectiveness due to the accuracy of several input parameters and they do not reflect general uncertainty around the assumptions made in the model. The uncertainty from these assumptions was explored in the univariate sensitivity analysis.

Figure 22. Cost-effectiveness acceptability curve for loperamide compared to no treatment
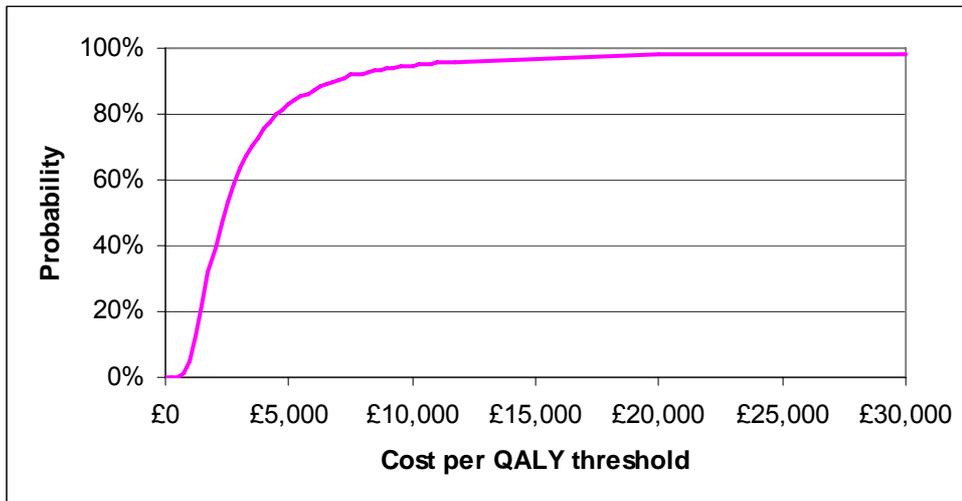
Univariate sensitivity analysis

Maintaining the 45% difference in response between the two arms but reducing the response rate in the no treatment arm from 45% to zero marginally decreased the cost per QALY to £1,593.

We carried out a threshold analysis to determine whether antimotility therapy is still cost-effective for lower gains in health related quality of life. In the basecase it is assumed that patients who respond to therapy accumulate 0.071 QALYs more per annum than patients who do not respond. For comparison, a gain of 0.135 QALYs would represent a complete remission

of IBS symptoms. If the QALY gain associated with a response to therapy was reduced to 0.006 QALYs, then the cost per QALY of providing antimotility therapy would be above £20,000 per QALY.

The basecase analysis assumes that the lowest cost preparation is prescribed. If, the highest cost preparation (Imodium syrup) is prescribed at the same dose, then the cost per QALY is increased to £3,173. If the dose is increased to the maximum dose of 16mg per day, which is double the highest dose used in the maintenance therapy trials, and the highest cost preparation is prescribed, then the monthly cost of loperamide rises to £23.85, but the cost per QALY is still well below £20,000 per QALY at £9,311.

If a patient takes a therapy on an as needed basis, it is assumed that they take the therapy on days when their quality of life is significantly affected by their IBS symptoms but not on days when their symptoms are less severe. They only accrue QALY benefits and drug costs on the days they take the therapy. However, it is still necessary to assess all patients for response after 1 month of initiating therapy. This means that it is less cost-effective to initiate therapy in patients who use the therapy on fewer days as the monitoring costs are just as high but the benefits are lower. This is shown by the cost per QALY of £5,297 for patients who take the therapy on 25% on days.

Table 2: Sensitivity results for loperamide compared to no treatment for 100 patients with IBS-D or IBS-A

Scenario	No Treatment		Intervention		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	1.60	£3,055	3.19	£1,914
No response in no treatment arm	£0	0	£2,542	1.60	£1,593
Response rate in no treatment arm from RCTs	£0	1.69	£3,117	3.38	£1,842
Treatment used 75% of days	£0	1.60	£2,741	2.79	£2,290
Treatment used 50% of days	£0	1.60	£2,428	2.39	£3,042
Treatment used on 25% of days	£0	1.60	£2,114	2.00	£5,297
Half of treatment obtained over the counter	£0	1.60	£2,428	3.19	£1,521
Higher cost formulations (same dose)	£0	1.60	£5,066	3.19	£3,173
High utility gain of 0.135	£0	3.02	£3,055	6.04	£1,011
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.006 and 0.007				

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. In studies for short-term symptom relief of diarrhoea (including people with IBS) for loperamide versus placebo there is fair evidence showing a significant number of patients without unformed stools.

2. In studies for short-term symptom relief of diarrhoea (including people with IBS) for co-phenotrope (lomotil) compared with placebo there is:
   - Limited evidence showing significant improvement in stool frequency
   - Fair evidence showing no significant difference in the number of patients without unformed stools

3. In studies for short-term symptom relief of diarrhoea (including people with IBS) for loperamide versus co-phenotrope (lomotil) there is:
   - Fair evidence showing a significant improvement in the number of patients without unformed stools.
   - Good evidence to show a significant improvement in stool score.

4. In studies for short-term symptom relief of diarrhoea (including participants with IBS) for co-phenotrope (lomotil) compared with Kaolin and Morphine there is fair evidence showing a significant improvement in stool consistency for co-phenotrope.

5. In studies for longer-term maintenance treatment of diarrhoea (including participants with IBS) for loperamide versus placebo there is:
   - Limited amount of fair quality evidence showing a highly significant improvement in global symptoms
   - Limited evidence showing significant improvement in pain and bowel habit.

6. In studies for longer-term maintenance treatment of diarrhoea (including people with IBS) for loperamide versus yoga there is a limited amount of weak evidence showing no significant difference in bowel score.

7. There is no clear evidence on adverse effects for antimotility agents.

HEALTH ECONOMIC STATEMENT

Evidence from a decision analytic model showed that loperamide is cost-effective as a long-term maintenance therapy for individuals with diarrhoea provided that response is assessed after the first month and every 6 months thereafter and treatment is discontinued in individuals for which it provides no therapeutic benefit.

GDG DISCUSSION

The GDG noted that, despite the small number of trials, loperamide is widely used and accepted as clinically effective for the treatment of diarrhoea in people with IBS. The GDG noted that it was good practice to titrate the dose of loperamide, according to symptom response, with the aim of achieving a well formed stool (Bristol Stool Form Scale type 4). In certain situations the daily dose of loperamide required may exceed 16 mg, and the GDG noted that this is an out of licence dose.

EVIDENCE TO RECOMMENDATIONS

The GDG took into consideration the clinical and cost effective evidence and their experience of the widespread use of loperamide. They noted the lack of evidence about adverse effects, but

did not consider this to be a significant factor in practice. The GDG wished to encourage primary care clinicians to advise people with IBS of the need to titrate doses and the method of doing so.

RECOMMENDATION

Loperamide should be the first choice of antimotility agent for diarrhoea in people with IBS.

RECOMMENDATION

People with IBS should be advised how to adjust their doses of laxative or antimotility agent 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 Form Scale type 4).

8.3 Antispasmodics

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the antispasmodics review and are reported below.

Types of studies

The GDG decided that the washout period for this review should be at least one week. Trials with shorter washout periods were not included in the analysis.

Types of intervention

Studies were to include the following interventions:

- Anti-muscarinic agents:
    - Atropine (synonyms: hyoscyamine)
    - Dicycloverine (synonyms: dicyclomine, diethylaminocarbethoxybicyclohexyl hydrochloride; trade name: Merbentyl®)
    - Hyoscine (synonyms: scopolamine; trade name: Buscopan®)
    - Propantheline (Trade name: Pro-Banthine®).
- Direct-action smooth muscle relaxants:
    - Alverine (Trade name: Spasmonal®)
    - Mebeverine: includes modified release and conventional drug delivery (Trade name: Colofac®)
    - Peppermint oil: includes modified release and enteric coated (Trade names: Colpermin®; Mintec®).

Studies with interventions not available in the UK were included in the review. These studies were listed, but not included in the sections on analysis or characteristics of studies.

The following comparisons were included

- Antispasmodic versus placebo (or nothing)
- Antispasmodic type 1 versus type 2
- Antispasmodic dose 1 versus dose 2
- Antispasmodic + another intervention versus the other intervention alone
- Antispasmodic modified release versus conventional delivery
- Duration of treatment 1 versus duration 2.

NB: In spite of the large placebo effect associated with IBS, comparisons with no treatment were included.

The antispasmodics review was concerned with both longer term maintenance treatment and short term symptom relief.

The GDG decided that there should be a minimum duration of treatment of four weeks for maintenance in this review. Studies of shorter durations were included but dealt with in sensitivity analyses.

Subgroup analyses

We planned to carry out subgroup analyses by type of antispasmodic (antimuscarinic and direct-acting smooth muscle relaxants); dose; mode of delivery (modified release/ conventional), and; duration of intervention.

Search strategy for identification of studies

The initial search identified a Cochrane Review (Quartero 2005, *Bulking agents, antispasmodic and antidepressant medication for the treatment of irritable bowel syndrome*). Searches were partly based on the terms in this review. 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. The search strategies are listed in Appendix B.

The titles and abstracts identified by the search strategy were assessed. One-hundred and three were identified as being potentially relevant to the review and these papers were retrieved in full. Twenty-three studies met the inclusion criteria for the review. The reference lists for each of the retrieved studies were inspected for further potential papers, but none were identified. The 80 excluded studies are listed in Appendix E, along with the reasons for exclusion

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

The Cochrane review recently published on this topic did not cover all the comparisons required for the guideline, as only comparisons with placebo were included. The interventions in the Cochrane Review also extended to drugs not licensed for use in the UK, and the subdivision by types of antispasmodic was different from that determined by the GDG. In addition, crossover trials were excluded from the Cochrane review unless first-period only results were given. We also did not agree with some of the other inclusions/exclusions. A simple update of the Cochrane review was therefore not appropriate to the needs of the guideline and the review was instead used mainly as a reference. Studies excluded from the Cochrane review as non-placebo comparisons or crossover trials were assessed for inclusion in the guideline review.

This review relates only to those studies using interventions licensed for use in the UK. The following 38 studies were therefore eliminated from the review:

- Cimetropium bromide (Centonze 1988; Dobrilla 1990; Passaretti 1989; Piai 1985; Piai 1987)
- Fenoverine (Galeone 1992)
- Octatropine methylbromide (Barbara 1979)

- Octilonium bromide (Baldi 1992; Capurso 1984; Capurso 1992a; D'Arienzo 1980; Defrance 1991; Ferrari 1986)
- Otilonium bromide (Baldi 1982; Baldi 1991; Battaglia 1998; Evangelista 1999; Glende 2002)
- Pinaverium bromide (Awad 1997; Corazza 1983; Delmont 1981; Dubarry 1977; Galeone 1986; Levy 1977; Lu 2000; Virat 1987)
- Prifinium bromide (Piai 1979; Sasaki 1985)
- Propinox (Pulpeiro 2000)
- Rociverine (Ghidini 1986)
- Secoverine (Ehsanullah 1985)
- Syntropium bromide (Galeone 1990)
- Trimebutine (Dumitrascu 2000; Ghidini 1986; Luttecke 1980; Moshal 1979; Schaffstein 1990; Schang 1993)
- Zamifenacin (Houghton 1997)

There were 23 included studies remaining. Nine were conducted in the UK (Dew 1984; Evans 1998; Flavell Matts 1967; Gilbody 2000; GP research group; Mitchell 2002; Nash 1986; Prout 1983; Ritchie 1979); eight in the rest of Europe, one in the USA, three in Australia and New Zealand, one in India, and one in Taiwan.

Generally, we excluded reports of studies if they were not in English, however, some studies translated and included in the Cochrane review were exceptions to this (Berthelot 1981; Czalbert 1990; Lech 1988; Schafer 1990). We had some reservations about doing this because of a difference of opinion with the Cochrane review on the eligibility of one of the English-language studies (Fielding 1980) due to the absence of reporting randomisation, but the study was included in the Cochrane review. However, we decided to include all the non-English language studies in the Cochrane review on the basis of trust. There was only one paper (Czalbert 1990) for which we were unable to verify that the patients were randomised to treatments, and we decided to include this and carry out a sensitivity analysis.

The majority of studies (17/23) had fewer than 100 patients, with six having 20 or fewer in the intervention arm (Carling 1989; Czalbert 1990; Evans 1998; Lech 1988; Ritchie 1979; Tasman-Jones 1973). Six studies had more than 100 patients in total and one had 712 patients (Schafer 1990). Some of the studies were of crossover design, so fewer patients are required to achieve adequate power.

Study Design
Setting: The majority of studies took place in secondary care; two studies were in primary care (Gilbody 2000; GP research group); five studies did not report the setting or it was not possible to determine it because the report was not in English (Berthelot 1981; Czalbert 1990; Evans 1998; Flavell Matts 1967; Schafer 1990).

There were ten crossover studies (Dew 1984; Evans 1998; Flavell Matts 1967; GP research group; Hennessy 1975; Lawson 1988; Nash 1986; Prout 1983; Tasman-Jones 1973; van Outryve 1995) in which participants were allocated to receive both the intervention and control treatments during the course of the study, in a random order. All of these studies had either no 'washout period' or it was not reported, in which case this was assumed to be none. As the GDG had specified a washout period of one week minimum, all these studies were eliminated from the analysis. One additional crossover study (Carling 1989) reported first period results for the global improvement of symptoms outcome, so these data were combined with those from parallel design studies.

The remaining 12 studies had a parallel design (Berthelot 1981; Czalbert 1990; Gilbody 2000; Inauen 1994; Kruis 1986; Lech 1988; Liu 1997; Mitchell 2002; Nigam 1984; Page 1981; Richie 1979; Schafer 1990).
Thirteen studies were therefore included in the analysis (12 parallel and one crossover trial, first period only).

Two studies had more than two arms: Schafer (1990) compared hyoscine plus paracetamol and hyoscine alone with placebo; Carling (1989) compared peppermint oil with hyoscamine and placebo; Ritchie (1979) compared hyoscine with placebo, and hyoscine+lorazepam with lorazepam alone. There were thus 17 comparisons in the antispasmodics review.

The rest of the description of studies will focus on these studies/comparisons.

Population
The definition of IBS varied between studies: one used the Rome I criteria (Mitchell 2002); one used the Rome II criteria (Gilbody 2000); three met criteria defined by the authors that were similar to the above (Inauen 1994; Nigam 1984; Page 1981). In seven comparisons, the authors stated that the patients had IBS, with no further explanation (Berthelot 1981; Czalbert 1990; Liu 1997; 2 x Ritchie 1979; 2 x Schafer 1990). The remaining five comparisons (3 x Carling 1989; Kruis 1986; Lech 1988) did not use a formal definition but described a range of symptoms consistent with IBS.

Most studies included a combination of IBS types, one study specified constipation (Page 1981); one study (Carling 1989) included only patients with IBS-C and IBS-A; and eight were unclear (Berthelot 1981; Czalbert 1990; Gilbody 2000; Lech 1988; Liu 1997; Nigam 1984; 2 x Ritchie 1979).

None of the studies stated that any participants had IBS as result of gastrointestinal infection.

The majority of studies did not state the number of participants with bloating. Four studies had some patients with bloating (Czalbert 1990; Inauen 1994; Kruis 1986; Liu 1997). Three comparisons (3 x Carling 1989) stated that all patients had bloating.

Most of the studies did not describe symptom severity. Two studies stated that participants had symptoms of mixed severity (Inauen 1994; Mitchell 2002).

The age range of participants across studies was 16 to 79 years, with the mean age (where given) ranging from 34.5 to 48 years (Czalbert 1990). None of the studies particularly identified elderly participants.

Most studies had more women than men; two studies had more men than women (Liu 1997; Nigam 1984).

Interventions

The studies varied in the type of antispasmodic used:

- Eight comparisons used anti-muscarinics:
  - One dicyclomine bromide (Page 1981)
  - Two hyoscamine (2 x Carling 1989)
  - Three hyoscine (Nigam 1984; Schafer 1990; Ritchie 1979)
  - One hyoscine plus paracetamol (Schafer 1990)
  - One hyoscine plus diazepam (Ritchie 1979).
- Nine had direct-action smooth muscle relaxants:
  - One alverine (Mitchell 2002)
  - Four mebeverine (Berthelot 1981; Kruis 1986; Gilbody 2000; Inauen 1994 (the last two were also modified release)
  - Four peppermint oil (Carling 1989; Czalbert 1990; Lech 1988; Liu 1997).

None of the included studies used antispasmodics for short-term symptom relief. The duration ranged from two weeks to 16 weeks (Kruis 1986). The most common durations were 12 weeks (four studies), four weeks (four studies) and two weeks (four studies). Studies of less than 4 weeks duration were considered in sensitivity analyses because the GDG preferred a minimum of four weeks duration. This meant the following: two weeks – four studies (3 x Carling 1989; Page 1981); three weeks – one study (Inauen 1994).

Comparisons

The included studies covered the following comparisons (we have indicated with an asterisk the studies with less than 4 weeks duration):

- 13 comparisons of antispasmodics versus placebo:

- o Five gave anti-muscarinics (Carling 1989*; Nigam 1984; Page 1981*; Richie 1979; Schafer 1990)
  - o Seven gave direct-action smooth muscle relaxants (Berthelot 1981; Carling 1989*; Czalbert 1990; Kruis 1986; Lech 1988; Liu 1997; Mitchell 2002)
- One study (Ritchie 1979) compared antispasmodic plus diazepam versus diazepam alone
- One study (Schafer 1990) compared antispasmodic plus paracetamol versus paracetamol alone
- One study compared different classes of antispasmodic:
  - o One compared an anti-muscarinic with a smooth muscle relaxant (Carling 1989*)
- Two studies compared different types of antispasmodic in the same class (smooth muscle relaxant):
  - o Two studies (Gilbody 2000; Inauen 1994*) compared mebeverine modified release (200 g bid) with mebeverine conventional (135g tid).

Methodological Quality

The results of the quality assessment for included trials are shown in Appendix D.

The method of randomisation was reported in none of the studies. Allocation concealment was reported in one study (Ritchie 1979), which reported an adequate method, in which the drugs were issued in random order by the pharmacist. All but two of the studies reported that the patients were blinded to the interventions. One stated that the patients were not blinded (Inauen 1994). One study (Czalbert 1990) was unclear about patient blinding. Only one study (Mitchell 2002) described an *a-priori* power calculation. Several studies included in the review demonstrated baseline comparability of the groups, but 11 did not give baseline characteristics or were in non-English languages (Berthelot 1981; Lech 1988; Nigam 1984; Ritchie 1979; Schafer 1990). One study was not comparable at baseline (Liu 1997) for the severity of stool frequency (more severe for Colpermin). Three studies reported no withdrawals (Czalbert 1990; Nigam 1984; Ritchie 1979). One study (Page 1981) reported that more than 20% of patients in at least one arm (or overall) were not analysed or were lost to follow-up (attrition bias): 33% in the intervention group and 39% in the placebo group.

The GDG preferred a minimum intervention period of four weeks as this was felt to be clinically significant relative to potential effect. This meant the following were treated with caution: two weeks: Carling (1989); Inauen (1994); Page (1981).

The risk of bias was assessed for each included study and no studies were excluded from the analysis. Four studies were assessed as being at higher risk of bias: Page (1981) (attrition bias and duration less than four weeks); Inauen (1994) (patients not blinded and duration less than four weeks); Liu 1997 (lack of baseline comparability), and; Carling (1989) (duration less than 4

weeks). All these studies were treated with caution and sensitivity analyses were carried out. This was also done for Czalbert as discussed earlier.

RESULTS

A. Antispasmodics versus Placebo

There were 12 studies included in the analysis that compared antispasmodics with placebo, and a further two studies that compared antispasmodics + another intervention versus that other intervention alone. The GDG decided that it was inappropriate to combine the two types of comparison, even though each could be considered to be a comparison of antispasmodics versus placebo. The GDG's view was that the drugs could interact and their effects would not simply be additive.

One of the studies was in patients with constipation-predominant IBS (Page 1981); eight did not specify the type of IBS and the remainder had patients of mixed IBS-type. Therefore the studies were not stratified by IBS type. Similarly, there was too little information to separate by severity, post-infective cause or bloating status. Only one study (Mitchell 2002) reported using established criteria to diagnose IBS (Rome I).

Where outcomes were measured at different times during the study, we took the end-study results unless there were significant numbers of withdrawals or problems with compliance. Therefore, for the Kruis (1986) study we took the values at four weeks.

1. Global symptoms

a) Number of patients with improvement in global symptoms

Eight studies with 731 patients reported this outcome for the comparison antispasmodics versus placebo. As described in the general section, patient assessments of improvement were combined with symptom score related methods (unlike the Cochrane review). Overall the relative risk was 1.32 (95%CI 1.18, 1.48), i.e. statistically significant, favouring antispasmodics. There was some heterogeneity, but it was not significant (p=0.09; I^2=43%). The sensitivity analysis without Page (1981) made little difference to the summary statistics.

The full meta-analysis corresponded to a number needed to treat (NNT) of 6 (95%CI 5, 10), for a placebo group rate of 0 to 71%. This meant that clinicians needed to treat 6 patients to get additional benefit in relief of symptoms for one patient. Typically this is viewed as a low NNT.

Figure 1


```
Review:        IBS
Comparison:    08 Antispasmodics vs Placebo (all IBS types)
Outcome:       01 Global improvement of IBS symptoms

Study                    Antispasmodic      Placebo       RR (fixed)      Weight      RR (fixed)
or sub-category              n/N              n/N           95% CI          %           95% CI

01 Antispasmodic vs placebo
Czalbert 1990               16/17            12/17                          3.76       1.33 [0.96, 1.85]
Kruis 1986                  16/40            16/40                          5.01       1.00 [0.58, 1.71]
Mitchell 2002               27/42            23/45                          6.96       1.26 [0.87, 1.81]
Ritchie 1979                 4/12             0/12                          0.16       9.00 [0.54, 150.81]
Schafer 1990a              138/182          114/178                        36.12       1.18 [1.03, 1.36]
Page 1981                   27/32            16/30                          5.18       1.58 [1.10, 2.28]
Nigam 1984                   7/21             0/21                          0.16      15.00 [0.91, 246.93]
Lech 1988                   13/19             6/23                          1.70       2.62 [1.24, 5.57]
Subtotal (95% CI)             365              366                         59.04       1.32 [1.18, 1.48]
Total events: 248 (Antispasmodic), 187 (Placebo)
Test for heterogeneity: Chi² = 12.33, df = 7 (P = 0.09), I² = 43.2%
Test for overall effect: Z = 4.73 (P < 0.00001)

02 Antispasmodic + diazepam vs diazepam
Ritchie 1979                 5/12             4/12                          1.25       1.25 [0.44, 3.55]
Subtotal (95% CI)              12               12                          1.25       1.25 [0.44, 3.55]
Total events: 5 (Antispasmodic), 4 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.42 (P = 0.68)

03 Antispasmodic + paracetamol vs paracetamol
Schafer 1990b              143/177          126/175                        39.71       1.12 [1.00, 1.26]
Subtotal (95% CI)             177              175                         39.71       1.12 [1.00, 1.26]
Total events: 143 (Antispasmodic), 126 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.93 (P = 0.05)

Total (95% CI)                554              553                        100.00       1.24 [1.14, 1.35]
Total events: 396 (Antispasmodic), 317 (Placebo)
Test for heterogeneity: Chi² = 14.52, df = 9 (P = 0.11), I² = 38.0%
Test for overall effect: Z = 5.03 (P < 0.00001)

                        0.1  0.2   0.5   1   2   5   10
                        Favours Placebo   Favours Spasmolytic
```

Subgroup analysis into anti-muscarinic agents and direct-acting smooth muscle relaxants, for all antispasmodics-placebo comparisons (Figure 2) suggested there was little difference between classes of antispasmodics, although there was some heterogeneity (I2=57%, p=0.08) for the anti-muscarinics group. The RR for the random effects model was 1.51 (95%CI 1.00, 2.28), i.e. this result was sensitive to the method of analysis.

Figure 2


```
Review:        IBS 31 August 2006
Comparison:    08 Antispasmodics vs Placebo (all IBS types)
Outcome:       02 Global improvement of IBS symptoms by type of antispasmodic

Study                    Antispasmodic      Placebo       RR (fixed)      Weight      RR (fixed)
or sub-category              n/N              n/N           95% CI          %           95% CI

01 anti-muscarinics
Nigam 1984                   7/21             0/21                          0.27      15.00 [0.91, 246.93]
Page 1981                   27/32            16/30                          8.84       1.58 [1.10, 2.28]
Ritchie 1979                 4/12             0/12                          0.27       9.00 [0.54, 150.81]
Schafer 1990a              143/177          114/178                        60.85       1.26 [1.11, 1.44]
Subtotal (95% CI)             242              241                         70.22       1.38 [1.22, 1.57]
Total events: 181 (Antispasmodic), 130 (Placebo)
Test for heterogeneity: Chi² = 6.89, df = 3 (P = 0.08), I² = 56.5%
Test for overall effect: Z = 4.99 (P < 0.00001)

02 Smooth muscle relaxants
Czalbert 1990               16/17            12/17                          6.42       1.33 [0.96, 1.85]
Kruis 1986                  16/40            16/40                          8.56       1.00 [0.58, 1.71]
Lech 1988                   13/19             6/23                          2.91       2.62 [1.24, 5.57]
Mitchell 2002               27/42            23/45                         11.89       1.26 [0.87, 1.81]
Subtotal (95% CI)             118              125                         29.78       1.33 [1.06, 1.68]
Total events: 72 (Antispasmodic), 57 (Placebo)
Test for heterogeneity: Chi² = 4.31, df = 3 (P = 0.23), I² = 30.3%
Test for overall effect: Z = 2.44 (P = 0.01)

Total (95% CI)                360              366                        100.00       1.37 [1.22, 1.53]
Total events: 253 (Antispasmodic), 187 (Placebo)
Test for heterogeneity: Chi² = 11.01, df = 7 (P = 0.14), I² = 36.4%
Test for overall effect: Z = 5.45 (P < 0.00001)

                        0.1  0.2   0.5   1   2   5   10
                        Favours Placebo   Favours Spasmolytic
```

We noted that three of these studies were in hyoscine and the remaining one (Page 1981) was in dicyclomine bromide. Page (1981) had over 30% missing data and was only two weeks

in duration. In its absence there was little difference to the fixed effects result, but a substantial difference to the random effects model.

b) Global symptom score (mean)

This outcome was recorded by one study (Carling 1989), which compared the change in symptom score (before and after intervention) for peppermint oil and hyoscamine (atropine) compared with placebo. This was a crossover study that reported first period results. The study reported p values for the difference:

- Peppermint oil versus placebo: mean difference for change score was -11.8 (on a scale of 105 for a week); p=0.063 (i.e. statistically significant).
- Atropine versus placebo: mean difference for change score was -1.0 (on a scale of 105 for a week); p=0.46 (i.e. not statistically significant).

2. Individual symptoms

a) Pain

The following studies measured pain:

i.   Number of patients with no pain: one study (Liu 1997)
ii.   Number of patients with less pain: four studies (Lech 1988; Liu 1997; Mitchell 2002; Page 1981 (physician assessed)). The Kruis (1986) study was not included because of poor compliance – 4 week results were not reported for this outcome.
iii. Pain score (change): one study (Berthelot 1981).

Figure 3 shows the number of patients with less pain. There was a statistically significant increase in the number with less pain, favouring antispasmodics. However, there was also significant heterogeneity within the smooth muscle relaxants group (p=0.07, I^2=63%). This was possibly a drug specific effect: Lech (1988) and Liu (1997) gave peppermint oil and Mitchell (2002) gave alverine citrate. However, the duration of Mitchell (2002) was also longer (12 weeks versus 4 weeks).

Figure 3

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 anti-muscarinic agents					
Page 1981	32/34	21/37		26.42	1.66 [1.24, 2.22]
Subtotal (95% CI)	34	37		26.42	1.66 [1.24, 2.22]
Total events: 32 (Antispasmodic), 21 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.38 (P = 0.0007)					
02 smooth muscle relaxants					
Lech 1988	12/19	6/23		7.13	2.42 [1.12, 5.22]
Liu 1997	41/52	21/49		28.40	1.84 [1.29, 2.62]
Mitchell 2002	35/42	30/45		38.05	1.25 [0.98, 1.60]
Subtotal (95% CI)	113	117		73.58	1.59 [1.29, 1.96]
Total events: 88 (Antispasmodic), 57 (Placebo)					
Test for heterogeneity: Chi² = 5.46, df = 2 (P = 0.07), I² = 63.4%					
Test for overall effect: Z = 4.41 (P < 0.0001)					
Total (95% CI)	147	154		100.00	1.61 [1.36, 1.91]
Total events: 120 (Antispasmodic), 78 (Placebo)					
Test for heterogeneity: Chi² = 5.69, df = 3 (P = 0.13), I² = 47.3%					
Test for overall effect: Z = 5.47 (P < 0.00001)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours spasmolytics

A sensitivity analysis without Mitchell (2002) shows that this was responsible for the heterogeneity, but it was not clear why. It was interesting to note that this study was the only one that carried out a power calculation.

Figure 4: Sensitivity analysis

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 anti-muscarinic agents					
Page 1981	32/34	21/37		42.64	1.66 [1.24, 2.22]
Subtotal (95% CI)	34	37		42.64	1.66 [1.24, 2.22]
Total events: 32 (Antispasmodic), 21 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.38 (P = 0.0007)					
02 smooth muscle relaxants					
Lech 1988	12/19	6/23		11.51	2.42 [1.12, 5.22]
Liu 1997	41/52	21/49		45.85	1.84 [1.29, 2.62]
Subtotal (95% CI)	71	72		57.36	1.96 [1.42, 2.70]
Total events: 53 (Antispasmodic), 27 (Placebo)					
Test for heterogeneity: Chi² = 0.41, df = 1 (P = 0.52), I² = 0%					
Test for overall effect: Z = 4.08 (P < 0.0001)					
Total (95% CI)	105	109		100.00	1.83 [1.46, 2.29]
Total events: 85 (Antispasmodic), 48 (Placebo)					
Test for heterogeneity: Chi² = 0.94, df = 2 (P = 0.62), I² = 0%					
Test for overall effect: Z = 5.25 (P < 0.00001)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours spasmolytics

The one study (Berthelot 1981) recording a pain score showed a statistically significant decrease in pain score, favouring the antispasmodic (figure 5). The mean difference was -0.56 (95%CI -1.03, -0.09) on a scale of 1 to 4.

Figure 5

Study or sub-category	N	Antispasmodic Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Change score from baseline							
Berthelot 1981	36	-1.54(0.97)	33	-0.98(1.02)		100.00	-0.56 [-1.03, -0.09]
Subtotal (95% CI)	36		33			100.00	-0.56 [-1.03, -0.09]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.33 (P = 0.02)							

-4 -2 0 2 4
Favours Spasmolytic   Favours Placebo
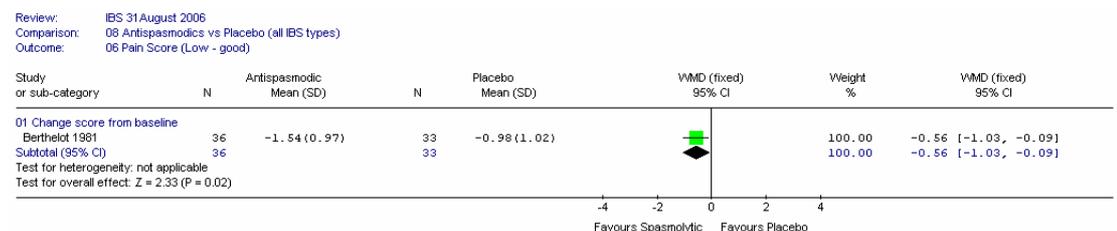
b) Bloating

Only two studies (Liu 1997 and Mitchell 2002) reported the number of patients with less bloating (Figure 6). Meta-analysis showed a high degree of heterogeneity between studies (p=0.0002, I^2=93%). It was not clear why these studies were so different, but see the discussion above regarding the pain outcome.

Figure 6

Review: IBS 31 August 2006
Comparison: 08 Antispasmodics vs Placebo (all IBS types)
Outcome: 05 No of patients with less bloating

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
02 smooth muscle relaxants					
Liu 1997	43/52	14/49		35.61	2.89 [1.83, 4.58]
Mitchell 2002	25/42	27/45		64.39	0.99 [0.70, 1.40]
Subtotal (95% CI)	94	94		100.00	1.67 [1.27, 2.20]

Total events: 68 (Antispasmodic), 41 (Placebo)
Test for heterogeneity: Chi² = 14.24, df = 1 (P = 0.0002), I² = 93.0%
Test for overall effect: Z = 3.65 (P = 0.0003)

| Total (95% CI) | 94 | 94 | | 100.00 | 1.67 [1.27, 2.20] |

Total events: 68 (Antispasmodic), 41 (Placebo)
Test for heterogeneity: Chi² = 14.24, df = 1 (P = 0.0002), I² = 93.0%
Test for overall effect: Z = 3.65 (P = 0.0003)

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours spasmolytics

c) Bowel habits

i. Number of patients with improved bowel habits

One study (Page 1981), with 71 patients and attrition bias, recorded the number of patients with improved bowel habits (figure 7). There was a statistically significant improvement with the antispasmodic.

Figure 7

Review: IBS 31 August 2006
Comparison: 08 Antispasmodics vs Placebo (all IBS types)
Outcome: 07 No of patients with improved bowel habits

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 anti-muscarinic agents					
Page 1981	29/34	20/37		100.00	1.58 [1.14, 2.19]
Subtotal (95% CI)	34	37		100.00	1.58 [1.14, 2.19]

Total events: 29 (Antispasmodic), 20 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.72 (P = 0.006)

| Total (95% CI) | 34 | 37 | | 100.00 | 1.58 [1.14, 2.19] |

Total events: 29 (Antispasmodic), 20 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.72 (P = 0.006)

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours spasmolytics

ii. Stool score

One study reported a stool score (Berthelot 1981) on a scale of 1-4. There was a statistically significant difference between antispasmodic and placebo, favouring the former.

Figure 8
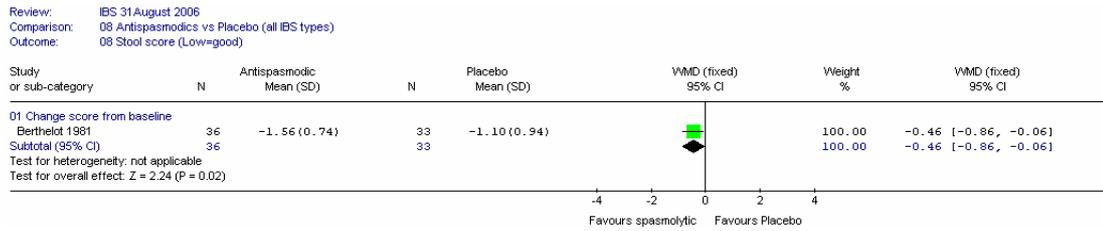
3. Adverse effects

Three studies (four comparisons) reported the number of patients with adverse effects. These were grouped by antispasmodic. In all cases there were wide confidence intervals, but it appeared that there were significantly more side effects for both atropine and dicyclomine bromide than placebo (Figure 9). In particular, the statistically significant effect of atropine was manifested as dry mouth and blurred vision, but the confidence intervals were very wide, as demonstrated in figure 10.

Figure 9
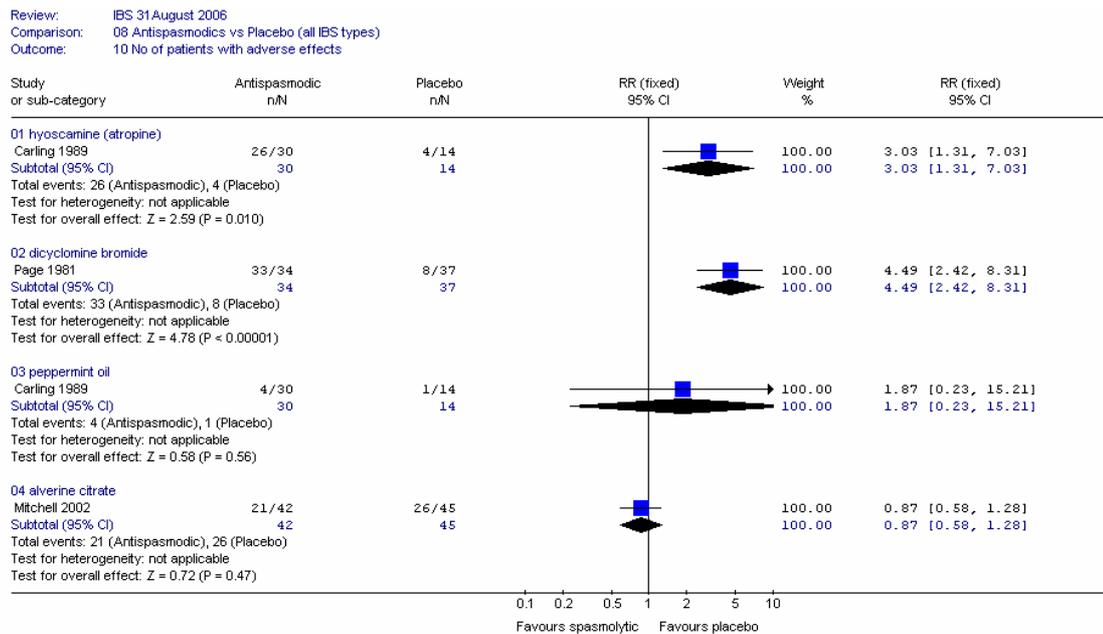
Figure 10: Specific side effects

Study or sub-category	Antispasmodic n/N	Placebo n/N	Peto OR 95% CI	Weight %	Peto OR 95% CI
01 hyoscamine (atropine) - blurred vision					
Carling 1989	14/30	0/14		100.00	8.18 [2.13, 31.46]
Subtotal (95% CI)	30	14		100.00	8.18 [2.13, 31.46]
Total events: 14 (Antispasmodic), 0 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.06 (P = 0.002)					
02 peppermint oil - blurred vision					
Carling 1989	1/30	0/14		100.00	4.33 [0.06, 291.39]
Subtotal (95% CI)	30	14		100.00	4.33 [0.06, 291.39]
Total events: 1 (Antispasmodic), 0 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.68 (P = 0.49)					
03 hyoscamine (atropine) - dry mouth					
Carling 1989	21/30	1/14		100.00	11.67 [3.33, 40.91]
Subtotal (95% CI)	30	14		100.00	11.67 [3.33, 40.91]
Total events: 21 (Antispasmodic), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.84 (P = 0.0001)					
04 peppermint oil - dry mouth					
Carling 1989	4/30	1/14		100.00	1.82 [0.25, 13.15]
Subtotal (95% CI)	30	14		100.00	1.82 [0.25, 13.15]
Total events: 4 (Antispasmodic), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.60 (P = 0.55)					

0.01   0.1   1   10   100

Favours spasmolytic   Favours placebo

B. Antispasmodic type 1 versus Antispasmodic type 2

One study compared different types of antispasmodic: Carling (1989) compared hyoscamine (atropine) with peppermint oil (smooth muscle relaxant).

1. Global outcomes

a) Global symptom score

The study did not record standard deviations or p-values.

2. Adverse effects

The study showed a statistically significant increase in side effects for atropine compared with peppermint oil, although the confidence interval was wide.

Figure 11

Study or sub-category	Antispasmodic 1 n/N	Antispasmodic 2 n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Atropine vs peppermint oil					
Carling 1989	26/30	4/30		100.00	6.50 [2.58, 16.36]
Subtotal (95% CI)	30	30		100.00	6.50 [2.58, 16.36]
Total events: 26 (Antispasmodic 1), 4 (Antispasmodic 2)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.97 (P < 0.0001)					
Total (95% CI)	30	30		100.00	6.50 [2.58, 16.36]
Total events: 26 (Antispasmodic 1), 4 (Antispasmodic 2)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.97 (P < 0.0001)					

0.1   0.2   0.5   1   2   5   10

Favours Spasmolytic1   Favours Spasmolytic2

C. Comparison of antispasmodics in the same class

Two studies (Gilbody 2000; Inauen 1994) compared modified release mebeverine twice daily (total 400 mg) with conventional mebeverine three times daily (total 405 mg). The two studies had different durations: Gilbody (2000) had duration of eight weeks, and; Inauen (1994) had duration of three weeks. Gilbody (2000) was also carried out in general practice.

1. Global improvement of symptoms

a) Number of patients with improved symptoms

There appeared to be little difference between the two formulations (figure 12).

Figure 12

Review: IBS 31 August 2006
Comparison: 10 Antispasmodic 1 vs 2 within same type
Outcome: 01 Global improvement of IBS symptoms

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 mr mebeverine vs conventional mebeverine					
Gilbody 2000	57/80	58/80		75.32	0.98 [0.81, 1.19]
Inauen 1994	22/24	19/24		24.68	1.16 [0.91, 1.47]
Subtotal (95% CI)	104	104		100.00	1.03 [0.88, 1.20]
Total events: 79 (Antispasmodic), 77 (Placebo)					
Test for heterogeneity: Chi² = 1.18, df = 1 (P = 0.28), I² = 15.3%					
Test for overall effect: Z = 0.32 (P = 0.75)					
Total (95% CI)	104	104		100.00	1.03 [0.88, 1.20]
Total events: 79 (Antispasmodic), 77 (Placebo)					
Test for heterogeneity: Chi² = 1.18, df = 1 (P = 0.28), I² = 15.3%					
Test for overall effect: Z = 0.32 (P = 0.75)					

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours Spasmolytic

2. Adverse effects

a) Pain

One study in general practice (Gilbody 2000) showed little difference between the interventions for the number of patients with adverse effects, although it was noted that the number of adverse effects was high in both groups and the study said this included symptoms associated with IBS (pain and diarrhoea).

Figure 13

Review: IBS 31 August 2006
Comparison: 10 Antispasmodic 1 vs 2 within same type
Outcome: 10 No of patients with adverse effects

Study or sub-category	Antispasmodic n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 mebeverine mr vs mebeverine conventional					
Gilbody 2000	63/80	66/80		100.00	0.95 [0.82, 1.11]
Subtotal (95% CI)	80	80		100.00	0.95 [0.82, 1.11]
Total events: 63 (Antispasmodic), 66 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.60 (P = 0.55)					

0.1 0.2 0.5 1 2 5 10
Favours spasmolytic   Favours placebo

Adverse Effects

An adverse effects review has been carried out and is reported in section 8.5.1. There were six included RCTs in the review of adverse effects of antispasmodics (Grillage 1990; Gilbody 2000; Schaffstein 1990; Mitchell 2002; Liu 1997; Van Outryve 1995). The interventions and comparators were extremely varied, as was the reporting of adverse effects outcomes. In view

of this, no meta-analysis was performed. Dry mouth, drowsiness, dizziness and constipation were the most common complaints reported amongst people taking antispasmodics.

One of the limitations of the adverse effects review was that many of the adverse outcomes of interest were very similar to the symptoms of the IBS itself. For instance, antispasmodics were associated with drowsiness and constipation, both of which are commonly seen in people with untreated IBS. This made it difficult for the investigators to differentiate between the progress of the condition and the harmful effects of the drug.

ECONOMIC LITERATURE FOR ANTISPASMODICS

No relevant health economic analyses were identified on the cost-effectiveness of antispasmodic therapy in the treatment of IBS.

COST-EFFECTIVENESS ANALYSIS FOR ANTISPASMODICS

This section describes the health economic analysis undertaken to inform recommendations on the use of antispasmodics as a long-term maintenance therapy in IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

- The following interventions were treated as a class of interventions with the same clinical effectiveness as there was insufficient evidence to demonstrate a significant difference in effectiveness between them; hyoscine, mebeverine (standard and slow release), peppermint oil, dicycloverine (dicylcomine), alverine.
- There was insufficient evidence to demonstrate that atropine (hyoscamine) was more effective than placebo. Therefore, the cost-effectiveness of atropine was not estimated.
- The studies included in the clinical effectiveness review did not stratify results by IBS subtype, so it was not possible to estimate the effectiveness for each of the subtypes separately. Therefore, it was assumed that antispasmodics are equally effective in all subtypes.

Modelled response rates

In the basecase scenario the response rate of 45% in the no treatment arm is taken from the Mearin (2004) cohort study. This represents the group of patients for whom we would expect an improvement in symptoms without any specific intervention. The RR of response for antispasmodics versus placebo is 1.32; therefore the response rate in the intervention arm is 59% (=45% x 1.32), giving an absolute difference in response between the intervention and no treatment arms of 14% (=59%-45%) during the 1st month. In the basecase scenario the response rate for the subsequent interventions is assumed to be equal to the response rate for the first intervention. Therefore an additional 14% of those who do not respond to the first

intervention achieve a therapeutic response to the second intervention, increasing the overall response rate to 65% (=59%+14%*[1-59%]) after the second month. The response rate over time for the basecase is given in Figure 14.

We have also considered an alternative scenario in which no patient in the comparator arm achieves an improvement in symptoms but there is still a 14% chance of response for the intervention arm.

Figure 14: Modelled response rate for antispasmodic therapy, when allowing up to 4 switches of antispasmodic therapy, and no treatment
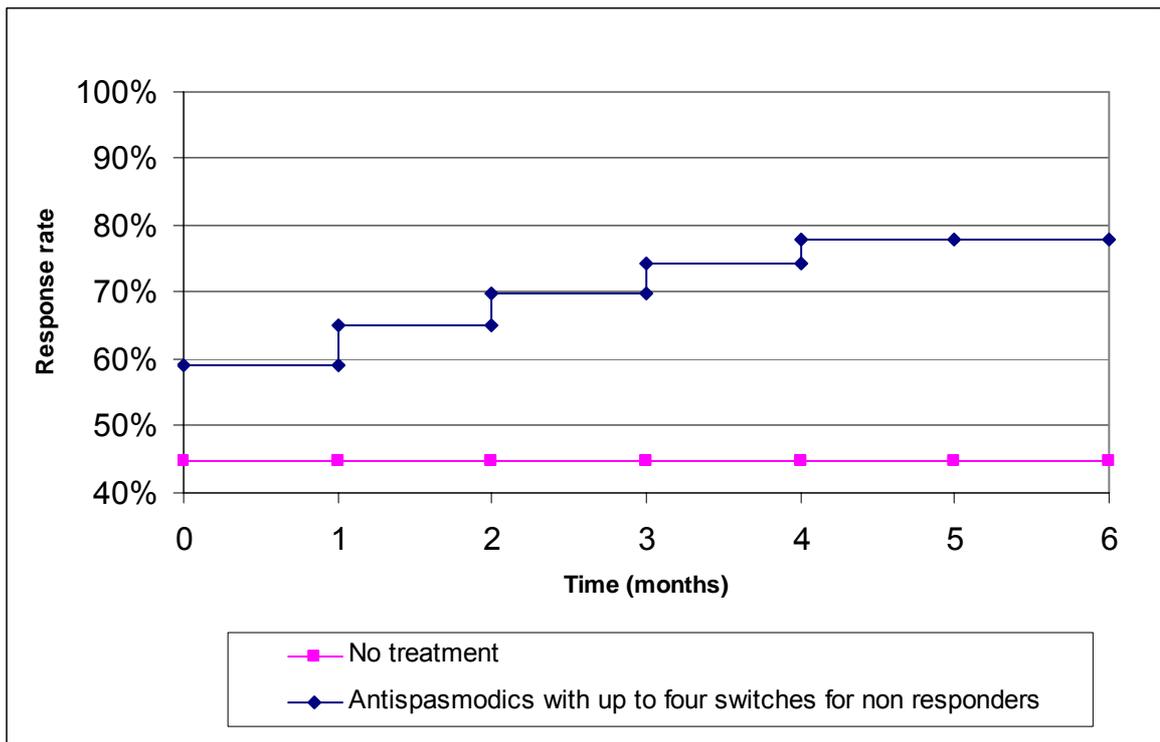
Table 1: Intervention specific parameters – antispasmodics

Description	Value		Evidence
RR of response for intervention vs placebo	1.32 (1.18 – 1.48)		Meta-analysis of RCT evidence for improvement in global symptoms
Maximum number of switches considered	4		Limited by number of effective interventions
Drug costs			
Intervention	**Dose per day**	**Cost per month* (assuming lowest cost preparation)**	
Hyoscine	30mg	£4.22	
Mebeverine	400mg	£6.76	
Peppermint oil	0.6ml	£7.65	
Alverine	360mg	£20.99	
Dicyclomine hydrochloride	160mg	£24.54	

* British National Formulary (Joint Formulary Committee 2007)

Results

Table 2: Incremental cost-effectiveness of allowing subsequent switches in antispasmodic therapy

Strategy	Cost	QALY	Incremental Cost per QALY compared to previous row	Incremental cost per QALY compared to no treatment
No treatment	£0.00	1.60	N/A	N/A
Intervention, no switches	£3,469	2.11	£6,792	£6,792
Intervention with up to 1 switch	£4,640	2.28	£6,728	£6,775
Intervention with up to 2 switches	£5,654	2.40	£8,497	£7,031
Intervention with up to 3 switches	£7,005	2.48	£17,620	£7,952
Intervention with up to 4 switches	£8,189	2.52	£27,047	£8,857

Table 2 gives the incremental cost-effectiveness for several treatment pathways with different numbers of therapy switches included. These results demonstrate that although a treatment pathway which allows up to 4 switches had a cost per QALY under £20,000 compared to no treatment, the incremental cost-effectiveness compared to a pathway which allows up to 3 switches was greater than £20,000. Therefore, at a willingness to pay threshold of £20,000 per QALY, it would only be cost-effective to allow up to 3 switches as the additional switch would not provide sufficient additional benefit.

These results are an estimate of the cost-effectiveness over the first 6 months after the initiation of antispasmodic therapy. The cost per QALY for continuing antispasmodic therapy beyond 6 months is lower than the cost per QALY during the initial 6 months provided that treatment is reviewed every 6 months and discontinued in patients who no longer experience a therapeutic benefit, either due to lack of effectiveness or a change in their symptom profile.

The probabilistic sensitivity analysis provided an estimate of the uncertainty in the cost per QALY estimate due to uncertainty in the efficacy estimate, the probability of response in the no treatment arm and the utility gain. The CEAC in Figure 15 shows the uncertainty surrounding the relative cost-effectiveness of allowing each additional switch in therapy. For example, the first curve on the left of Figure 15 shows that it is highly likely that using a first antispasmodic therapy would be cost-effective compared to no treatment as the probability of the cost per QALY being under a £20K threshold is over 90%. The second curve from the left is very close to the first and it shows that it is highly likely that allowing patients to switch once to an alternative antispasmodic therapy would be cost-effective compared to no further antispasmodic therapy for those who do not respond to the first. Similarly, the second switch is also highly likely to be cost-effective as it has a 91% probability of being under £20K. However, once the third switch of therapy is considered, it is only fairly likely to be cost-effective as there is a 53% likelihood that the true cost per QALY is under £20K and a 79% likelihood that it is under £30K. Providing four switches is fairly unlikely to be cost-effective compared to 3 switches as the cost per QALY has a 23% probability of being under £20K and a 51% probability of being under £30K. However, it should be noted that these estimates only consider the uncertainty in cost-effectiveness due to the accuracy of several input parameters and they do not reflect general uncertainty around the assumptions made in the model. The uncertainty from these assumptions was explored in the univariate sensitivity analysis.

Figure 15: Cost-effectiveness acceptability curves for antispasmodics (AS) compared to no treatment (NT) and for each additional switch of antispasmodic for non responders (AS)
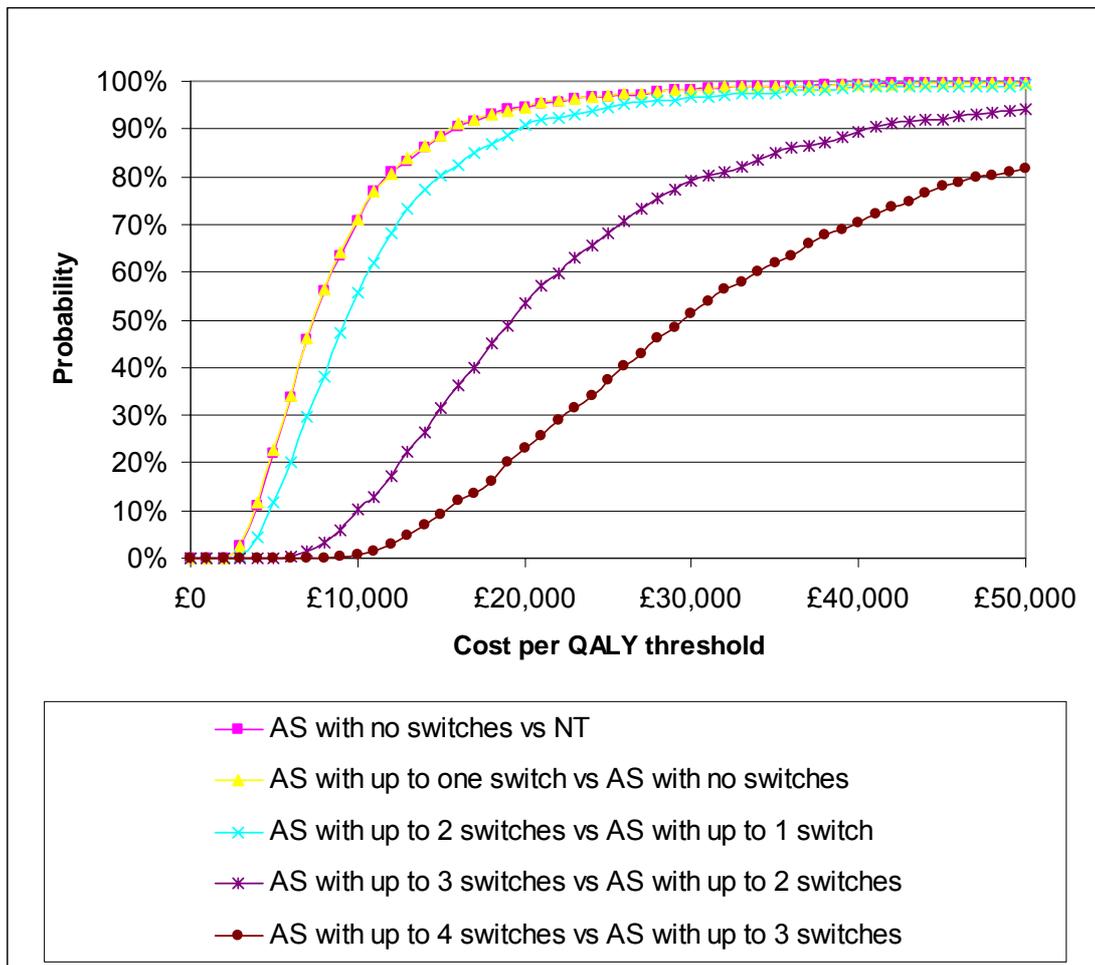
Univariate sensitivity results for up to three switches

The results in Table 3 show that initiating treatment with an antispasmodic and allowing up to three switches in therapy for non-responders had a cost per QALY of £7,952 compared to no treatment. When the response rate in the no treatment arm was taken from the average response rate in the RCTs of pharmacological interventions (47%), rather than from the cohort study (45%), the cost per QALY was very similar at £7,539. Maintaining the 14% difference in response between the two arms but reducing the response rate in the no treatment arm from 45% to zero marginally decreased the cost per QALY to £7,772. Reducing the response rate for subsequent antispasmodics, in patients who have not responded to initial therapy, by 50% significantly increased the cost per QALY for each subsequent switch of therapy, such that only 2 rather than 3 switches of therapy could be provided for a cost per QALY under £20,000. However, the cost per QALY for 3 switches compared to no treatment was only moderately increased to £10,003 per QALY and even when the response rate for subsequent therapy was

set to zero, the cost per QALY for antispasmodic therapy with up to 3 switches remained under £20,000 when compared to no treatment.

We carried out a threshold analysis to determine whether antispasmodic therapy would still be cost-effective for lower gains in health related quality of life. In the basecase it was assumed that patients who respond to therapy accumulate 0.071 QALYs more per annum than patients who do not respond. For comparison, a gain of 0.135 QALYs would represent a complete remission of IBS symptoms. When the QALY gain associated with a response to therapy was reduced to 0.028 QALYs, the cost per QALY of providing antispasmodic therapy with up to 3 switches was estimated to be above £20,000 per QALY compared to no treatment.

If a patient takes a therapy on an as needed basis, it would be reasonable to assume that they take the therapy on days when their quality of life is significantly affected by their IBS symptoms but not on days when their symptoms are less severe. It has therefore been assumed in the model that they only accrue QALY benefits and drug costs on the days they take the therapy. However, it is still necessary to assess all patients for response after 1 month of initiating therapy. This means that it would be less cost-effective to initiate therapy in patients who use the therapy on fewer days, as the monitoring costs are just as high but the benefits are lower. This is shown by the estimated cost per QALY of £20,578 for patients who take the therapy on 25% on days. However, a more detailed look at the results for these patients (data not tabled) shows that up to 1 switch of therapy could be provided for a cost per QALY of £19,414 with a 45% likelihood of being under £20,000 per QALY and a 73% likelihood of being under £30K per QALY. So, cost-effective treatment strategies may be available for patients who do not experience severe symptoms as frequently.

If a patient also takes another medication (such as a laxative or anti-motility agent), then these medications can be reviewed at the same time, so it may be cost-effective to provide both therapies. For example, if laxatives are prescribed with the antispasmodic and both used on 25% of days then allowing up to 2 switches of both treatments was estimated to be cost-effective with a cost per QALY of £10,107 compared to no treatment a cost per QALY of £17,393 compared to 1 switch.

Table 3: Sensitivity results for antispasmodic therapy with up to 3 switches compared to no treatment for 100 patients with IBS (all subtypes)

Scenario	No Treatment		Intervention		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	1.60	£7,005	2.48	£7,952
No response in no treatment arm	£0	0.00	£9,931	1.29	£7,722
Response rate in no treatment arm from RCTs	£0	1.69	£6,736	2.59	£7,539
Response to subsequent treatment half as likely as response to first	£0	1.60	£7,077	2.30	£10,003
Response to subsequent treatment zero	£0	1.60	£7,125	2.11	£13,949
Treatment used 75% of days	£0	1.60	£6,180	2.26	£9,355
Treatment used 50% of days	£0	1.60	£5,356	2.04	£12,161
Treatment used on 25% of days	£0	1.60	£4,532	1.82	£20,578
Half of treatments obtained over the counter	£0	1.60	£5,356	2.48	£6,080
Higher cost formulations (same dose)	£0	1.60	£7,298	2.48	£8,285
High utility gain of 0.135	£0.00	3.02	£7,005	4.69	£4,200
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.028 and 0.029.				

GDG DISCUSSION

The GDG consensus was that antispasmodics should be used as first line therapy alongside dietary and lifestyle advice for people with IBS, particularly those with pain occurring as spasm. Antispasmodics should be taken as needed.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. There is a fair amount of good quality evidence showing significant improvement in symptoms for antispasmodics, in general, and smooth muscle relaxants, in particular, when compared to placebo.

2. There is a fair amount of evidence showing significant global improvement in symptoms for antimuscarinic agents compared with placebo.

3. Sub-group analysis suggests there is little difference in the effect of antimuscarinic agents and smooth muscle relaxants for global improvement of symptoms.

4. There is a moderate amount of good quality evidence showing a significant reduction in pain and an improvement in bowel habit for antispasmodics compared with placebo.

5. There is a moderate amount of good quality evidence showing that conventional and modified release mebeverine were equally effective.

ADVERSE EFFECTS EVIDENCE STATEMENT

There is limited evidence that antispasmodics are associated with dry mouth, dizziness and drowsiness.

HEALTH ECONOMIC STATEMENT

Evidence from a decision analytic model showed that antispasmodics (hyoscine, mebeverine, peppermint oil, dicycloverine, alverine) are cost-effective for long-term maintenance use in individuals with IBS. The cost-effectiveness estimate is based on a clinical pathway in which response is assessed after one month and non-responders are switched to an alternative antispasmodic with the lower cost antispasmodics used before higher cost antispasmodics. The analysis assumes that the response to each subsequent therapy is independent of the response to previous antispasmodics. Trying a further antispasmodic is unlikely to be cost-effective in individuals who have not responded to four previous antispasmodics. The cost-effectiveness analysis assumes that treatment is reviewed every 6 months to establish whether antispasmodic therapy is still relevant to the individual's symptom profile.

EVIDENCE TO RECOMMENDATIONS

The GDG took into consideration the clinical and cost effective evidence. They noted the limited evidence about adverse effects, but did not consider this to be a significant factor in practice. The GDG wished to encourage primary care clinicians to give antispasmodics as a first line therapy alongside dietary and lifestyle advice.

RECOMMENDATION

Healthcare professionals should consider prescribing antispasmodic agents for people with IBS. These should be taken as required, alongside dietary and lifestyle advice.

8.4 Antidepressants

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some criteria specific to the antidepressants review are reported below.

Types of patients

For this review, patients were required to have IBS and not to have inflammatory bowel disease or major psychiatric disorders.

Types of studies

Studies that investigated drugs not listed in the BNF were excluded. These included: Amineptine; Amoxapine; Desipramine, and; Pirenzepine.

The GDG decided that crossover studies were acceptable and that the washout period for this review should be at least one week. Trials with shorter washout periods were not included in the analysis.

Types of intervention

Studies included the following interventions:

Tricyclics and related antidepressants:
- Amitriptyline (Triptafen®, Triptafen-M® (with perphenazine)
- Clomipramine (Anafranil®)
- Dosulephin (Prothiaden®)
- Doxepin (Sinepin®)
- Imipramine
- Lofepramine (Feprapax®, Lomont®, Gamanil®)
- Nortriptyline (Allegron®)
- Trimipramine(Surmontil®)
- Mianserin
- Trazodone (Molipaxin®).

Selective serotonin re-uptake inhibitors (SSRIs):
- Citalopram (Cipramil®)
- Escitalopram (Cipralex®)
- Fluoxetine (Prozac®)
- Fluvoxamine (Faverin®)
- Paroxetine (Seroxat®)

- Sertraline (Lustral®).

Monoamine oxidase inhibitors (MAOIs):
- Phenelzine (Nardil®)
- Isocarboxazid
- Tranylcypromine.

Reversible MAOIs:
- Moclobemide (Manerix®).

Others:
- Duloxetine (Cymbalta®)
- Flupentixol (Fluanxol®)
- Mirtazapine (Zispin Soltab®)
- Reboxetine (Edronax®)
- L-Tryptophan (Optimax®).

The following comparisons were included:
- Antidepressant versus placebo (or nothing)
- Antidepressant type 1 versus Antidepressant type 2
- Antidepressant dose 1 versus Antidepressant dose 2
- Antidepressant versus other interventions.

NB: In spite of the large placebo effect associated with IBS, comparisons with no treatment were included.

The antidepressants review was concerned with medium and longer term symptom relief. Medium term treatment was defined as three months and long term as between six months and one year.

Subgroup analyses

We planned to carry out subgroup analyses by type of antidepressant; dose; mode of delivery, and; duration of intervention.

Search strategy for identification of studies

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. The search strategies are listed in Appendix B.

The search strategy identified 1458 studies. The titles and abstracts of these studies were assessed. Thirty were identified to be potentially relevant to the review and these papers were retrieved in full. The reference lists for each of the retrieved studies were inspected for further potential papers, but none were identified. The 17 excluded studies are listed in the Appendix, along with reasons for exclusion. The remaining 13 studies were included (Boerner 1988; Creed 2003; Kuiken 2003; Myren 1982; Myren 1984; Rajagopalan 1998; Schrivastava 1984; Steinhart 1982; Tabas 2004; Tanum and Malt 1996; Tripathi 1983; Vij 1991; Quartero 2007). One of these studies was a Cochrane review (Quartero 2007). The Myren (1982) study did not state that it was randomised although it was double blind. Since this study was included in the Cochrane review (under the name of Block (1983), which is an identical paper in Norwegian) we included it too, but treated it with caution.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

One Cochrane review was identified (Quartero 2007) and this guideline review is an update, with revision to make it appropriate to the UK. This mainly involved elimination of one of the included studies (Heefner 1978) because the antidepressant, desipramine, is not in the BNF. Therefore the analysis for the guideline review was based on six studies included in the Cochrane review (Boerner 1988; Myren 1982; Myren 1984; Rajagopalan 1998; Shrivastava 1984; Vij 1991) and six additional studies (Creed 2003; Kuiken 2003; Steinhart 1982; Tabas 2004; Tanum and Malt 1996; Tripathi 1983). Tripathi (1983) and Shrivastava (1984) were two reports of a single trial, i.e. there were 11 included studies.

One study was conducted in the UK (Creed 2003); one was conducted in Germany (Boerner 1988); one in The Netherlands; three in Norway (Myren 1982; Myren 1984; Tanum and Malt 1996): three in India (Rajagopalan 1998; Tripathi 1983; Vij 1991) and the remaining two in the United States (Steinhart 1982; Tabas 2004).

Eight studies had fewer than 100 patients (Myren 1982; Boerner 1988; Kuiken 2003; Steinhart 1982; Rajagopalan 1998; Tripathi 1983; Tabas 2004; Tanum and Malt 1996). Two studies had fewer than 20 patients in the intervention arm (Kuiken 2003; Steinhart 1982).The latter was a crossover design so fewer patients were required to achieve adequate power. Creed (2003) had 257 patients and Myren (1984) had 258.

Study Design

Setting: The majority of studies took place in secondary care, one of which treated inpatients (Tripathi 1983); one was stated to take place in primary care in a later paper (Myren 1982); three included patients from both primary and secondary care settings (Boerner 1988; Myren 1984; Tabas 2004) and one study (Steinhart 1982) did not report the setting.

Population

The age range of patients across the IBS studies was 13 to 75 years, with the mean age, where given, ranging from 35 to 41 years. The study with a lower age of 13 was Tripathi (1983); in this study the range was 13 to 60 years, with a mean of 37. The GDG did not consider this level of children to be important. No study particularly identified elderly patients.

Five studies had more women than men (Creed 2003; Myren 1982; Steinhart 1982; Tabas 2004; Tanum and Malt 1996); three studies had about the same number of men and women (Kuiken 2003; Rajagopalan 1998; Myren 1984) and two studies did not give the proportions of patients by gender (Boerner 1988; Tripathi 1983).

Six studies identified the type of IBS as mixed (Boerner 1988; Creed 2003; Kuiken 2003; Steinhart 1982; Tabas 2004; Vij 1991). The remaining studies gave no information regarding type of IBS. In Tanum and Malt (1996) only 60% of the patients had IBS; the rest had non ulcer dyspepsia (NUD). The authors reported that there was no *significant* difference in response between IBS and NUD patients, although there was a trend towards a slightly better response in the NUD patients.

Two studies stated that the patients had severe IBS (Creed 2003; Steinhart 1981); one study had mixed severity patients (Vij 1991). Creed (2003) was stratified by pain level before randomisation. Tabas (2004) also selected patients who were non-responders to placebo.

Four studies reported the inclusion of refractory IBS patients: Creed (2003) included patients that had failed to respond to usual treatment and had a median duration of IBS of eight years; Kuiken (2003) stated that the patients had all been treated unsuccessfully previously. Steinhart (1981) reported a mean duration of IBS of 5 years and stated that all patients had received antispasmodics previously; Tabas (2004) included patients who had failed to respond to a high fibre diet.

Three studies reported a long duration of IBS: Tanum and Malt (1996) had a mean duration of symptoms of about 8 years; Rajagopalan (1998) had a mean duration of about 4 years; Kuiken (2003) had a mean duration of symptoms of 5.9 years.

One study implied that some of the patients did not have refractory IBS: in Myren (1984), 45 to 61% were not taking other drugs before the study started, and the other studies did not report previous treatments.

Five studies stated that the patients had some depression: Tabas (2004) reported that 27/81 (33%) had a score greater than 10 on the Beck Depression Inventory (although major psychiatric illnesses were excluded); Creed (2003) reported that 47% had anxiety or depression; Steinhart (1981) stated that 57% had depression and 79% anxiety; Vij (1991) had 57% with psychiatric co-morbidities; Boerner (1988) reported that some patients had depression (mid point on the Hamilton depression scale). Kuiken (2003) excluded patients with depression.

One study stated that all the patients had anxiety (Tripathi 1983). Three studies reported that the patients did not have psychiatric disorders: Rajagopalan (1998) stated that patients had no major medical or psychiatric illnesses; Tanum and Malt (1986) excluded patients with schizophrenia, anxiety or depression, and; Myren (1982) and Myren (1984) both showed low scores on depression and anxiety scales.

Interventions

The interventions included:

- Tricyclic Amitriptyline up to 75mg for 3 months (Rajagopalan 1998) and 50mg for 1 month (Steinhart 1982) – c.f. BNF levels for depression treatment: initially 75mg, increasing to 150 to 200mg
- Tricyclic Trimipramine 30 to 50mg for 4 to 6 weeks in three studies (Myren 1982; Myren 1984; Tripathi 1983) – c.f. BNF levels for depression: initially 50-75mg, then 150 to 300mg
- Tricyclic Doxepin 50mg for 8 weeks (Boerner 1988) and 75mg for 6 weeks (Vij 1991) – c.f. BNF levels for depression: initially 75mg, then 30 to 300mg
- Tricyclic-related Mianserin 30 mg initially, then up to 120mg for weeks 2 to 7, then tapered in week 8 – c.f. BNF levels for depression: 30 to 40mg initially; usual dose 30 to 90mg
- SSRI Paroxetine 20mg per day for 3 months (Creed 2003) and up to 40mg for 3 months; 23% 10mg; 43% 20mg; 33% 40mg (Tabas 2004) – c.f. BNF levels for depression: initially 20mg, then up to 50mg
- SSRI Fluoxetine 20mg per day for 6 weeks – c.f. BNF levels for major depression: 20mg once daily increased after 3 weeks if necessary, usual dose 20 to 60mg.

Comparisons

The majority of comparisons were of antidepressants versus placebo. One study (Creed 2003) compared antidepressants with usual care. One study compared different doses of antidepressants (Myren 1984).

Antidepressant versus placebo

- Tricyclics versus placebo:
  - Trimepramine (Myren 1982; Myren 1984; Tripathi 1983)
  - Amitriptyline (Rajagopalan 1998; Steinhart 1982)
  - Doxepin (Boerner 1988; Vij 1991)
  - Mianserin (Tanum and Malt 1996).

- SSRI versus placebo:
  - Paroxetine (Tabas 2004) included high fibre diet (>25g daily) in both the treatment and placebo groups (NB. these patients had already been identified as non-responders to high fibre)
  - Fluoxetine (Kuiken 2003).

Antidepressant versus usual care

- SSRI versus usual care:
  - Paroxetine (Creed 2003) versus 'routine care' by gastroenterologist and GP including antispasmodics, laxatives, antidiarrhoeal medication or additional analgesics.

Antidepressant versus psychotherapy (this is covered in the psychotherapy review)

- Paroxetine versus psychotherapy (Creed 2003).

It was decided to combine the SSRI studies with comparators of placebo and usual care using subgroup analyses.

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D.

The method of randomisation was adequate in four of the studies: three used computer generated random numbers (Creed 2003; Kuiken 2003; Tabas 2004) and the other used a random number table (Vij 1991). Myren (1982) did not state that the study was randomised, but stated that the study was double blinded; we included this study because it was stated to be an RCT in the Cochrane review. The remaining studies did not state the method of randomisation. Allocation concealment was reported in four studies (Creed 2003; Myren 1982; Tabas 2004). Creed (2003) and Kuiken (2003) reported that an independent third party carried out the randomisation; Tabas (2004) reported that identical capsules were sealed in sequentially numbered identical boxes.

The majority of studies reported that the patients were blinded to the interventions, with the exception of Creed (2003), in which blinding was not possible due to the nature of the comparisons. However, the outcome assessors were blinded in this study.

Three studies reported a sample size calculation (Creed 2003; Kuiken 2003; Tabas 2004), but Kuiken (2003) was powered for a different primary outcome (rectal sensitivity) and was underpowered for symptoms. The remaining studies gave no details of an *a priori* sample size calculation (Boerner 1988; Myren 1982; Myren 1984; Rajagopalan 1998; Tripathi 1983; Steinhart 1982; Tanum and Malt 1996; Vij 1991).

The majority of studies included in the review demonstrated baseline comparability of the groups, apart from Myren (1982) which was not comparable at baseline for vomiting, but levels of this were low in both groups (0.5 and 0.1 on 10cm VAS), and; Rajagopalan (1998) which was not comparable on stool type, with the antidepressant group having looser stools. In two studies there were no details of baseline characteristics (Tripathi 1983; Boerner 1988).

Only one study had total missing data of more than 20% (Rajagopalan 1998), comprising 9/20 in each group (45%). Those who dropped out had a significantly shorter duration of symptoms at recruitment, but otherwise there was no difference between completers and dropouts. Two studies had no missing data (Tripathi 1983; Myren 1982). Four reported missing data of less than 20% (Kuiken 2003; Tanum and Malt 1996; Tabas 2004; Vij 1991). However, there were 24% missing data in each arm of Vij (1991) for the outcome of pain. Myren (1984) and Steinhart (1982) provided no information regarding the number of drop outs.

In Creed (2003) there were missing data, 16% (14/86) in the paroxetine group; 14% (12/85) psychotherapy. 0% in the routine care group did not start the trial. A further 29/86 (34%) in the paroxetine group and 14/85 (16%) in the psychotherapy arm discontinued treatment, but these patients still appear to have been followed. Overall, loss to follow-up at three months was 12/86 (14%) for paroxetine, 11/85 (13%) psychotherapy and 7/86 (8%) usual care arm. At 15 months the authors contacted more of the patients. The authors reported that there were no significant differences at baseline between those who did and did not complete the treatments. For the 3 month pain score and SF36 outcome measures respectively, the patients included in the analysis were 74 and 59 (69%) paroxetine; 74 and 58 (68%) psychotherapy and 79 and 63 (73%) usual care, but some of these patients had discontinued treatment. We decided to include the results from this study, with some reservations, especially about the paroxetine arm and about the SF36 results. The study also recorded the number of patients with an improvement in global symptoms, based on the results from 74, 74 and 80 patients respectively. The GDG decided that this outcome was more representative because patients that dropped out due to side effects would not have rated their global symptoms as improved. The follow-up period in Creed (2003) allowed the patients to have paroxetine in all arms: 42% in paroxetine group, 19% in psychotherapy and 22% in the usual care group, i.e. the follow-up period should be considered to be partly confounded. Therefore we did not report the results for the follow-up period for the comparison paroxetine versus placebo.

Overall, we considered that Rajagopalan (1998) was at high risk of bias because of the extent and nature of the missing data and the baseline differences, and we decided not to include the results from this study in the analysis. Three other studies were treated with caution: Myren (1982), which was not stated to be randomised; Creed (2003) which had missing data and non compliance for pain and SF36; and some confounding in the follow up period; and Vij (1991), which had 24% missing data in each arm for the pain outcome. We examined the latter three studies with sensitivity analyses.

RESULTS

A. Antidepressants versus placebo or usual care

1. Global symptoms

a) Number of patients with global improvement of symptoms (pain, bloating and bowel habit)

Six studies, in 434 patients, reported the number of patients with improvement in global symptoms (Boerner 1988; Creed 2003; Kuiken 2003; Tabas 2004; Myren 1982; Vij 1991). The studies were combined in a meta-analysis, but as separate subgroups by type of antidepressant. The controlled trial (Myren 1982) was included in the tricyclics subgroup and examined in a sensitivity analysis. The comparisons of paroxetine with placebo and with usual care were also considered in sensitivity analyses. 'Usual care' was defined as patients receiving IBS treatment that was deemed appropriate by either their gastroenterologist consultant or general practitioner.

The difference between the antidepressants and placebo was statistically significant overall and for each subgroup. Within subgroups there was no heterogeneity, but between groups there was some (I^2=42%, p=0.12). The overall relative risk (RR) for the meta-analysis of 434 patients was 1.55 (95%CI 1.30, 1.84), which corresponded to a number needed to treat of 5 (95%CI 4, 7), for a control group rate of 22 to 68%.

For the tricyclic subgroup (n=180) the number with global improvement of symptoms was statistically significantly higher for the antidepressant group; RR 1.31 (95%CI 1.04, 1.64), which gave an NNT of 6 (95%CI 4, 34), for a control group rate of 22-68%. In the absence of the Myren (1982) study, which was not stated to be randomised, the RR for this group was 1.37 (95%CI 0.99 1.91), i.e. no longer significant and with some heterogeneity (I^2=62%, p=0.11).

For the meta-analysis of the three SSRI studies (n=254), there was a significant difference favouring antidepressant, RR 1.80 (95%CI 1.38, 2.34), with no heterogeneity (I^2=0%, p=0.48). This corresponded to an NNT of 4 (95%CI 3, 7), for a control group rate of 28-41%. In the absence of the Creed (2003) study the effect was slightly bigger (RR 1.85 (95%CI 1.17, 2.91). We decided to use the results in Figure 1 in the health economic modelling.

Figure 1

Study or sub-category	Antidepressant n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Tricyclic antidepressant s including Myren CCT					
Boerner et al 1988	26/38	22/37		24.06	1.15 [0.82, 1.62]
Myren CCT 1982	25/30	21/31		22.30	1.23 [0.92, 1.65]
Vij et al 1991	11/21	5/23		5.15	2.41 [1.00, 5.79]
Subtotal (95% CI)	89	91		51.51	1.31 [1.04, 1.64]
Total events: 62 (Antidepressant), 48 (Placebo)					
Test for heterogeneity: Chi² = 2.59, df = 2 (P = 0.27), I² = 22.9%					
Test for overall effect: Z = 2.34 (P = 0.02)					
02 SSRI antidepressants vs placebo or usual care (Creed)					
Creed et al 2003	49/74	30/80		31.12	1.77 [1.27, 2.45]
Kuiken 2003	9/17	7/17		7.56	1.29 [0.62, 2.65]
Tabas 2004	19/30	10/36		9.81	2.28 [1.26, 4.13]
Subtotal (95% CI)	121	133		48.49	1.80 [1.38, 2.34]
Total events: 77 (Antidepressant), 47 (Placebo)					
Test for heterogeneity: Chi² = 1.45, df = 2 (P = 0.48), I² = 0%					
Test for overall effect: Z = 4.32 (P < 0.0001)					
Total (95% CI)	210	224		100.00	1.55 [1.30, 1.84]
Total events: 139 (Antidepressant), 95 (Placebo)					
Test for heterogeneity: Chi² = 8.74, df = 5 (P = 0.12), I² = 42.8%					
Test for overall effect: Z = 4.86 (P < 0.00001)					

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours Drug

b) Global symptom score

One study with 28 patients reported the global symptom score, but no details of the scale were given. There was no significant difference between interventions.

Figure 2

Study or sub-category	N	Antidepressant Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 Tricyclic antidepressant (amitriptyline) 50mg once daily							
Steinhart 1982	14	60.50(19.99)	14	69.36(20.92)		100.00	-8.86 [-24.02, 6.30]
Subtotal (95% CI)	14		14			100.00	-8.86 [-24.02, 6.30]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.15 (P = 0.25)							
Total (95% CI)	14		14			100.00	-8.86 [-24.02, 6.30]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.15 (P = 0.25)							

-100 -50 0 50 100
Favours antidepress   Favours placebo
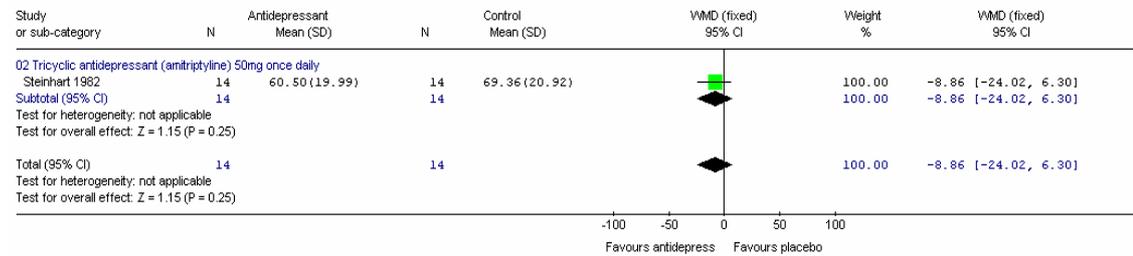
2. Individual symptoms

a) Pain

i. Number of patients with less pain

Three studies reported this outcome. Two were tricyclics versus placebo (Vij 1991; Tanum and Malt 1996) and the other was an SSRI versus placebo (Tabas 2004). These were included as subgroups in a meta-analysis in 150 patients. Overall there was significant heterogeneity (I²=84%, p=0.004), which is attributed to the type of antidepressant.

There was a large significant difference in favour of tricyclics compared to placebo in the number of patients with reduced pain; RR 3.91 (95%CI 1.93, 7.93), with no heterogeneity (I²=0%, p=0.81), although the confidence interval was fairly wide. This corresponded to an NNT of 2 (95%CI 2, 4), for a placebo group rate of 16-18%. We noted that this analysis

included Tanum and Malt (1996) which had only 60% of patients with IBS, and; Vij (1991) which had 24% missing data in each arm.

There was no significant difference between the SSRI and placebo in one study in 66 patients (Tabas 2004); RR 0.88 (95%CI 0.54, 1.45).

Figure 3

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 04 Number of patients with less Pain

Study or sub-category	Antidepressant n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Tricyclic					
Tanum 60% IBS 1996	19/25	4/22		17.41	4.18 [1.68, 10.42]
Vij et al 1991	10/18	3/19		11.94	3.52 [1.15, 10.76]
Subtotal (95% CI)	43	41		29.35	3.91 [1.93, 7.93]
Total events: 29 (Antidepressant), 7 (Placebo)					
Test for heterogeneity: Chi² = 0.05, df = 1 (P = 0.81), I² = 0%					
Test for overall effect: Z = 3.78 (P = 0.0002)					
02 SSRI (+ high fibre in each arm)					
Tabas 2004	14/30	19/36		70.65	0.88 [0.54, 1.45]
Subtotal (95% CI)	30	36		70.65	0.88 [0.54, 1.45]
Total events: 14 (Antidepressant), 19 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.49 (P = 0.62)					
Total (95% CI)	73	77		100.00	1.77 [1.20, 2.63]
Total events: 43 (Antidepressant), 26 (Placebo)					
Test for heterogeneity: Chi² = 12.52, df = 2 (P = 0.002), I² = 84.0%					
Test for overall effect: Z = 2.85 (P = 0.004)					

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours drug

ii. Number of patients with pain

One study in 34 non-depressed patients (Kuiken 2003) reported the number of patients with 'significant pain'. There was no statistically significant difference between interventions, although the confidence interval was fairly wide.

Figure 4

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 21 Number of patients with pain

Study or sub-category	Antidepressant n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 SSRI versus placebo					
Kuiken 2003	9/17	13/17		100.00	0.69 [0.41, 1.16]
Subtotal (95% CI)	17	17		100.00	0.69 [0.41, 1.16]
Total events: 9 (Antidepressant), 13 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.39 (P = 0.17)					
Total (95% CI)	17	17		100.00	0.69 [0.41, 1.16]
Total events: 9 (Antidepressant), 13 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.39 (P = 0.17)					

0.1 0.2 0.5 1 2 5 10
Favours antidepress   Favours control

iii. Pain score

Four studies (Creed 2003; Tanum and Malt 1996; Myren 1982; Myren 1984) reported a type of pain score, all appeared to use a VAS of 100 mm or 10cm. Myren (1982) recorded an abdominal obstruction discomfort measurement, which the GDG decided was a different

outcome to pain. Myren (1984) gave means and p values only. We decided to combine the first two studies in a meta-analysis by subgroup (Creed 2003 compared an SSRI with usual care). We also noted that Tanum and Malt (1996) had only 60% of patients with IBS.

Figure 5

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 10 Pain Score

Study or sub-category	N	tricyclic Mean (SD)	N	placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Tricyclic versus placebo							
Tanum 60% IBS 1996	25	16.90(24.30)	22	42.80(20.90)		33.40	-25.90 [-38.82, -12.98]
Subtotal (95% CI)	25		22			33.40	-25.90 [-38.82, -12.98]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.93 (P < 0.0001)							
02 SSRI vs usual care							
Creed et al 2003	74	-20.60(30.19)	79	-11.40(27.38)		66.60	-9.20 [-18.35, -0.05]
Subtotal (95% CI)	74		79			66.60	-9.20 [-18.35, -0.05]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.97 (P = 0.05)							
Total (95% CI)	99		101			100.00	-14.78 [-22.25, -7.31]
Test for heterogeneity: Chi² = 4.27, df = 1 (P = 0.04), I² = 76.6%							
Test for overall effect: Z = 3.88 (P = 0.0001)							

-100   -50   0   50   100

Favours treatment   Favours control

At the end of treatment, meta-analysis showed significant heterogeneity (I^2=77%, p=0.04), with different effect sizes being found for the two studies. This may be an effect of type of antidepressant, type of comparator, severity of IBS, or it may be that Tanum and Malt (1996) overestimated the effect because there were only 60% of patients with IBS. Individually, there was a statistically significant difference for the Tanum and Malt (1996) study (tricyclic versus placebo in 60% IBS): mean difference -25.90 (95%CI -38.82, -12.98) and for the Creed (2003) study: mean difference -9.20 (95%CI -18.35, -0.05).

Boerner (1988) compared a tricyclic with placebo, and recorded the *median* improvement in pain on a scale of 0 to 4. There was a difference in median change score of 0.3 units, in favour of the tricyclic antidepressant, which was reported to be statistically significant at the level of p<0.05.

b) Bowel habit

i. Number of patients with improvement in bowel habit

Meta-analysis of two studies in 110 patients showed a statistically significant increase in the number of patients with improvement in bowel habit; RR 1.92 (95%CI 1.19, 3.07), with no heterogeneity between the tricyclics and SSRI subgroups (I^2=0; p=0.51) corresponding to an NNT of 4 (95%CI 3, 13). The confidence interval was fairly wide. The results for individual classes of antidepressant were not statistically significant, although the antidepressant was favoured; Vij (1991) (tricyclics) had a wide confidence interval and Tabas (2004) (SSRI) was fairly wide.

Figure 6

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 11 Improved bowel habit (No of Patients)

Study or sub-category	Treatment n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Tricyclic vs placebo					
Vij et al 1991	11/21	5/23		30.43	2.41 [1.00, 5.79]
Subtotal (95% CI)	21	23		30.43	2.41 [1.00, 5.79]
Total events: 11 (Treatment), 5 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.97 (P = 0.05)					
02 SSRI vs placebo					
Tabas 2004	17/30	12/36		69.57	1.70 [0.97, 2.97]
Subtotal (95% CI)	30	36		69.57	1.70 [0.97, 2.97]
Total events: 17 (Treatment), 12 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.86 (P = 0.06)					
Total (95% CI)	51	59		100.00	1.92 [1.19, 3.07]
Total events: 28 (Treatment), 17 (Control)					
Test for heterogeneity: Chi² = 0.44, df = 1 (P = 0.51), I² = 0%					
Test for overall effect: Z = 2.70 (P = 0.007)					

0.1 0.2 0.5 1 2 5 10
Favours Control   Favours Treatment

c) Bloating

i. Number of patients with less bloating

Two studies recorded the number of patients with less bloating, one a tricyclic (Boerner 1998) and the other using an SSRI (Tabas 2004), both compared with placebo. Meta-analysis could not be carried out because Boerner (1998) reported the median.

Boerner (1988) recorded the median improvement in the feeling of fullness on a scale of 0 to 4. There was a difference in median change score of 0.23 units, in favour of the tricyclic, but this was reported to be not statistically significant.

Tabas (2004), comparing SSRI with placebo, did not demonstrate a significant effect on the number of patients with less bloating, but the confidence interval was fairly wide so this conclusion was uncertain.

Figure 7

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 07 Less bloating (No of Patients)

Study or sub-category	Treatment n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Tabas 2004	11/30	14/36		100.00	0.94 [0.51, 1.76]
Total (95% CI)	30	36		100.00	0.94 [0.51, 1.76]
Total events: 11 (Treatment), 14 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.18 (P = 0.85)					

0.1 0.2 0.5 1 2 5 10
Favours treatment   Favours control
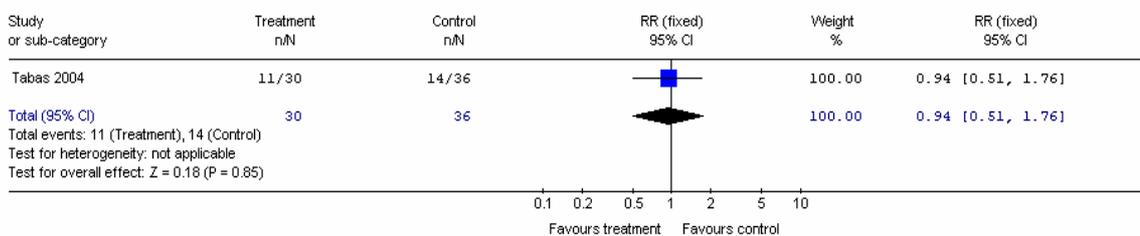
ii. Number of patients with bloating

One study (Kuiken 2003) reported the number of patients with bloating in 34 non-depressed patients. There was no significant difference between the SSRI and placebo, although the confidence interval was fairly wide. We note that the study was sponsored by Eli Lilly

(manufacturers of fluoxetine).

Figure 8

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 23 Number of patients with bloating

Study or sub-category	Antidepressant n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 SSRI versus placebo					
Kuiken 2003	10/17	8/17		100.00	1.25 [0.66, 2.38]
Subtotal (95% CI)	17	17		100.00	1.25 [0.66, 2.38]
Total events: 10 (Antidepressant), 8 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.68 (P = 0.50)					
Total (95% CI)	17	17		100.00	1.25 [0.66, 2.38]
Total events: 10 (Antidepressant), 8 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.68 (P = 0.50)					

0.1 0.2 0.5 1 2 5 10
Favours antidepress   Favours control

3. Mental health

a) Psychological Distress

This was measured by the SCL-90 global severity index (90 item, 5 point rating scale; range 90 to 450, high = bad). There was a small, statistically significant difference between SSRI (paroxetine) and usual care at 3 months, WMD: -0.28 (95%CI -0.43, -0.09), favouring the antidepressant.

Figure 9

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 06 Psychological Distress (3months)

Study or sub-category	N	Treatment Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 SSRI vs usual care							
Creed et al 2003	67	-0.24 (0.49)	70	0.02 (0.50)		100.00	-0.26 [-0.43, -0.09]
Subtotal (95% CI)	67		70			100.00	-0.26 [-0.43, -0.09]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.07 (P = 0.002)							
Total (95% CI)	67		70			100.00	-0.26 [-0.43, -0.09]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.07 (P = 0.002)							

-10 -5 0 5 10
Favours treatment   Favours control

b) Depression score

One study (Myren 1982, CCT) recorded depression on a VAS of 0-10cm for tricyclics versus placebo, in which a high score indicated increased depression. There was statistically significantly less depression for the patients taking antidepressants. We noted that this study, at baseline, showed low scores on depression and anxiety scales.

Figure 10

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 14 Depression 0-10cm VAS

Study or sub-category	N	Antidepressant Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Tricyclic vs placebo							
Myren CCT 1982	30	0.42(2.30)	31	2.25(2.34)		100.00	-1.83 [-2.99, -0.67]
Subtotal (95% CI)	30		31			100.00	-1.83 [-2.99, -0.67]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.08 (P = 0.002)							
Total (95% CI)	30		31			100.00	-1.83 [-2.99, -0.67]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.08 (P = 0.002)							

-10   -5   0   5   10
Favours antidepress   Favours control

c) Anxiety

Two studies reported anxiety; one (Myren 1982, CCT) compared tricyclics versus placebo on a VAS of 0 to 10 and the other (Tabas 2004) reported the number of patients with anxiety for the comparison SSRI with placebo. The former study, at baseline, showed low scores on depression and anxiety scales, and the latter reported that 27/81 (33%) had a score greater than 10 on the Beck Depression Inventory.

i. Anxiety score

There was no significant difference in the degree of anxiety.

Figure 11

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 15 Anxiety 0-10cm VAS

Study or sub-category	N	Antidepressant Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Tricyclic vs placebo							
Myren CCT 1982	30	0.99(1.53)	31	0.85(1.56)		100.00	0.14 [-0.64, 0.92]
Subtotal (95% CI)	30		31			100.00	0.14 [-0.64, 0.92]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.35 (P = 0.72)							
Total (95% CI)	30		31			100.00	0.14 [-0.64, 0.92]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.35 (P = 0.72)							

-10   -5   0   5   10
Favours antidepress   Favours control

ii. Number of patients with anxiety

The confidence interval was wide and the result was not statistically significant.

Figure 12

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 16 Number of patients with less anxiety

Study or sub-category	antidepressant n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 SSRI vs placebo					
Tabas 2004	13/30	8/36		100.00	1.95 [0.93, 4.07]
Subtotal (95% CI)	30	36		100.00	1.95 [0.93, 4.07]
Total events: 13 (antidepressant), 8 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.78 (P = 0.08)					
Total (95% CI)	30	36		100.00	1.95 [0.93, 4.07]
Total events: 13 (antidepressant), 8 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.78 (P = 0.08)					

```
          0.1  0.2  0.5   1    2    5   10
          Favours control   Favours antidepress
```

4. Quality of life

Two studies reported quality of life measures, both for SSRIs (Creed 2003; Tabas 2004).

a) IBS-QoL

One study (Tabas 2004) comparing SSRI and placebo examined the change compared with baseline in three components of IBS-QoL: food avoidance; work function; social function. Tabas (2004) found the following differences for SSRI versus placebo:

- Food avoidance subscale – improvement of 12.7%; p value = 0.03 (statistically significant)
- Work function subscale – improvement of 2.1%; p value = 0.08 (not significant)
- Social function subscale – improvement of 13.4%; p value = 0.76 (not significant).

b) Change in SF36 Mental component score at 3 months

There was no significant difference in the mental health quality of life SF-36 score (scale 0 to 100, high=good) at 3 months between SSRI (paroxetine) and usual care, although there were 29% who discontinued treatment and 32% loss to follow up in the paroxetine arm. This conclusion does not agree with the p values reported in Creed (2003) (p=0.007).

Figure 13

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 12 Quality of Life - change from baseline SF36 mental health (3months)

Study or sub-category	N	antidepressant Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Creed et al 2003	59	3.70 (13.44)	63	-0.50 (12.70)		100.00	4.20 [-0.45, 8.85]
Total (95% CI)	59		63			100.00	4.20 [-0.45, 8.85]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.77 (P = 0.08)							

```
         -100   -50    0    50   100
          Favours control   Favours antidepress
```
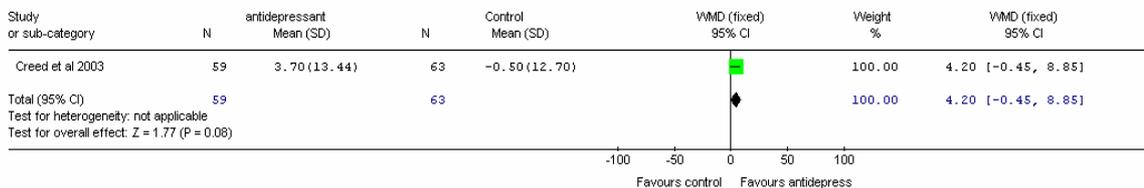
c) Change in SF36 physical health score at 3 months

The change in the physical health component also showed no significant difference at 3 months. This did not agree with the p value reported in Creed (2003) (p=0.24).

Figure 14

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 17 Quality of Life - change from baseline SF36 physical health

Study or sub-category	N	antidepressant Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Creed et al 2003	59	2.40(10.22)	63	-0.50(6.98)		100.00	2.90 [-0.23, 6.03]
Total (95% CI)	59		63			100.00	2.90 [-0.23, 6.03]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.82 (P = 0.07)

```
        -100   -50    0    50   100
      Favours control  Favours antidepress
```

B. Comparison of different doses of Tricyclic antidepressant

One study (Myren 1984) compared different doses of Trimipramine in 428 participants.

1. Global symptoms

The study reported the 'total effect of treatments' in the opinion of the physicians. Medians with their 95% confidence intervals were reported for each group. We used this to calculate the standard deviation, which we used with the median value to compare groups. *Figure 16* shows there was no significant difference between any of the doses.

2. Individual symptoms

a) Number of patients with abdominal pain

The study reported a statistically significant reduction in pain in patients that were given 50mg per day either as a single dose or in two divided doses (p < 0.01). Patients taking 35mg in a single dose had significantly less pain than placebo (p< 0.05), and for patients taking 30mg per day in three divided doses there was stated to be no difference between the drug and placebo (p= 0.10).

Figure 15

Study or sub-category	N	Antidepressant 1 Mean (SD)	N	antidepressant 2 Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 50mg at bedtime vs 10mg in morning + 40mg in evening							
Myren 1984	85	6.00(5.17)	89	6.00(4.57)		100.00	0.00 [-1.45, 1.45]
Subtotal (95% CI)	85		89			100.00	0.00 [-1.45, 1.45]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.00 (P = 1.00)							
02 50mg at bedtime vs 35mg at bedtime							
Myren 1984	85	6.00(5.17)	86	5.80(7.57)		100.00	0.20 [-1.74, 2.14]
Subtotal (95% CI)	85		86			100.00	0.20 [-1.74, 2.14]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.20 (P = 0.84)							
03 50mg at bedtime vs 10mg three times daily							
Myren 1984	85	6.00(5.17)	69	5.00(4.66)		100.00	1.00 [-0.55, 2.55]
Subtotal (95% CI)	85		69			100.00	1.00 [-0.55, 2.55]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.26 (P = 0.21)							
04 10mg in morning + 40mg in evening vs 35mg at bedtime							
Myren 1984	89	6.00(4.57)	86	5.80(7.57)		100.00	0.20 [-1.66, 2.06]
Subtotal (95% CI)	89		86			100.00	0.20 [-1.66, 2.06]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.21 (P = 0.83)							
05 10mg in morning + 40mg in evening vs 10mg three times daily							
Myren 1984	89	6.00(4.57)	69	5.00(4.66)		100.00	1.00 [-0.45, 2.45]
Subtotal (95% CI)	89		69			100.00	1.00 [-0.45, 2.45]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.35 (P = 0.18)							
06 35mg at bedtime vs 10mg three times daily							
Myren 1984	86	5.80(7.57)	69	5.00(4.66)		100.00	0.80 [-1.14, 2.74]
Subtotal (95% CI)	86		69			100.00	0.80 [-1.14, 2.74]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.81 (P = 0.42)							

-10   -5   0   5   10
Favours antidep 2   Favours antidep 1

ADVERSE EFFECTS

The evidence on adverse effects of tricyclics and SSRIs is provided in the adverse effects review (section 8.5.2).

ECONOMIC LITERATURE FOR TRICYCLICS AND SSRIS

One relevant health economic analysis was identified on the cost-effectiveness of SSRIs in the treatment of IBS (Creed 2003), but none were identified which considered the cost-effectiveness of tricyclics. Creed (2003) was a trial based economic evaluation conducted in the UK which recruited patients from secondary and tertiary care with severe IBS. This study aimed to assess whether an SSRI (paroxetine) would be superior to usual care in reducing abdominal pain and improving quality of life and whether these improvements could be achieved at no additional cost due to treatment costs being offset by reduced health care costs. (It also included a comparison of psychotherapy with usual care). The patient population considered were secondary and tertiary care patients with severe IBS who had not responded to usual treatment. The included patients had a mean duration of IBS of 8 years. This study was considered to be relevant to patients with refractory IBS only. The SSRI intervention consisted of 20mg of paroxetine daily for 3 months which was prescribed and monitored either by the patient's gastroenterologist or their GP. After three months, patients in the SSRI arm returned to their GP and received usual care for one year during which time they were followed-up. In the comparator arm patients received usual care from either their gastroenterologist or their GP for the three month treatment period and the following year of follow-up. The primary outcome was abdominal pain measured on a VAS of severity with secondary outcomes considering days with

pain, overall change in symptoms and HRQofL measured by the SF-36. Direct costs of health care and intervention costs were recorded.

The number of people with an improvement in global symptoms was significantly higher for SSRI at the end of treatment compared to usual care. The clinical outcomes from this trial have been summarised in detail in the clinical effectiveness review. Direct health care costs were not significantly increased for SSRI compared to usual care during the intervention period or the following year. However, the results for the follow-up period may have been partially confounded as patients in the usual care arm were also allowed SSRIs.

This study was a partial economic evaluation as it did not assess the incremental cost of any benefit achieved in the form of a cost-effectiveness ratio. The evidence provided by this study was considered to be indirect as the patients were recruited from secondary and tertiary care and costs may differ for refractory patients managed in primary care. No potential areas of significant bias were identified but the results for the follow-up year were considered to be partly confounded by the use of SSRIs in the comparator arm during follow-up. Direct health care costs were not significantly increased by SSRIs during the intervention period. However, the study was powered to detect a specific change in clinical rather than cost outcomes. As this study did not provide an estimate of the cost per QALY for SSRIs compared to usual care, it was not particularly useful in determining whether recommending SSRIs for use in the NHS would result in the efficient use of NHS resources.

COST-EFFECTIVENESS ANALYSIS FOR TRICYCLICS AND SSRIS

This section describes the health economic analysis undertaken to inform recommendations on the use of tricyclics and SSRIs as long-term maintenance therapies in IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

The general approach is the same as for other maintenance therapies except that the clinical pathway has been modified to allow for gradual dose increases and an additional follow-up appointment at 12 weeks. The clinical pathway was modified as follows:

- Treatment is initiated at 10mg with dose increases of 10mg no more frequently than every 2 weeks. Patients are encouraged to increase the dose until an effective dose has been established or side-effects become problematic up to a maximum dose of 30mg for tricyclics and 20 mg for SSRIs. (Alternative maximum doses of 50mg for tricyclics and 40mg for SSRIs are considered in a sensitivity analysis).
- Follow-up appointments are required every month until an effective dose has been established and a further follow-up appointment is required 12 weeks after that point. The usual 6 monthly review is carried out as for other pharmacological interventions. A sensitivity

analysis in which follow-up appointments are twice as frequent was considered in order to estimate whether monitoring patients more intensively would have a significant impact on cost-effectiveness.

- Patients who do not respond to the maximum dose are switched to an alternative tricyclic or SSRI. An incremental analysis was carried out to determine whether allowing patients who do not respond to switch to a second treatment, provides sufficient additional benefit to be cost-effective given the additional costs.

- The model assumes that the response seen in the clinical trial is gradually achieved over the range of doses considered, even if the maximum dose considered is not as high as the trial dose. For example, the RCT for tricyclics versus placebo used a range of doses from 30 to 75mg, compared to the maximum dose of 30mg used in the basecase model. In the economic model, we assumed that equal numbers of patients respond to the 10mg, 20mg and 30mg doses with the total number of responders equal to that predicted by applying the RR from the RCTs to the response rate in the no treatment arm. Sensitivity analyses were carried out assuming all patients respond to the lowest or highest dose.

Tricyclic or SSRI therapy initiated
Initial dose 10mg, patient controlled dose increases of 10mg no more frequently than every 2 weeks.
Follow-up appointment booked for 1 month later

1 month later

GP assesses if patient has responded to treatment

Responded

Didn't respond

Patient continues on treatment and follow-ups booked for 12 weeks after effective dose established and 6 months after starting TC or SSRI

Has the max dose been tried (30mg tricyclics and 20mg SSRI)

Yes

No

Switch to alternative TC or SSRI
Return to top

Continue dose increases at 10mg every 2 weeks and appointment booked for 1 month later

1 month later

6 months after first TC or SSRI

GP assesses whether TC or SSRI therapy is still appropriate

6 months later

Still appropriate

No longer appropriate

TC or SSRI therapy continued for further 6 months

TC or SSRI therapy discontinued

The following assumptions have been made regarding the effectiveness of tricyclics and SSRIs based on the clinical effectiveness review:

- The following tricyclics were treated as a class of interventions with the same clinical effectiveness as there was insufficient evidence to demonstrate a significant difference in effectiveness between them: trimipramine, amitriptyline, doxepin.
- There was evidence that there is no difference in the effectiveness of tricyclics for doses above 30mg compared to 30mg (Myren 1984), so this was the maximum dose modelled in the basecase. A more conservative assumption in which doses of 50mg are required to achieve the effectiveness seen in the RCTs have been considered in a sensitivity analysis.
- The following SSRIs were treated as a class of interventions with the same clinical effectiveness as there was insufficient evidence to demonstrate a significant difference in effectiveness between them: fluoxetine, paroxetine. A maximum dose of 20mg was modelled in the basecase as this was the most commonly prescribed dose in the trials, but a sensitivity analysis was carried out on doses of up to 40mg using the dose distribution from Tabas (2004).
- The studies included in the clinical effectiveness review did not stratify results by IBS subtype, so it was not possible to estimate the effectiveness for each of the subtypes separately. Therefore, it was assumed that these interventions are equally effective across all IBS subtypes.

Modelled response rates

Figures 18 to 23 below give the modelled response rates for six different strategies for prescribing tricyclics and SSRIs for the management of chronic pain in IBS.

1. Tricyclic up to 30mg
2. Tricyclic up to 30mg followed by second tricyclic up to 30mg
3. Tricyclic up to 30mg followed by SSRI up to 20mg
4. SSRI up to 20mg
5. SSRI up to 20mg followed by tricyclic
6. SSRI up to 20mg followed by second SSRI up to 20mg

The RR of response for SSRIs versus placebo (or usual care) was much higher than the RR of response for tricyclics versus placebo (see Table 1) giving much larger increases in the number responding at each dose. However, as there was no head-to-head comparison of SSRIs and tricyclics, it was not possible to tell whether the apparent superiority of SSRIs is simply because they have been tested in a population that is more likely to respond to a pharmacological intervention. Indirect comparisons between SSRIs and tricyclics should be interpreted with caution due to the potential for bias.

Figure 17: Response rates for a tricyclic (up to 30mg) compared to no treatment
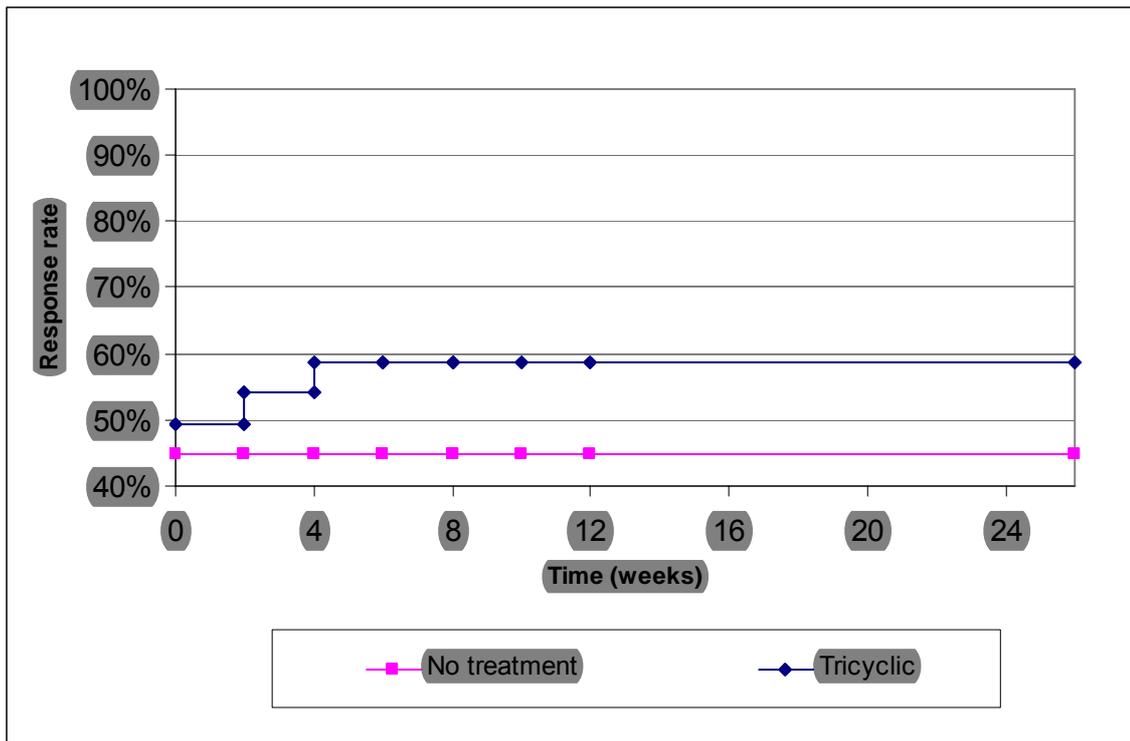
Figure 18: Response rates for a tricyclic up to 30mg followed by second tricyclic up to 30mg in non-responders compared to no treatment
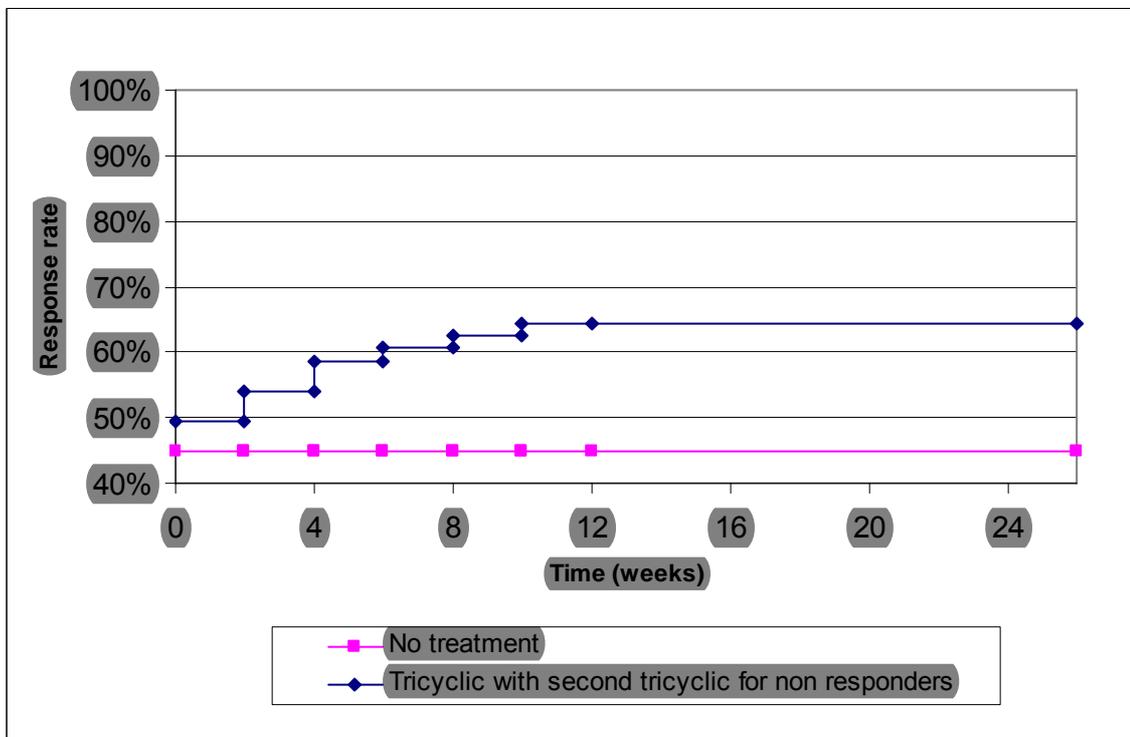
Figure 19: Response rates for a tricyclic up to 30mg followed by an SSRI up to 20mg in non-responders compared to no treatment
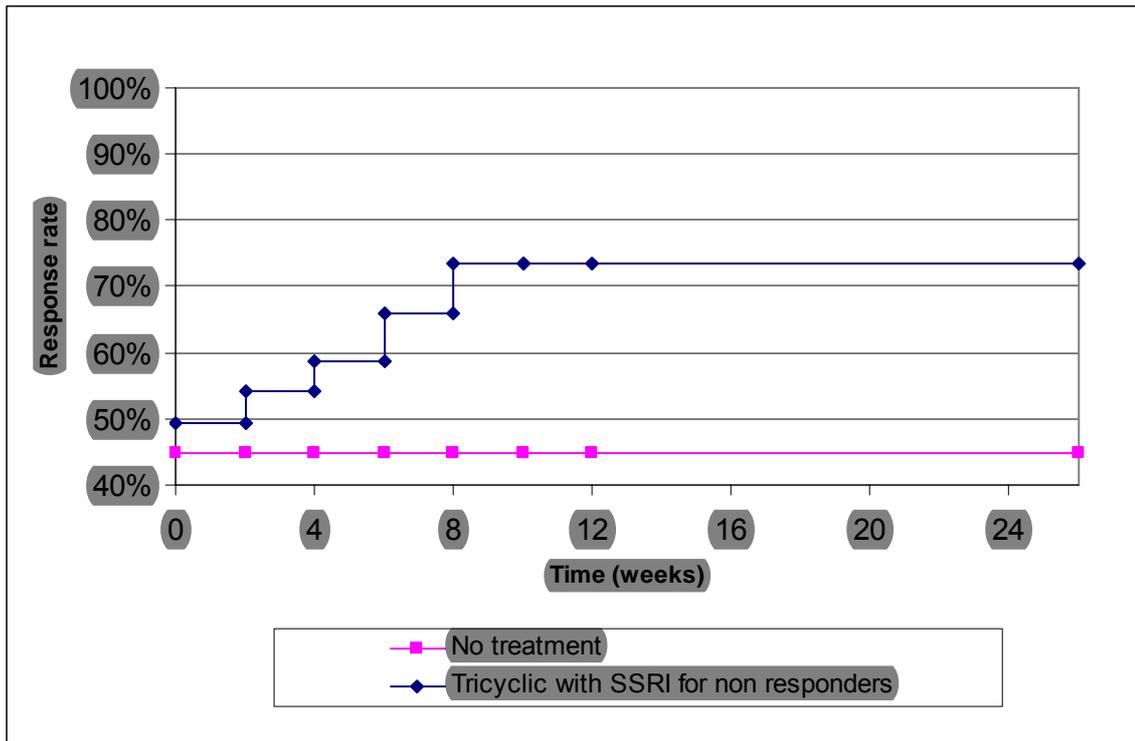
Figure 20: Response rates for an SSRI up to 20mg compared to no treatment
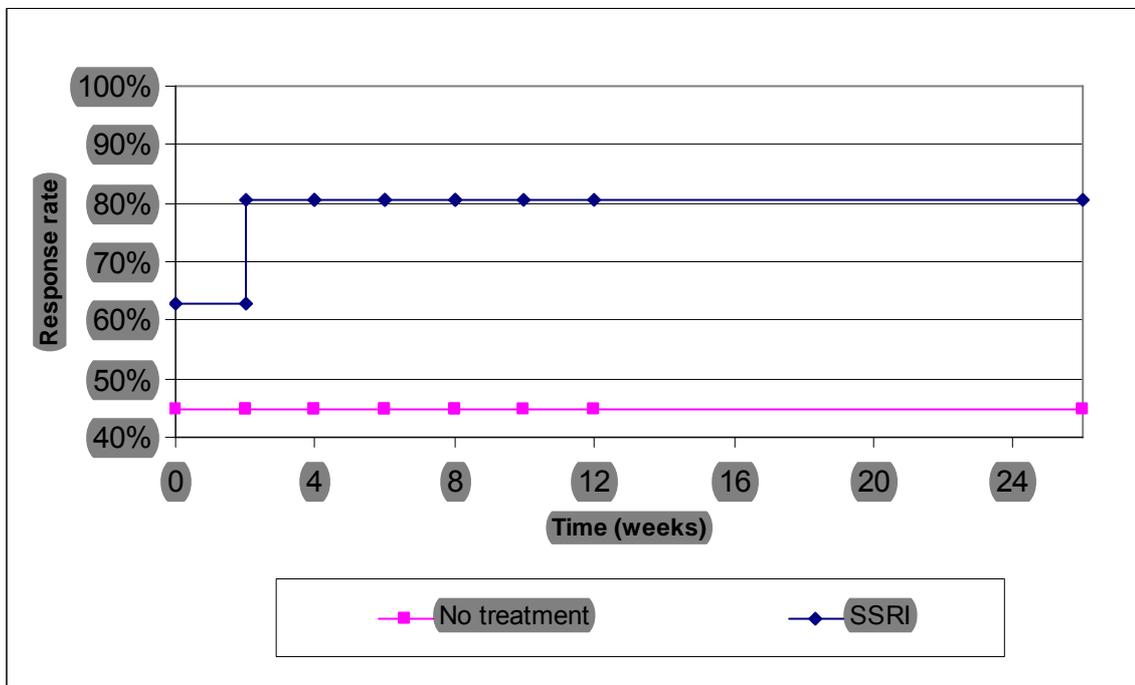
Figure 21: Response rates for an SSRI up to 20mg followed by a tricyclic up to 30mg in non-responders compared to no treatment
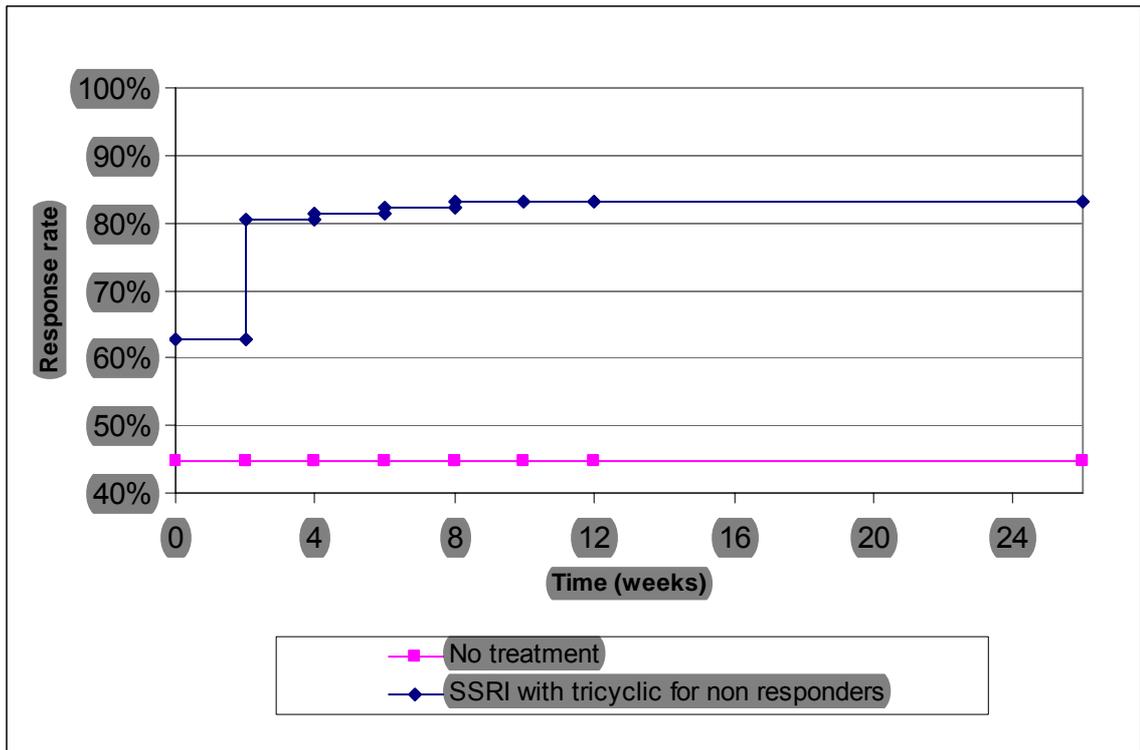
Figure 22: Response rates for an SSRI (up to 20mg) followed by a second SSRI in non-responders compared to no treatment
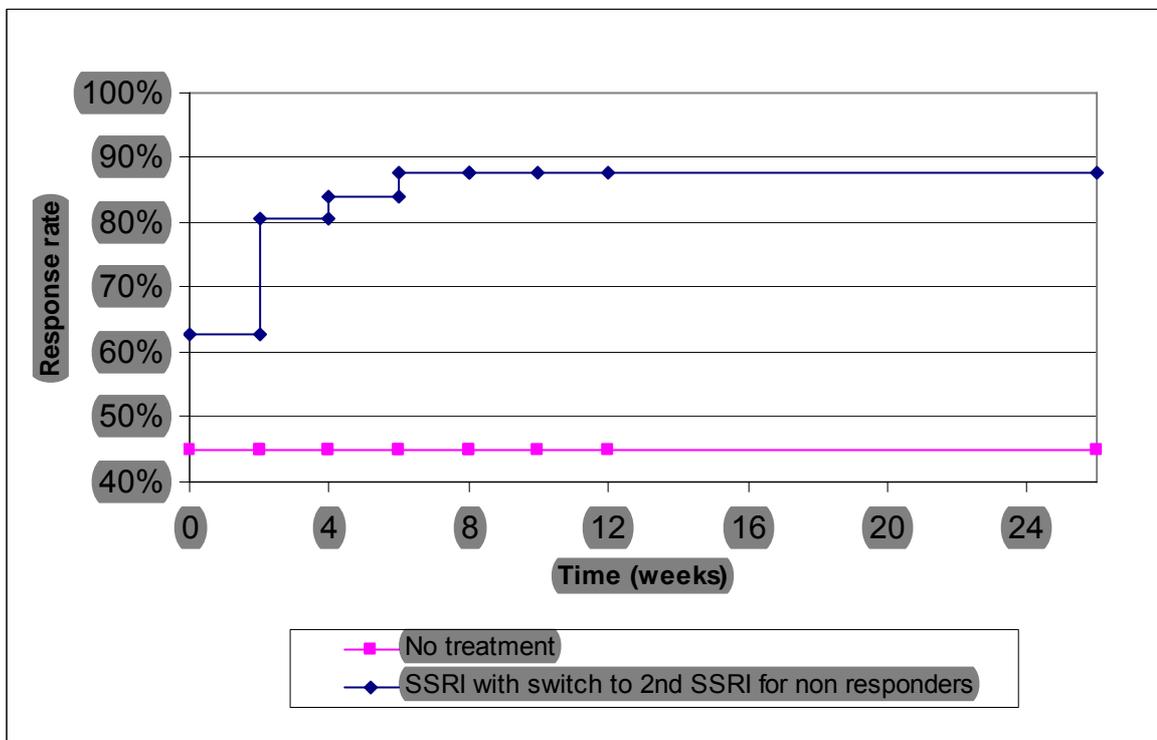
Amitriptyline tablets were the lowest cost tricyclic at doses of 10mg to 20mg. Doxepin had a lower cost per 10mg but the smallest tablet size is 50mg, so it was not considered practical to assume that doxepin is commonly prescribed at a doses of 10 to 30mg per day. It was assumed in the model that 10mg amitriptyline tablets (non-proprietary) are prescribed as the first line tricyclic, with trimipramine (Surmontil®) prescribed second line. The highest cost tricyclic was amitriptyline solution (non-proprietary) at £5.03 per month for a dose of 10mg per day, so this preparation was used in the high drug cost sensitivity analysis.

Of the two SSRIs considered, the lowest cost preparation is fluoxetine (non-proprietary) and it was assumed that this is prescribed first line, with the lowest cost paroxetine preparation (non-proprietary) prescribed second line. A sensitivity analysis was carried out assuming that the highest cost preparation is prescribed (Branded liquid preparations of fluoxetine and paroxetine at £14.40 and £9.62 per month respectively).

Table 1: Intervention specific parameters – Tricyclics

Description	Value			Evidence
RR of response for TCA vs placebo	1.31			Meta-analysis of RCT evidence for improvement in global symptoms
Maximum number of switches considered	1 to second tricyclic or SSRI			Assumption based
Drug costs				
Intervention	Dose per day	Cost per month*(assuming lowest cost preparation)		
Amitriptyline	10mg to 30mg	£1.43 to £4.30		
Doxepin		£1.24 to £3.72		not avail under 50mg
Trimipramine		£3.87 to £11.61		

* British National Formulary (Joint Formulary Committee 2007)

Table 2: Intervention specific parameters – SSRIs

Description	Value			Evidence
RR of response for TCA vs placebo	1.80			Meta-analysis of RCT evidence for improvement in global symptoms
Maximum number of switches considered	1 to Tricyclic or second SSRI			Assumption based
Drug costs				
Intervention	Dose per day	Cost per month*(assuming lowest cost preparation)		
Fluoxetine	10mg to 20mg	£0.75 to £1.50		
Paroxetine	10mg to 20mg	£3.05 to £6.10		

* British National Formulary (Joint Formulary Committee 2007)

RESULTS

Tricyclic up to 30mg but no further treatment if no response

A strategy of using up to 30mg of tricyclic is estimated to provide 0.46 additional QALYs (difference between QALYs gained for intervention and no treatment), at a cost of £4,459, compared to no treatment for a cohort of 100 patients over a 6 month period, provided treatment is not continued in those who do not respond to a dose of 30mg. This strategy has a cost per QALY of £9,762, compared to no treatment, in the first 6 months after initiating treatment. Treatment can be continued in the next 6 months for a cost per QALY of £3,395 provided that treatment is reviewed every 6 months and discontinued in patients who no longer experience a therapeutic benefit, either due to lack of effectiveness or a change in their symptom profile.

The probabilistic sensitivity analysis estimates the uncertainty in the cost-effectiveness due to uncertainty surrounding the efficacy estimate, the utility gain, and the response rate in the no treatment arm. The results of this analysis are presented in Figure 24 as a CEAC, which shows the probability of the cost per QALY falling under various thresholds. The CEAC shows that a strategy of prescribing up to 30mg of tricyclic has a 76% probability of being under £20K compared to no treatment over a 6 month timeframe.

Figure 23: CEAC for up to 30mg of tricyclic compared to no treatment
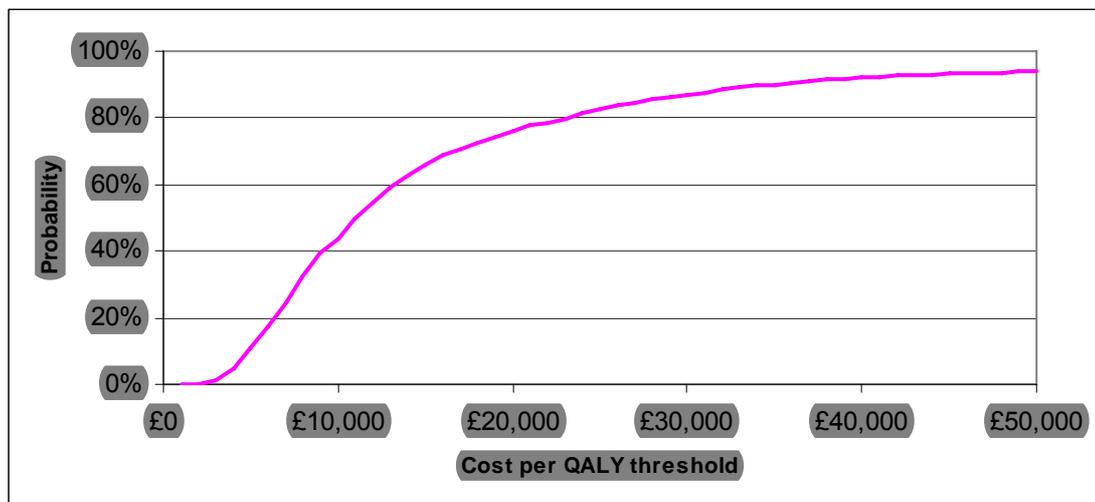
Univariate sensitivity results for up to 30mg of tricyclics

The basecase cost-effectiveness estimates assume that an equal number of patients respond to the 10mg, 20mg and 30mg doses. However, if we assume that patients only respond after reaching the 30mg dose, then the cost per QALY is higher at £11,296. If patients are allowed to increase their dose up to 50mg in order to achieve a response and we assume that no patient responds to a lower dose then the cost per QALY is significantly higher at £17,937.

If the highest cost preparation is prescribed instead of the lowest cost preparation, then the cost per QALY is increased to £14,022. If the GP follow-up is twice as frequent as modelled in the

pathway (i.e. every two weeks until an effective dose is established instead of every four weeks and two follow-up appointments in the next 12 weeks instead of one), then the cost per QALY is increased to £17,826. These sensitivity analyses suggest that using tricyclics at doses of up to 30mg is likely to be cost-effective compared to no treatment, even if the costs of care are higher than estimated in the modelled care pathway.

The threshold analysis on the utility gain associated with an improvement in global symptoms shows that the cost per QALY would be over £20,000 if the utility gain is less than 0.035. For comparison, we used a utility gain of 0.071 in the basecase and a utility gain of 0.135 would be equivalent to a complete resolution of IBS symptoms.

Table 3: Results for up to 30mg tricyclic (amitriptyline) with no switches compared to no treatment in a cohort of 100 patients with IBS (all subtypes)

Scenario	No Treatment		Intervention		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	0	1.60	4,459	2.05	9,762
No response in no treatment arm	0	0	4,252	0.46	9,308
Response rate in no treatment arm from RCTs	0	1.69	4,487	2.18	9,266
None respond until 30mg dose	0	1.60	4,730	2.01	11,296
Gradual response up to 50mg	0	1.60	5,661	2.01	13,521
None respond until 50mg	0	1.60	6,145	1.94	17,937
Higher cost formulation (liq)	0	1.60	6,405	2.05	14,022
Follow-up twice as freq	0	1.60	£8,143	2.05	17,826
High utility gain of 0.135	0	3.02	4,459	3.89	5,156
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.034 and 0.035.				

Tricyclic up to 30mg with switch to a second tricyclic if no response to first

If patients who do not respond to 30mg of tricyclic are allowed to switch to a second tricyclic drug, and we assume that the second tricyclic is just as likely to be effective, then the cost rises to £5,858 per 100 patients over the first 6 months, but the total QALY gain is increased to 0.60. The cost per QALY of treating patients with up to two tricyclics to gain a response is £9,789 per QALY compared to no treatment, and £9,873 compared to stopping treatment after the first

tricyclic. Even if the chance of response to the second tricyclic is half the chance of response to the first, the cost per QALY is £10,912 compared to no treatment and £18,324 compared to stopping after a failure on the first tricyclic.

A probabilistic sensitivity analysis was carried out on the cost-effectiveness of allowing non-responders to switch to a second tricyclic compared to no further tricyclic therapy for those who do not respond to the first tricyclic. The CEAC in Figure 25 shows that a strategy of prescribing up to 30mg of tricyclic with a switch to a second tricyclic for non-responders has an 87% likelihood of being under £20K compared to no further treatment for non-responders to 30mg of tricyclic.

Figure 24: CEAC for up to 30mg a tricyclic with switch to a second tricyclic if no response compared to up to 30mg tricyclic without switch for non-responders

Tricyclic up to 30mg followed by an SSRI if no response to tricyclic
If patients who do not respond to 30mg of tricyclic are allowed to switch to an SSRI (up to 20mg), and we assume that the SSRI effectiveness is independent of the response to the tricyclic, then the cost rises to £5,648 per 100 patients over the first 6 months, but the total QALY gain is increased to 0.84. The cost per QALY of treating patients with a tricyclic at doses of up to 30mg followed by an SSRI of up to 20mg is £6,703 per QALY compared to no treatment, and £3,080 compared to stopping after the tricyclic. Even if the chance of response to the SSRI is half the chance of response seen in the SSRI trials, then the cost per QALY is £8,450 compared to no treatment and £5,342 compared to stopping after a failure on the tricyclic.

A probabilistic sensitivity analysis was carried out on the cost-effectiveness of allowing non-responders to switch to a SSRI compared to no further SSRI therapy for those who do not respond to the first tricyclic. The CEAC in Figure 26 shows that a strategy of prescribing up to

30mg of tricyclic with a switch to an SSRI for non-responders has a 95% likelihood of being under £10K compared to no further treatment for non-responders to 30mg of tricyclic.

Figure 25: CEAC for up to 30mg of tricyclic with switch to an SSRI if no response compared to up to 30mg of tricyclic without switch for non-responders
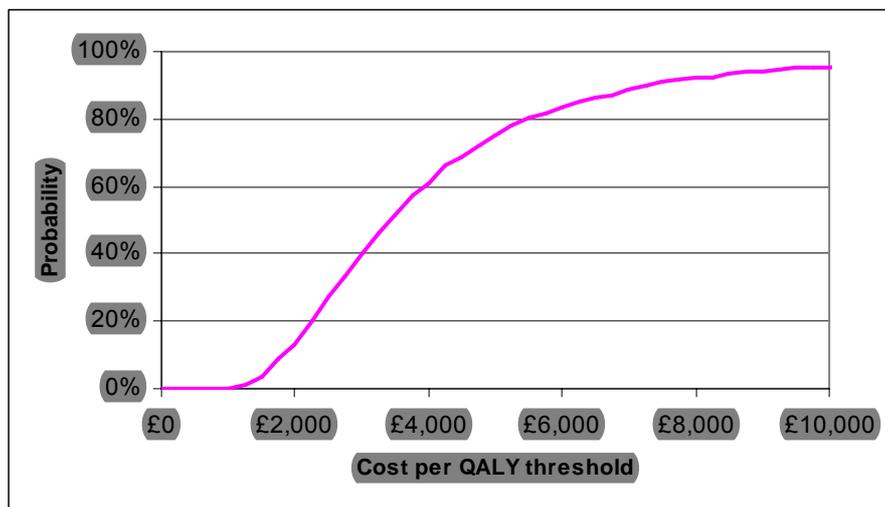
SSRI up to 20mg but no further treatment if no response

A strategy of using up to 20mg of SSRI is estimated to provide 1.23 additional QALYs, at a cost of £3,708, compared to no treatment for a cohort of 100 patients over a 6 month period, provided treatment is not continued in those who do not respond to a dose of 20mg. This strategy has a cost per QALY of £3,020 compared to no treatment, in the first 6 months after initiating treatment. Treatment can be continued in the next 6 months for a cost per QALY of £1,483 provided that treatment is reviewed every 6 months and discontinued in patients who no longer experience a therapeutic benefit, either due to lack of effectiveness or a change in their symptom profile.

These estimates assume that an equal number of patients respond to the 10mg and 20mg doses. However, if we assume that patients only respond after reaching the 20mg dose, then the cost per QALY is higher at £3,209. If patients are allowed to increase their dose up to 40mg in order to achieve a response, then the cost per QALY is significantly higher at £3,790 when using the dose distribution from Tabas (2004). However, if no patient responds until a dose of 40mg is reached then the cost per QALY is higher still at £4,698.

If the highest cost preparation is prescribed instead of the lowest cost preparation, then the cost per QALY is increased to £9,799. If the GP follow up is twice as frequent as modelled in the pathway (i.e every two weeks until an effective dose is established instead of every four weeks and two follow-up appointments in the next 12 weeks instead of one), then the cost per QALY is increased to £5,667. These sensitivity analyses suggest that using SSRIs at doses of up to

20mg is likely to be cost-effective compared to no treatment, even if the costs of care are higher than estimated in the modelled care pathway.

The threshold analysis on the utility gain associated with an improvement in global symptoms shows that the cost per QALY would be over £20,000 if the utility gain is less than 0.011. For comparison, we have used a utility gain of 0.071 in the basecase and a utility gain of 0.135 would be equivalent to a complete resolution of IBS symptoms.

Table 4: Results for up to 20mg SSRI (fluoxetine) with no switches compared to no treatment in a cohort of 100 patients with IBS (all subtypes)

Scenario	No Treatment		Intervention		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	0	1.60	3,708	2.82	3,020
No response in no treatment arm	0	0	2,747	1.23	2,237
Response rate in no treatment arm from RCTs	0	1.69	3,817	2.99	2,932
None respond until 20mg dose	0	1.60	3,783	2.78	3,209
Gradual response up to 40mg	0	1.60	4,430	2.77	3,790
None respond until 40mg	0	1.60	5,076	2.68	4,698
Higher cost formulation (liq)	0	1.60	12,032	2.82	9,799
Follow-up twice as freq	0	1.60	6,959	2.82	5,667
High utility gain of 0.135	0	3.02	3,708	5.35	1,595
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.010 and 0.011.				

A probabilistic sensitivity analysis was carried out on the cost-effectiveness of treatment with an SSRI (up to 20mg) compared to no treatment. The CEAC in Figure 27 shows the probability of the cost per QALY falling under various thresholds. The CEAC shows that a strategy of prescribing up to 20mg of SSRI has a 95% likelihood of being under £10K compared to no treatment over a 6 month timeframe.

Figure 26: CEAC for SSRI up to 20mg (with no switching for non-responders) compared to no treatment
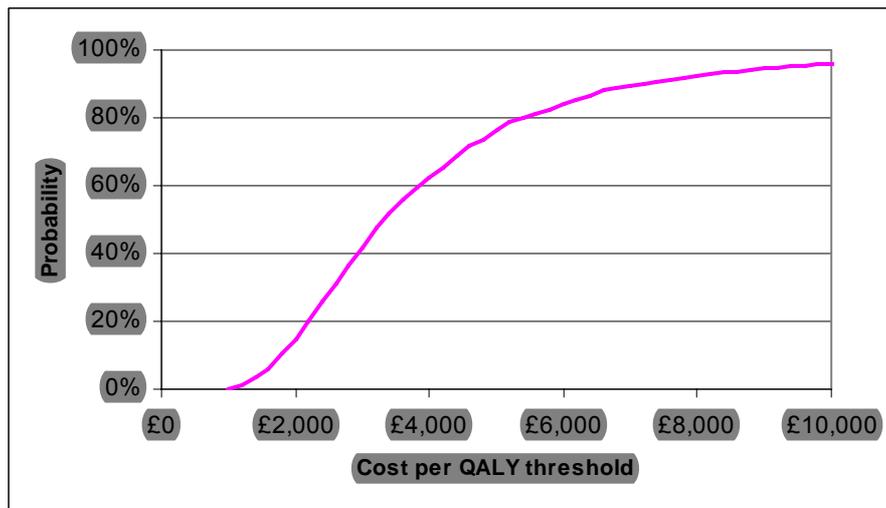
SSRI up to 20mg with switching to a tricyclic (up to 30mg) if no response to SSRI

If patients who do not respond to 20mg of SSRI are allowed to switch to a tricyclic drug (up to 30mg), and we assume that the tricyclic effectiveness is independent of the response to the SSRI, then the cost rises to £4,526 per 100 patients over the first 6 months, but the total QALY gain is increased to 1.30. The cost per QALY of allowing patients to try an SSRI followed by a tricyclic is £3,477 per QALY compared to no treatment, and £11,073 compared to stopping after the SSRI. Even if the chance of response to the tricyclic in patients who haven't responded to an SSRI is half the chance of response seen in the tricyclic trials, the cost per QALY is £3562 compared to no treatment and £21,574 compared to stopping after a failure on the first tricyclic. If the frequency of follow-up is twice that estimated in the basecase pathway, then the cost per QALY for treatment with an SSRI, followed by a TCA in non responders, is £6,523 compared to no treatment and £20,750 compared to stopping treatment after the SSRI.

A probabilistic sensitivity analysis was carried out on the cost-effectiveness of allowing non-responders to switch to a tricyclic compared to no further tricyclic therapy for those who do not respond to the SSRI. The CEAC in Figure 28 shows that a strategy of prescribing up to 20mg of SSRI with a switch to a tricyclic for non-responders has a 59.6% probability of being under £20K compared to no further treatment for non-responders to 20mg of SSRI. In the probabilistic analysis, there is a 12.8% probability that all of the patients respond to the first SSRI resulting in no further benefit to be gained by switching patients to a second SSRI. In drawing the CEAC we have assumed that the cost per QALY of allowing patients to switch would be above any reasonable threshold when no benefit can be achieved for that switch.

Figure 27: CEAC for up to 20mg of SSRI with switch to tricyclic if no response compared to up to 20mg SSRI without switch for non-responders

SSRI up to 20mg with a second SSRI (up to 20mg) if no response to first SSRI

If patients who do not respond to 20mg of SSRI are allowed to switch to a second SSRI (up to 20mg), and we assume that the effectiveness of the second SSRI is independent of the response to the first SSRI, then the cost rises to £4,677 per 100 patients over the first 6 months, but the total QALY gain is increased to 1.43. The cost per QALY of allowing patients to try an SSRI followed by a tricyclic is £3,275 per QALY compared to no treatment, and £4,843 compared to stopping after the first SSRI. Even if the chance of response to the second SSRI in patients who haven't responded to the first SSRI is half the chance of response seen in the SSRI trials, the cost per QALY is £3,361 compared to no treatment and £7,544 compared to stopping after a failure on the first SSRI.

A probabilistic sensitivity analysis was carried out on the cost-effectiveness of allowing non-responders to switch to a second SSRI compared to no further SSRI therapy for those who do not respond to the first SSRI. The CEAC in Figure 29 shows that a strategy of prescribing up to 20mg of SSRI with a switch to a second SSRI for non-responders has a 75% likelihood of being under £10K compared to no further treatment for non-responders to 20mg of SSRI. Again, there is a 12.8% probability that all of the patients respond to the first SSRI resulting in no further benefit to be gained by switching patients to a second SSRI. In drawing the CEAC we have assumed that the cost per QALY of allowing patients to switch would be above any reasonable threshold when there is no benefit can be achieved for that switch.

Figure 28: CEAC for up to 20mg SSRI with switch to a second SSRI if no response compared to up to 20mg SSRI without switch for non-responders
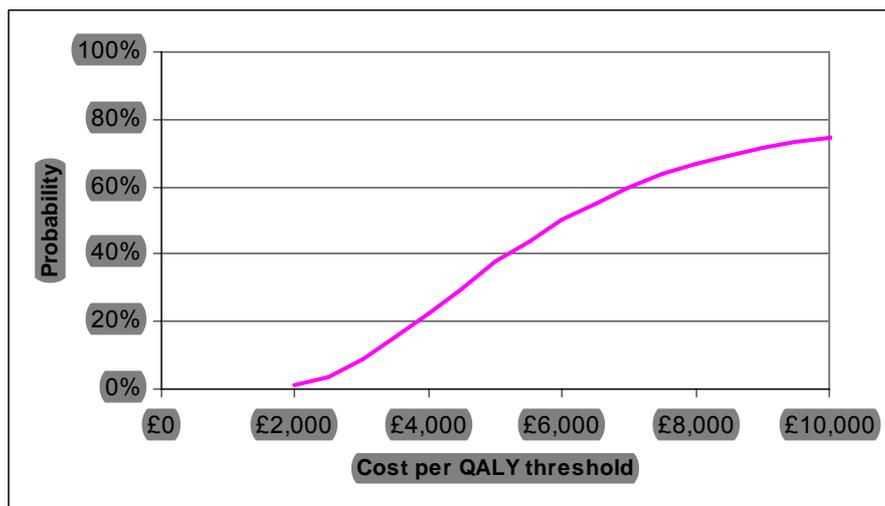
Indirect comparison of tricyclics and SSRIs

The results presented above (see Tables 3 and 4) for tricyclics compared to no treatment and SSRIs compared to no treatment, suggest that SSRIs are more cost-effective than tricyclics as they have a larger QALY gain and lower cost compared to usual care. However, as discussed earlier, this conclusion should be treated with caution as it is based on an indirect comparison and these have a high potential for bias.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. There is a moderate amount of good quality evidence, mainly in patients with refractory IBS and with some depression, showing a significant global improvement in symptoms for both tricyclics and SSRIs when compared with placebo.

2. There is limited evidence to show:
   - A significant reduction in pain and bloating for tricyclics when compared with placebo.
   - A borderline improvement in bowel habit for tricyclics when compared with placebo.

3. There is a moderate amount of good quality evidence showing show no significant improvement in global symptoms for 50mg tricyclics (Trimipramine) compared with 30mg.

4. The evidence is inconclusive as to whether there is an improvement in pain for SSRIs compared with placebo.

5. There is a moderate amount of good quality evidence to show there is no significant reduction in bloating or improvement in bowel habit for SSRIs when compared with placebo.

6. There is a moderate amount of weak evidence to show there is no significant improvement in quality of life for SSRIs when compared with usual care.

7. There is a moderate amount of good quality evidence to show there are significantly more patients discontinuing treatment with SSRIs compared with usual care.

ADVERSE EFFECTS EVIDENCE STATEMENTS (BASED ON NICE CLINICAL GUIDELINE 23 'DEPRESSION')

8. Patients started on antidepressants who are not considered to be at increased risk of suicide should normally be seen after 4 weeks. Thereafter they should be seen on an appropriate and regular basis.

9. In patients in primary care, there is insufficient evidence to determine whether there is a clinically significant difference between other antidepressants and amitriptyline (TCA) on reducing the likelihood of leaving treatment early either for any reason or due to side effects.

10. There is good evidence in trials of eight weeks and longer, that there is no clinically significant difference between SSRIs and placebo on reducing the likelihood of leaving treatment early. This is not consistent when analysing the reasons for leaving treatment, which demonstrate a clinically significant difference favouring placebo over SSRIs in relation to leaving the treatment early due to side effects.

HEALTH ECONOMIC STATEMENT

Evidence from a decision analytic model showed that low dose tricyclics and SSRIs (trimipramine, amitriptyline, doxepin, paroxetine, fluoxetine) are cost-effective for long-term maintenance use in individuals with IBS. The cost-effectiveness estimate is based on a clinical pathway in which dose is increased gradually and response is assessed every four weeks until an effective dose is established or the maximum dose is reached. Trying a second tricyclic or SSRI in an individual who has not responded to a previous tricyclic or SSRI is also likely to be cost-effective when assuming that response to the second treatment is independent of a lack of response to the first. The cost-effectiveness analysis assumes that treatment is reviewed every 6 months to establish whether it is still relevant to the individual's symptom profile.

EVIDENCE TO RECOMMENDATIONS

There is evidence from the clinical and cost effectiveness that tricyclics and SSRIs are effective in the symptom management of IBS. The GDG also took into consideration the reported adverse effects of tricyclics and SSRIs, but noted that the doses used for the management of IBS are

similar to those used in the treatment of chronic pain and are thus much lower than starting doses used for the management of depression. The GDG was concerned that primary care clinicians consider the side effects when prescribing tricyclics and SSRIs. Although the reported side effects of tricyclics are more common (at higher doses) than SSRIs, the GDG reported that, in primary care, tricyclics are used in preference to SSRIs in low doses. Therefore the GDG advised that Tricyclics and SSRIs should be prescribed as second line treatment, starting with Tricyclics, and moving on to SSRIs only when the former had been shown to be ineffective.

RECOMMENDATION

Healthcare professionals should consider tricyclic antidepressants (TCAs)* as second-line treatment for people with IBS if laxatives, loperamide or antispasmodics have not helped. TCAs are primarily used for treatment of depression but are only recommended here for their analgesic effect. Treatment should be started at a low dose (5–10 mg equivalent of amitriptyline), which should be taken once at night and reviewed regularly. The dose may be increased, but does not usually need to exceed 30 mg.

* At the time of publication (February 2008) TCAs did not have UK marketing authorisation for the indications described. Informed consent should be obtained and documented.

RECOMMENDATION

Selective serotonin reuptake inhibitors (SSRIs) should be considered for people with IBS only if TCAs have been shown to be ineffective.

** At the time of publication (February 2008) SSRIs did not have UK marketing authorisation for the indication described. Informed consent should be obtained and documented.

RECOMMENDATION

Healthcare professionals should take into account the possible side effects when prescribing TCAs or SSRIs***. After prescribing either of these drugs for the first time at low doses for the treatment of pain or discomfort in IBS, the person should be followed up after 4 weeks and then at 6–12 monthly intervals thereafter.

*** At the time of publication (February 2008) SSRIs did not have UK marketing authorisation for the indication described. Informed consent should be obtained and documented.

8.5 Adverse effects of pharmacological interventions

BACKGROUND

A wide variety of pharmacological interventions are available for the treatment of irritable bowel syndrome (IBS). As the various classes of agents have different pharmacological mechanisms of action, people with IBS may potentially be troubled by a wide range of adverse effects, depending on which treatment they are taking. While people with IBS are unlikely to experience significant harm from short-term or intermittent therapy, potential problems may arise if the drugs were taken over a longer-term period.

In making informed treatment decisions, health care professionals and people with IBS need to carefully weigh up evidence on the anticipated benefits against that of any relevant concerns about the safety and tolerability of IBS drug therapy. There are a few fundamental questions that can potentially be usefully answered from a review of adverse effects data:

- For people with IBS and health care professionals choosing to use a particular drug therapy for IBS, the review can inform them of potential adverse effects that could be anticipated from that therapy
- Availability of comparative data among different drugs can help in reaching a treatment decision based on which safety profile (or nature and frequency of adverse effects) is more acceptable.

While the overall frequencies of any adverse effects have been evaluated in the parallel efficacy reviews of IBS, there was limited detail given on what the specific adverse effects were, and whether the classes of drugs differ in their safety and tolerability profile.

8.5.1 Adverse effects: antispasmodics, antimotility agents and laxatives

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the evaluation of adverse effects and are reported in the following sections.

Types of studies

We did not apply any specific inclusion criteria based on study design; however, we preferred to exclude:

- Single case reports, as there was substantial scope for reporting and publication bias (towards the esoteric, interesting cases), and such cases may not have been representative of the general patient population
- Crossover studies, as it was impossible to discriminate between events that arise as a medium-long term complication of the first (previous) treatment, or as events resulting from the present therapy. This was particularly so as the protocol was primarily interested in evaluation of long-term effects from chronic administration of drugs for IBS. Moreover,

adverse events are usually measured as dichotomous outcomes, (presence or absence of an adverse effect) and it was technically challenging to incorporate dichotomous measures from crossover studies into meta-analyses.

For the broad overall review, we accepted all studies evaluating the safety or tolerability of any drug therapy in populations of people with IBS.

We also looked at adverse effects for three specific classes of drugs for IBS; their selection criteria are listed below by each class of agent:

1. Antispasmodics
Interventions of interest:
- Dicyclomine bromide
- Hyoscamine (atropine)
- Hyoscine
- Alverine
- Mebeverine (including modified release)
- Peppermint oil.

Duration of intervention: minimum 4 weeks.

Outcomes
All outcomes reported within the categories of 'adverse effects, side effects, adverse events, complications, safety, or tolerability'.

2. Antimotility Agents
Interventions of interest:
1. Loperamide
2. Co-phenotrope/Lomotil/Diphenoxylate.

Duration of intervention: minimum 4 weeks.

Outcomes
All outcomes reported within the categories of 'adverse effects, side effects, adverse events, complications, safety, or tolerability'.

3. Laxatives
Types of participants
People with symptoms of IBS, including those with single symptom of IBS i.e. chronic constipation with no physical cause.

Excluded: studies of long-stay patients in hospital or nursing home setting, and palliative care patients (evidence too indirect as the groups are very different from people with IBS).

Interventions of interest:

1. Polyethylene glycol (PEG) laxatives
2. Lactulose.

Duration of intervention: minimum 4 weeks.

Outcomes

All outcomes reported within the categories of 'adverse effects, side effects, adverse events, complications, safety, or tolerability'.

Quality of Adverse Effects Data

The techniques used in this review were generally based on advice within the Cochrane Handbook of Systematic Reviews regarding the assessment of adverse effects. This states, in particular, that the value of the data relies heavily on two major factors:

- How thorough were the methods used in monitoring adverse effects?
- How complete or detailed was the reporting?

In view of this, we concentrated on recording the following parameters:

- What methods (if any) did the trials stipulate for the specific assessment of AEs?
- Did the investigators pre-specify any possible adverse events that they were particularly looking out for?
- What categories of adverse effects were reported?

Identification of studies

We used a mixed strategy of checking articles that had already been retrieved for the efficacy reviews, and a new search of MEDLINE, EMBASE, CINAHL and the Cochrane Controlled Trials Register (search string: *'drug-class'/ adverse-effects)* with a total of 7206 hits. The search strategies are listed in Appendix B.

A total of 71 full text articles were screened; adverse effects data were extracted from 17 papers.

Study Design

The following types of studies were included in the adverse effects analysis:

- One non-randomised study which carried out an adverse effects survey in people with IBS
- Six randomised trials of antispasmodic agents
- One randomised trial of antimotility agents

- Seven randomised trials of laxatives (including one with a crossover design that should be considered with caution) and one observational long-term follow-up study.

Population

The underlying diagnosis was stated to be IBS for people in the observational study, as well as in the antispasmodic and antimotility agent trials.

However, the included laxative trials were predominantly in people with chronic or simple constipation, with no obvious physical cause. It is assumed that some of these people would have IBS.

Intervention and Comparisons

There was a very diverse range of interventions and associated comparator agents across the trials.

Assessment and Reporting of Adverse Effects

The results for the included trials are shown in Table 4 to Table 9.

Although some trials stated that they specifically enquired about adverse effects, none of them actually stated whether this was an open question e.g. 'Are you having any problems with the treatment?' of if the enquiry was targeted at particular symptoms e.g. 'Have you had any diarrhoea?'

Moreover, some trials stated that people were asked to record problems in a diary, but the trial report gives no details about whether people had to record specific events (including when and how often) or if it was left to their discretion.

RESULTS

Overall Effects of Medication on people with IBS

Only one study looked at the overall impact of medication-related adverse effects on people with IBS (Lembo 2004). This was an online survey carried out in the United States, with people identified from a computer database. This survey collected data on medication use (questions on therapies used for symptoms of IBS) and self-reported assessment of adverse effects. Respondents were given a list of GI and non-GI adverse effects (including a free text entry box) and asked if they had experienced any side effects.

All respondents stated that they had been diagnosed with IBS by their physicians. Most of the participants were women (88%) with a median age of 45 years. The average number of medications they had tried was 3.3. Of the 668 respondents, 504 reported constipation as their primary symptom i.e. most people were of the IBS-C subtype.

The survey covered more than 10 drug classes, including laxatives, antispasmodics, antimotility agents, antidepressants etc., but the article reported mainly on laxatives and antispamodics. Overall 51% of people on antispasmodics complained of at least one adverse effect, in contrast to 59% of people on laxatives. Raw data from this survey was reported based on drug class, and the author did not specify what the individual agents and dosages were. The overall figures were:

Table 1.

Side Effect	Laxative		Antispasmodic	
	(n=171)	*%*	(n=189)	*%*
Any	100	*58*	96	*51*
Drowsiness			43	*23*
Dizziness	5	*3*	17	*9*
Insomnia	1	*1*	2	*1*
Nausea	32	*19*	4	*2*
Abdominal cramps	67	*39*	9	*5*
Abdominal pain	36	*21*	6	*3*
Abdominal discomfort	55	*32*	8	*4*
Bloating	36	*21*	6	*3*
Dry mouth	8	*5*	47	*25*
Headache	4	*2*	8	*4*
Decreased sexual interest	1	*1*	2	*1*
Constipation	7	*4*	9	*5*
Diarrhea	31	*18*	2	*1*
None	71	*42*	93	*49*
SEVERITY				
Mild	62	*36*	113	*60*
Moderate	94	*55*	55	*29*
Severe	14	*8*	21	*11*

The four most frequent adverse effects of antispasmodics are listed below. Figures for laxative users are given for purposes of comparison:

Table 2.

	Antispasmodic (N= 189)	Laxative (N= 171)	Relative risk (95% CI)
Dry mouth	47 (25%)	8 (5%)	5.32 (2.59 – 10.94)
Drowsiness	43 (23%)	0 (0%)	Not estimable (Odds Ratio 8.65)
Dizziness	17 (17%)	6 (3%)	2.56 (1.03- 6.34)
Constipation	9 (5%)	7 (4%)	1.16 (0.44-3.05)

The six most frequent adverse effects of laxatives are listed below. Figures for antispasmodic users are given for purposes of comparison.

Table 3.

	Laxative (N=171)	Antispasmodic (N=189)	Relative risk (95% CI)
Abdominal cramp	67 (39%)	9 (5%)	8.23 (4.24-16.0)
Abdominal discomfort	56 (32%)	8 (4%)	7.74 (3.80-15.8)
Bloating	36 (21%)	6 (3%)	6.63 (2.86-15.3)
Diarrhoea	31 (18%)	2 (1%)	17.13 (4.16-70.5)
Abdominal pain	36 (21%)	6 (3%)	6.63 (2.86-15.3)
Nausea	33 (19%)	4 (2%)	9.12 (3.3 – 25.2)

These findings indicated that laxatives wee significantly associated with GI adverse effects, while antispasmodics were associated with dry mouth, dizziness and drowsiness. The Relative Risk figures must be interpreted with caution as the treatment groups were not randomly allocated, and people with particular IBS symptoms may have been selectively channelled towards a specific class of treatment. However, it seems unlikely that channelling based on IBS subtype and existing gastrointestinal symptoms would have accounted for the marked difference seen with regards to side effects symptoms such as dry mouth, drowsiness and dizziness. Moreover, 504/668 (75%) of the respondents stated that constipation was their primary symptom. As the respondents had tried a mean of 3.9 IBS therapies, their adverse experiences with different classes of drugs could have been reflected within the survey responses.

Strengths of the study were that it surveyed people outside a trial setting, and they may potentially have been able to take the treatments over a longer time period. One important feature of the study was that it specifically enquired about certain adverse effects, covering both GI and non-GI problems.

The main weaknesses were that there was considerable potential for selection bias as the respondents may not have been typical of people with IBS. Recall bias was also a major problem if the treatments were taken a long time ago. We were also not certain of the diagnosis of IBS, as this was based on the respondents reporting that they had been diagnosed by a physician. As the survey was carried out in the US, larger or different dose regimens may have been used. Perhaps the most important limitation was that the individual agents were not listed, and lumping by drug class obscured any differences in the safety profiles of individual agents within a class.

RESULTS OF SPECIFIC CLASSES OF IBS THERAPIES

1. Antispasmodics
There were six included RCTs.

The interventions and comparators were extremely varied, as was the reporting of adverse effects outcomes. In view of this, no meta-analysis was performed and a descriptive summary is given in the appendix.

Mebeverine versus Dicyclomine
There was a trend towards a lower rate of adverse effects in the mebeverine group (RR 0.33, 95% CI 0.10, 1.04; p=0.06). GI symptoms were more frequently seen with dicyclomine.

Mebeverine standard preparation versus Mebeverine sustained release
No clear difference in safety or tolerability could be seen.

Mebeverine versus Trimebutine
Adverse effects were non-significantly more common in people taking mebeverine, with dry mouth being the most common complaint with mebeverine. Note: Trimebutine is not listed in the British National Formulary.

Alverine versus placebo
The authors reported 5 adverse events with alverine related to the nervous system but did not give any details.

2. Antimotility agents
We only identified two RCTs with adverse effects data. One was a crossover trial which should be treated with caution (Cann 1984).

No clear trend could be identified.

3. Laxatives

There were 7 included RCTs. Most of them did not specifically state that the participants had IBS, although it was very likely that some of the people with chronic constipation would fulfil the criteria for IBS.

The interventions and comparators were extremely varied, as was the reporting of adverse effects. A descriptive summary is given in the appendix.

Non-randomised study of TransiPEG.

One long-term (6 month) observational study on 231 people taking TransiPEG provided limited information as there was no comparator group (Paille 1999). In this study, 21 people (9%) reported adverse effects with 14 (6%) stopping therapy due to adverse events. The most frequently reported problems were abdominal pain (8 people), flatulence (5 people), and diarrhoea (4 people).

Lactulose versus Isphagula husk.

Two RCTs were evaluated, one of which was a crossover study (Quah 2006), and the findings must therefore be viewed cautiously. Abdominal pain occurred at a higher rate in the lactulose arms than the isphagula husk arms.

PEG versus lactulose (Ferguson and Attar 1999; Bouhnik 2004)

There were non-significant trends towards lactulose causing more abdominal pain and bloating, as compared to PEG. Flatus was also more common with lactulose (RR 1.72, 95% CI 0.99, 2.72).

PMF versus Placebo.

No clear trend could be identified from the two small studies (Corazziari 1996; Corazziari 2000).

Different formulations of PEG

PEG 3350 and PEG 4000 showed similar rates of adverse effects. Low dose PEG in either form was associated with fewer adverse effects compared to high dose (Chaussade 2003).

The RCT data on lactulose was consistent with the findings of the non-randomised data with regards to increased risk of abdominal symptoms.

Limitations of the results

There were four major limitations that arose in this adverse effects review. The first problem was that the studies were primarily aimed at assessing and reporting on the efficacy of the drug treatments. Evaluation of safety took a back seat, and reporting of adverse effects data was often cursory or non-existent, even in instances where the methods sections had explicitly

stated the intention of monitoring for adverse effects. Trial reports did not follow a structured format (e.g. by WHO system organ class) of reporting adverse effects, thus making it impossible to pool outcomes data.

The second major issue ws that none of the included studies appeared to anticipate, or pre-specify particular adverse effects of interest or concern. It was often possible to predict, based on pharmacological mode of action, the potential adverse effects of a drug therapy. Trial investigators could have designed specific aspects of the protocol to concentrate on detecting these adverse effects. While some of the trials did specify general measures for monitoring overall adverse effects, none of the reports stated whether they were checking for any specific problems. Given the wide-ranging nature of possible adverse effects, it may have been very difficult for trials to reliably pick up safety issues, unless there was some prior awareness of what the potential problems might have been.

The third major limitation was that many of the adverse outcomes of interest were very similar to the symptoms of the IBS itself. For instance, laxatives were associated with flatulence, cramps and abdominal pain – all of which are commonly seen in people with untreated IBS, and also form part of the efficacy assessment. This made it very difficult for the investigators to differentiate between the progress of the condition and the harmful effects of the drug. Moreover, some of these symptoms were listed within the efficacy section of the trial report, and it was not possible to determine whether a deterioration in these symptoms was due to lack of efficacy, the natural history of the condition, or the adverse effect of the drug.

The final limitation was the lack of specific non-randomised studies aimed at eliciting adverse effects of drug therapy in IBS. The only such study we looked at was not focused on particular drugs, and was only able to provide data on broad classes of IBS therapies, without naming specific drugs.

Table 4: Details of RCTs of Antispasmodic agents

Author	Year	Study design	Drug & Dose	Control	Age	Mean duration
Grillage	1990	Randomised, double-blind, parallel-group	Mebeverine 135 mg 3x daily	Dicyclomine 10 mg 3x daily	26	8 weeks
Gilbody	2000	Randomised, double-blind, double-dummy	Mebeverine 200 mg b.i.d.	Mebeverine 135 mg t.i.d.	33	8 weeks
Schaffstein	1990	Randomised, double-blind, parallel-group	135 mg mebeverine tablet	200 mg trimebutine tablet	Not stated	28 days
Mitchell	2002	Double-blind, randomised, placebo-controlled, parallel group	120 mg alverine citrate 3x daily	Placebo capsules 3x daily	40	12 weeks
Liu	1997	Randomised, double-blind, placebo-controlled	Colpermin 3-4x daily	Placebo capsules 3-4x daily	Not stated	1 month
Van Outryve	1995	Randomised, double-blind, crossover	Mebeverine plain 135 mg, 2 capsules t.i.d.	Mebeverine sustained release 200 mg, 2 capsules b.i.d.	49	6 weeks

Table 5: Details of RCTs of Antispasmodic agents

Study	Methods used for AEs	Reports	Drug	Comparator
Grillage 1990	Diary system		**Mebeverine**	**Dicyclomine**
		Total participants	23	28
		with any AE	3	11
		with severe AE	0	0
		stopped due to AE	1	2
		Specific details	Weight increase, headache, flatulence, tremor. Reason for drop-out: weight increase.	Gastrointestinal disturbance (nausea, dysphagia, flatulence, dyspepsia, diarrhoea) - 7 Back pain, headache, agitation. Reason for drop-out: nausea, agitation.
Gilbody 2000	Diary system, laboratory tests		**Mebeverine 200mg bd**	**Mebeverine 135mg tds**
		Total participants	106	107
		with any AE	63	66
		with severe AE	4	2
		stopped due to AE	0	0
		Specific details	Abdominal pain Diarrhoea (Total drop-outs appear unrelated to study medication: 4 for elective surgery, 1 pregnancy, 1 bloody diarrhoea - ulcerative colitis)	Abdominal pain Diarrhoea

Study	Methods used for AEs	Reports	Drug	Comparator
Schaffstein 1990	Specific enquiry at follow up, spontaneous reporting, urine biochemistry		**Mebeverine 135 mg**	**Trimebutine 200 mg**
		Total participants	100	99
		with any AE	8	2
		with severe AE	0	0
		stopped due to AE	1	1
		Specific details	Reason for drop-out: diarrhoea. Upper abdominal heaviness - 1 Dry mouth + nausea - 1 Erythrocyte count < normal on day 28 - 3	Reason for drop-out: headache, drowsiness, fever, pruritus. Dry mouth - 4 Hot flush - 3 Opthalmic pain, bitter taste, nausea - 1 Erythrocyte count < normal on day 28 - 9
Mitchell 2002	Spontaneous reports (unexpected events from people with IBS or investigators)		**Alverine**	**Placebo**
		Total participants	53	54
		with any AE	21 (40%)	26 (48%)
		with severe AE	0	0
		stopped due to AE	0	0
		Specific details	5 AEs related to nervous system	
Liu 1997	Liver function test		**Colpermin**	**Placebo**
		Total participants	52	49
		with any AE	0	0
		with severe AE	0	0
		stopped due to AE	0	0
		Specific details	Heatburn from chewing capsule - 1 Skin rash (mild, transient) - 1	
Van Outryve 1995	Clinical exam, spontaneous reporting, biological tests		**Mebeverine 270mg tds**	**Mebeverine 200 mg bd**
		Total. of participants	60	60

		No. with any AE	12	8
		No. with severe AE	0	0
		No. stopped due to AE	0	0
		Specific details	Limb cramps - 2 Diarrhoea - 1 Anorexia + nausea - 1 Flu/pharyngitis/rhinitis/sinusitis - 1 Headache - 2 Menorrhagia - 1 Palpitations - 1 Alopexia - 1 Cystitis - 2	Diarrhoea - 1 Abdominal pain, cramps - 1 Eructation - 1 Flu/pharyngitis/rhinitis/sinusitis - 1 Headache - 1 Headache + nasal obstruction - 1 Left paresthesia + asthenia - 1 Anaemia + leucopenia - 1

Table 6: Details of RCTs of Antimotility agents

Author	Year	Study design	Drug & Dose	Control	Age	Mean duration
Cann	1984	Randomised, double-blind, crossover	Loperamide 2mg 1-6X daily	Placebo	35	5 weeks
Lavo	1987	Randomised, double-blind, placebo controlled	Loperamide 2mg 1-4capsules at night	Placebo	43	13 weeks

Table 7: Details of RCTs of Antimotility agents

Study	Methods used for AEs	Reports	Drug	Comparator
Cann 1984	Not stated		**Loperamide**	**Placebo**
		Total participants	28	28
		with any AE	8	10
		with severe AE	NA	NA
		stopped due to AE	NA	NA
		Specific details	NA	NA
Lavo 1987	Specific enquiry at follow-up		**Loperamide**	**Placebo**
		Total participants	11	10
		with any AE	NA	NA
		with severe AE	NA	NA
		stopped due to AE	NA	NA
		Specific details	Constipation - 1 Swollen fingers - 1 Sleep disturbance - 1 Blisters in the mouth - 1	Blisters in the mouth - 1 Vertigo - 1 Numbness in the legs - 1 Headache - 1

Table 8: Details of RCTs on Laxatives

Author	Year	Study design	Diagnosis	Drug & Dose	Control	Mean duration	Age
Chaussade	2003	Double-blind, randomised, parallel-group	Chronic idiopathic constipation	PEG 3350 + electrolytes (Transipeg) standard dose (1 5.9g sachet / day) or Transipeg maximum dose (2 5.9g sachets / day)	PEG 4000 (Forlax) standard dose (1 10g sachet / day or Forlax maximum dose (2 10g sachets / day)	1 month	52
Bouhnik	2004	Randomised controlled parallel-group study	Chronic idiopathic constipation	PEG 4000 (Forlax) 10 g sachets. 2 sachets/day, option to take 1-3 sachets daily	Lactulose 10 g sachets. 2 sachets/day, option to take 1-3 sachets daily	4 weeks	57
Dettmar	1998	open study with random allocation	Simple constipation	Lactulose and other laxatives	Ispaghula husk 3.5 g sachet twice daily	4 weeks	Not stated
Quah	2006	Randomised crossover		Lactulose 10 ml twice daily, maximum 30 ml twice daily	Isphagula husk 3.5 g sachet once daily, maximum 2 sachets	4 weeks	50
Corazziari	2000	Double blind, placebo controlled, parallel group, randomised		PMF-100 17.5 g sachet twice daily, can be reduced to once daily	Placebo sachet twice daily, can be reduced to once daily	20 weeks	43
Corazziari	1996	Randomised, double blind, placebo controlled, parallel group	Chronic idiopathic constipation	PMF-100 17.5 g sachet twice daily, can be reduced to once daily	Placebo sachet	8 weeks	42
Ferguson, Attar	1999	Single blind, randomised, controlled	Chronic idiopathic constipation	PEGes 2 sachets (13 g PEG 3350 each). Adjusted to 1-3 sachets daily if required	Lactulose 2 sachets (10 g each). Adjusted to 1-3 sachets daily if required	4 weeks	Not stated

Table 9: Details of RCTs of Antimotility agents

Study	Methods used for monitoring adverse effects	Reports	Drug		Comparator	
Chaussade 2003	Diary system, spontaneous reporting, clinical examination, specific enquiry at follow-up, global impression (visual analogue scale		**PEG 3350 low dose**	**PEG 3350 high dose**	**PEG 4000 low dose**	**PEG 4000 high dose**
		Total participants	65	70	66	67
		with any AE	33	42	33	36
		with severe AE	NA	NA	NA	NA
		stopped due to AE	NA	NA	NA	NA
		Specific details	Diarrhoea / liquid stools – 9 (14%)	Diarrhoea / liquid stools – 25 (36%)	Diarrhoea / liquid stools – 11 (17%)	Diarrhoea / liquid stools – 24 (36%)
Bouhnik 2004	Diary system, faecal studies		**PEG 4000**		**Lactulose**	
		Total participants	32		33	
		with any AE	5		8	
		with severe AE	0		0	
		stopped due to AE	3		4	
		Specific details	Borborygmi - 9 Bloating - 10 Abdominal pain - 6 Flatus in excess - 15		Borborygmi - 12 Bloating - 11 Abdominal pain - 9 Flatus in excess - 16	

Study	Methods used for monitoring adverse effects	Reports	Drug	Comparator
Dettmar 1998	Diary system		**Isphagula husk**	**Lactulose**
		Total participants	224	91
		with any AE	NA	NA
		with severe AE	NA	NA
		stopped due to AE	NA	NA
		Specific details	Abdominal pain / griping – 34 Distension – 35 Diarrhoea – 5 Flatulence – 63 Indigestion – 25 Nausea – 18	Abdominal pain / griping – 20 Distension – 12 Diarrhoea – 6 Flatulence – 20 Indigestion – 8 Nausea – 7
Quah 2006	Diary system		**Isphagula husk**	**Lactulose**
		Total participants	39	39
		with any AE	NA	NA
		with severe AE	NA	NA
		stopped due to AE	NA	NA
		Specific details	Abdominal bloating – 11 Abdominal pain / cramps – 5 Nausea – 1 Wind / flatulence – 10 Headache – 2 Anorexia – 0 Urgency – 8	Abdominal bloating – 11 Abdominal pain / cramps – 10 Nausea – 3 Wind / flatulence – 16 Headache – 2 Anorexia – 1 Urgency – 14

Study	Methods used for monitoring adverse effects	Reports	Drug	Comparator
Corazziari 2000	Diary system, blood tests, specific enquiry, examination		**PMF-100**	**Placebo**
		Total participants	33	37
		with any AE	57 (events, not people)	41 (events, not people)
		with severe AE	0	0
		stopped due to AE	0	0
		Specific details	Number of adverse events (11 people reported 2 symptoms) Nausea - 22 Vomiting - 1 Anal pain - 5 Haematochezia - 7 Call to evacuate absent - 4 Anal itching - 2 Headache - 0 Epigastric pain / discomfort - 13 Faecal incontinence - 3 (Stopped due to nausea, incontinence, anal pain)	Nausea - 17 Vomiting - 1 Anal pain - 0 Haematochezia - 2 Call to evacuate absent - 0 Anal itching - 2 Headache - 3 Epigastric pain / discomfort - 16 Faecal incontinence - 0
Corazziari 1996	Uncertain		**PMF-100**	**Placebo**
		Total. participants	25	23
		with any AE	NA	NA
		with severe AE	NA	NA
		stopped due to AE	NA	NA
		Specific details	Abdominal pain - 8 Abdominal bloating - 16 Flatulence - 9 Borborygmi - 3 Anorexia - 3 Headache - 3 Asthenia - 1	Abdominal pain - 6 Abdominal bloating - 12 Flatulence - 5 Borborygmi - 8 Anorexia - 1 Headache - 1 Asthenia - 1
Study	**Methods used for monitoring adverse effects**	**Reports**	**Drug**	**Comparator**
Ferguson , Attar 1999	Diary system, blood tests (cell counts, electrolytes, glucose, urea nitrogen,		**PEG 3350**	**Lactulose**

protein, creatinine, folate, B12)	Total participants	60	55
	with any AE	NA	NA
	With severe AE	NA	NA
	Stopped due to AE	2	1
	Specific details	(stop due to AE - reason: 1 acute diarrhoea with vomiting and fever 1 abdominal pain) Liquid stools - 7 Bloating - 11 Abdominal pain - 7 Flatus - 8 Rumbling - 2 Mild hypokalaemia - 1 (taking diuretics)	(stop due to AE - reason: depression) Liquid stools - 5 Bloating - 20 Abdominal pain - 11 Flatus - 15 Rumbling - 8 Mild hypokalaemia - 1 (taking diuretics)

8.5.2 Adverse effects: tricyclics and selective serotonin re-uptake inhibitors

CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

Types of studies

Adverse effects data have been extracted from the randomised trials included in the clinical effectiveness review of antidepressants.

Search strategy for identification of studies

In discussion with colleagues at NICE, it was agreed that for adverse effects evidence, direct reference to Clinical Guideline 23 (Depression) was appropriate in order to supplement the RCT data. Contact with the National Collaborating Centre for Mental Health clinical effectiveness lead established the search strategy supporting their work, and evidence statements were lifted from the guideline in relation to the adverse effects for both TCAs and SSRIs.

PARAMETERS OF THE REVIEW

It is recognised that these drugs are typically used to treat large populations with psychiatric morbidities. The context for this review is to recognise the potential harmful effects of these drug types when prescribed at low dose for the treatment of IBS symptoms, namely pain and discomfort. In the last twenty years, tricyclics and SSRIs have been prescribed in the treatment of functional GI disorders such as IBS. The prevalence of anxiety and depressive disorders is high in people with severe and/or intractable IBS and may be present to some degree in all people with IBS. A recognised pharmokinetic effect from these drug types is an analgesic effect that is separate from inhibitor effects typically desired in the treatment of depression. There is growing evidence to suggest that visceral pain syndromes such as IBS may be effectively treated using these drugs, that appear to modulate the interactions between the central and enteric nervous systems.

The two drug classes (tricyclics, e.g. trimipramine, amitriptyline, doxepin, and SSRIs, e.g. paroxetine, fluoxetine) are among the medications that have been used as long-term maintenance therapy (i.e. for 3 months or more) for IBS.

Treatment (for example with amitriptyline, the lowest cost tricyclic) is generally initiated at 10mg with dose increases of 10mg no more frequently than every 2 weeks. Patients are encouraged to increase the dose until an effective dose is established or side-effects become problematic, up to a maximum dose of 30mg for tricyclics and 20mg for SSRIs. An exception to this is doxepin, where the smallest tablet size is 50mg. Patients who do not respond to maximum dose are switched to an alternative drug, with switches from first to second tricyclic considered and switches between tricyclics and SSRIs considered.

A. Tricyclics

1. Adverse effects

All TCAs cause, to varying degrees, anticholinergic side effects (dry mouth, blurred vision, constipation, urinary retention, sweating), sedation and postural hypotension. These side effects necessitate starting with a low dose and increasing slowly. Mianserin is a sedating tricyclic.

Evidence is available for the adverse effects occurring for doxepin (Boerner 1988) and mianserin (Tanum and Malt 1996). Tanum and Malt (1996) included the adverse effect of mild sedation, and it is known that mianserin is a sedating tricyclic. All the trials compared tricyclics with placebo control.

In Boerner (1988), a dose of 50mg doxepin was compared with placebo. In Tanum and Malt (1996), the dose of mianserin was up to 120mg; we note that in this study only 60% of the participants had IBS. These doses are higher than those generally recommended for the use of these drugs in IBS.

Figure 1.

Review: IBS MW 080507
Comparison: 39 Tricyclic versus placebo
Outcome: 01 Adverse events

Study or sub-category	Tricyclic n/N	Control n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
Boerner 1988	7/38	4/37		83.82	1.86 [0.50, 6.99]
Tanum and Malt 1996	20/25	3/22		16.18	25.33 [5.31, 120.92]
Total (95% CI)	63	59		100.00	5.66 [2.21, 14.48]

Total events: 27 (Tricyclic), 7 (Control)
Test for heterogeneity: Chi² = 6.24, df = 1 (P = 0.01), I² = 84.0%
Test for overall effect: Z = 3.62 (P = 0.0003)

0.001 0.01 0.1 1 10 100 1000
Favours treatment   Favours control

The Tanum and Malt (1996) study reported that significantly more patients were found to have adverse effects of mild sedation. In both studies the confidence intervals were wide.

One further study (Myren 1984) compared different doses of trimipramine (30 to 50mg). There were no significant differences between any of the trimipramine groups and placebo for palpitations, dizziness or dryness in the mouth. There was a significant increase in tiredness and morning drowsiness in the first 1 to 2 weeks for the antidepressant groups compared with placebo, but there were no significant differences at the end of treatment.

2. Number of patients withdrawing

Two studies recorded the numbers of withdrawals from the trials, for people receiving 120mg minaserin (Tanum and Malt 1996) and 10mg trimipramine (Tripathi 1983), compared with placebo. Meta-analyis showed a statistically significant effect in favour of placebo, but the confidence interval was wide.

Figure 2

Review: IBS MW 080507
Comparison: 39 Tricyclic versus placebo
Outcome: 02 Withdrawals

Study or sub-category	Tricyclic n/N	Control n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
Tanum and Malt 1996	2/25	0/25		12.65	5.43 [0.25, 118.96]
Tripathi 1983	19/25	13/25		87.35	2.92 [0.87, 9.78]
Total (95% CI)	50	50		100.00	3.24 [1.06, 9.91]

Total events: 21 (Tricyclic), 13 (Control)
Test for heterogeneity: Chi² = 0.14, df = 1 (P = 0.71), I² = 0%
Test for overall effect: Z = 2.06 (P = 0.04)

0.001 0.01 0.1 1 10 100 1000
Favours treatment   Favours control

B. SSRIs

1. Adverse effects

Evidence is available for the adverse effects occurring for fluoxetine, in comparison with placebo (Kuiken 2003). It was used at a dose of 20mg in this study. The confidence interval was fairly wide, but there was no significant difference between interventions.

Figure 3

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 22 Number of adverse effects

Study or sub-category	Antidepressant n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 SSRI vs placebo (general adverse effects)					
Kuiken 2003	8/19	10/21		100.00	0.88 [0.44, 1.77]
Subtotal (95% CI)	19	21		100.00	0.88 [0.44, 1.77]

Total events: 8 (Antidepressant), 10 (Control)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.35 (P = 0.73)

0.1 0.2 0.5 1 2 5 10
Favours antidepress   Favours control
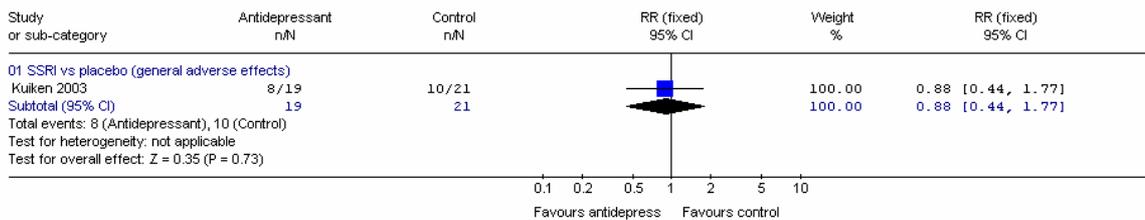
2. Number of people withdrawing because of side effects.

One study also investigated the number of participants withdrawing because of side effects for paroxetine compared with placebo (Tabas 2004); and Creed 2003 recorded the number who did not complete the treatment, for paroxetine compared with usual care. The confidence intervals were either wide or very wide, so conclusions could not be drawn about the non-significant results, but we noted there were significantly more people discontinuing treatment in the paroxetine group of the Creed (2003) study than in the usual care group. In Creed (2003), the dose of paroxetine was 20mg; and in Tabas (2004), the dose was up to 40mg.

Figure 4

Review: Antidepressants (Version 02)
Comparison: 04 Anti depressant versus placebo / usual care
Outcome: 19 Number of withdrawals

Study or sub-category	Antidepressant n/N	Control n/N	Peto OR 95% CI	Weight %	Peto OR 95% CI
01 SSRI vs placebo (withdrawals due to adverse effects)					
Tabas 2004	4/38	4/43		100.00	1.15 [0.27, 4.90]
Subtotal (95% CI)	38	43		100.00	1.15 [0.27, 4.90]
Total events: 4 (Antidepressant), 4 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.18 (P = 0.85)					
02 SSRI vs usual care (discontinued treatment)					
Creed et al 2003	29/86	0/86		100.00	10.93 [4.93, 24.23]
Subtotal (95% CI)	86	86		100.00	10.93 [4.93, 24.23]
Total events: 29 (Antidepressant), 0 (Control)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 5.89 (P < 0.00001)					

0.01 0.1 1 10 100

Favours antidepress   Favours control

C. Use of these drugs in the treatment of depression

Further evidence is available on the use of these medications (often at higher doses) for depression (NICE CG 23). For example, the effective dose of tricyclics for depression is usually taken to be 125mg. When used for depression, SSRIs are generally as effective as tricyclic antidepressants and are less likely to be discontinued because of side effects.

Tricyclics

Tricyclics are usually considered to have more side effects than other classes of drugs when used for depression, and people are less likely to stay in treatment on tricyclics than on other classes of drugs (RR= 0.71; 95% CI 0.65 to 0.78).

SSRIs

There is evidence suggesting that people are more likely to have side effects while on SSRIs than on placebo (RR=1.19; 95% CI 1.13 to 1.25), and are also more likely to withdraw from treatment due to side effects (RR=2.45; 95% CI 2.08 to 2.89). However, overall, a similar number of patients withdrew from treatment, so these withdrawals due to side effects of SSRIs may be offset by the number of people on placebo who withdrew for other reasons, for example due to lack of effectiveness of the treatment. Fewer people on SSRIs than on other antidepressants withdrew from treatment due to side effects (RR= 0.78; 95% CI, 0.71 to 0.85).

EVIDENCE STATEMENTS

Evidence statements for this review are mostly based on those in NICE Clinical Guideline 23 'Depression' (statements 2 to 4).

1. There is a moderate amount of good quality evidence to show there are significantly more patients discontinuing treatment with SSRIs compared with usual care.

2. Patients started on antidepressants who are not considered to be at increased risk of suicide should normally be seen after 2 weeks. Thereafter they should be seen on an appropriate and regular basis.

3. In people in primary care, there is insufficient evidence to determine whether there is a clinically significant difference between other antidepressants and amitriptyline (TCA) on reducing the likelihood of leaving treatment early either for any reason or due to side effects.

4. There is good evidence in trials of eight weeks and longer that there is no clinically significant difference between SSRIs and placebo on reducing the likelihood of leaving treatment early. This is not consistent when analysing the reasons for leaving treatment, which demonstrate a clinically significant difference favouring placebo over SSRIs in relation to leaving the treatment early due to side effects.

9 Psychological Interventions

Clinical Questions

1. Does CBT have a role in managing symptoms?

2. Do psychological interventions have a role in managing symptoms?

3. Does hypnotherapy have a role in managing IBS symptoms?

4. Does bio-feedback have a role in managing symptoms?

5. Does relaxation therapy have a role in managing symptoms?

BACKGROUND

Psychosocial factors are integral to the way in which people experience and interpret symptoms and they influence both illness behaviour and response to treatment. The effects on gastrointestinal function caused by emotional and psychological response include fluctuation in acid secretion; changes in motor activity and gut transit and have been well documented (Wolf 1981). Although it has been shown that there are no greater psychological disturbances in people with IBS than in the general population (Wilhelmsen 2000), anxiety and depression can be major contributing factors in the symptom profiles of IBS. Psychotherapy has been suggested as a possible treatment to reduce pain and symptoms and also to improve quality of life.

There are a range of psychological treatments which can be used in the management of IBS. Psychological therapies may be defined as the treatment of mental and emotional disorders through the use of psychological techniques designed to encourage communication of conflicts and insight into problems, with the goal being relief of symptoms, changes in behaviour leading to improved social and vocational functioning and personality growth.

Relaxation therapy is the simplest form of psychotherapy. The premise is that if response to stress contributes to IBS, reducing autonomic stress responses by relaxation will reduce symptoms, induce a feeling of well-being and increased confidence which will allow people with IBS to feel more able to control the condition. Relaxation can be taught using audio tapes and there are many readily available which people with IBS can access (Jones 2000).

More complex psychological interventions include biofeedback, cognitive behavioural therapy, dynamic psychotherapy and hypnotherapy are usually initiated for people with moderate or severe symptoms who have not responded to other management programmes. These therapies are effective, but time consuming to provide, require specialist input and currently availability varies widely across the UK.

Biofeedback

Biofeedback includes a number of techniques in which a physiological process is monitored and information regarding unconscious bodily functions are shown by audiovisual display to the patient. The patient is taught to bring about changes in the physiological process by using a number of strategies e.g. thoughts, sensations, feelings. The rationale is that the physiological process being monitored is causally related to a clinical condition, in this case IBS, and that alteration of the process can lead to a reduction or resolution of symptoms.

Cognitive Therapy

Cognitive therapy is a therapy that assumes that faulty thought patterns (called cognitive patterns) cause maladaptive behaviour and emotional responses. The treatment focuses on changing thoughts in order to solve psychological and personality problems. Behaviour therapy is also a goal-oriented, therapeutic approach, and it treats emotional and behavioural disorders as maladaptive learned responses that can be replaced by healthier ones with appropriate training. Cognitive-behavioural therapy (CBT) integrates features of behaviour modification into the traditional cognitive restructuring approach. Cognitive-behavioural therapy attempts to change clients' unhealthy behaviour through cognitive restructuring (examining assumptions behind the thought patterns) and through the use of behaviour therapy techniques. CBT can be used as a long-term treatment for irritable bowel syndrome. Different programmes comprise different elements in a variety of combinations, including: helping patients recognise the causes of disease; cognitive restructuring techniques to address unhelpful beliefs; changing underlying depressive or threatening 'life scripts'; psychotherapy to cope with emotional problems and find new solutions; stress management or relaxation training, using progressive muscle relaxation techniques; breaking habits of learned illness behaviours; practising more adaptive behaviours; assertion and coping skills training. CBT can be administered to patients individually or as a group.

Hypnotherapy

Hypnosis describes a range of naturally occurring states of altered awareness which may vary from momentary distractions and 'absences' through much enhanced states of relaxation to very deep states of inward focus and awareness. The mental processes which can occur in any of these states, appropriately utilised are generally far more flexible and potentially far more powerful in effecting change than those we can achieve in most everyday states of active conscious awareness. These states may be induced quite formally or quite naturalistically, in an almost unnoticeable way, depending on the requirement of the problem, the capability of the practitioner and the needs of the client (UK Council for Psychotherapy (UKCP) 1992).

Gut-directed hypnotherapy is a specific form of hypnotherapy developed for the management of gastrointestinal disorders. It uses the therapeutic qualities of hypnotherapy, such as deep relaxation, and adds gut-specific treatments and suggestions. 'Gut-directed hypnotherapy' can

be used as a treatment for irritable bowel syndrome. IBS is ideal for treatment with hypnosis, as there is no structural damage to the body. During hypnotherapy people learn how to influence and gain control of their gut function and then seem to be able to change the way the brain modulates their gut activity (Whorwell 2005). Firstly, patients are given a brief outline of the anatomy and physiology of the gut and a schematic representation of their symptoms, using a diagram of the colon showing how smooth muscle spasm can give pain, bloating and a disordered bowel habit. Patients are told that the reduction of this spasm and normalisation of smooth muscle activity will reduce pain and bloating and encourage a more normal flow through the bowel. Hypnosis is induced by a standard technique, then over successive sessions, patients are asked to place a hand on their abdomen and feel warmth; then this warmth is related to reduction of spasm and the ability to alleviate pain and distension; patients are told that bowel habit will normalise as their control gradually improves; they visualise the gut as a meandering river and they can adjust the flow along it to a comfortable setting as one would open and close lock gates on a river. Patients may be given a self-hypnosis tape to use at home. Ego-strengthening and confidence-building comments can be made at the end of the sessions. Hypnotherapy can be administered to patients individually or as a group.

Dynamic Psychotherapy

In the NHS psychodynamic psychotherapy is practised by psychiatrists, psychologists, social workers and other professionals who have received additional specialised training in these techniques. Long-term dynamic psychotherapy aims to bring about extensive change in several aspects of a person's functioning. It is a prolonged treatment typically comprising of hourly meetings every week for periods of time up to three years. Short-term or focal dynamic psychotherapy is a modification of the approach in which attention is focused on only one area of the person's experience. This shortens the amount of time required and usually this form of treatment requires between 10 and 20 sessions (University of Newcastle 2005).

The selection of the appropriate psychological approach will depend on the individual person. They may express a preference for a particular intervention but in order to be able to make informed choices people with irritable bowel syndrome need to be made aware of the existence of these psychological treatments and the rationale for their use. It is important that they be made aware that using a psychological treatment does not mean that the syndrome is "all in the mind." Addressing psychosocial factors is increasingly recognised as an important part of the management of irritable bowel syndrome.

9.1 Relaxation

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, except that crossover studies were excluded as inappropriate due to the carry-over effect of the relaxation interventions.

The following comparisons were included:

- Relaxation versus waiting list control, or symptom monitoring only
- Relaxation versus usual medical care
- Relaxation versus another intervention.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library (*1966 to current day with guidance from the GDG). Additionally, the PSYCINFO database was searched for this review. The search strategies are listed in Appendix B.

Study Design

Three parallel group design randomised trials were included (Blanchard 1993; Forbes 2000; Keefer 2001). Further details are given in the included studies table. Forbes (2000) was conducted in the UK, the other two studies were carried out in the USA. Trials lasted between 6 and 12 weeks. One study was conducted among patients recruited from their personal physician or media publicity (Blanchard 1993); one recruited from gastroenterologists and local media (Keefer 2001), and one recruited from secondary care (Forbes 2000). The total number of patients in the studies was 16 in Blanchard (1993) and Keefer (2001). Forbes (2000) included 25 and 27 patients in the two treatment arms respectively.

Population

All the studies included only people with IBS. Blanchard (1993) and Forbes (2000) did not report the number of participants with bloating; Keefer (2001) reported that seven (of 16) had bloating. None of the studies reported whether the symptoms were post-infective. The mean age of participants was 51.5 years in Keefer (2001), with participants aged between 34 and 76 years; the people with IBS were aged 22 to 64 years in Blanchard (1993), and the median age in Forbes (2000) was 37 years, with a range from 19 to 71 years. All the studies included more women than men. The patients in the Blanchard (1993) study had had IBS for more a mean of around 13 years, and in Keefer (2001) 15.8 years. The patients in Forbes (2000) had had IBS for more than six months.

In the Blanchard (1993) study, 56% of participants had an Axis I diagnosis; in Keefer (2001), 77% had an Axis I diagnosis, and participants were excluded if they had bipolar I or II,

schizophrenia or other psychoses, or if they were actively suicidal. Co-morbidities were not stated in Forbes (2000).

Interventions

Blanchard (1993) used progressive muscle relaxation, on an individual basis with two sessions per week for the first two weeks and then once a week for six additional weeks; regular home practice was emphasised with an audiotape to guide this. Keefer (2001) used relaxation response meditation, in six weekly 30-minute treatment sessions.

The following comparisons were included:
- Relaxation versus symptom monitoring only: two studies (Blanchard 1993; Keefer 2001).
- Relaxation versus another intervention
    - Relaxation versus hypnotherapy (Forbes 2000).

Outcomes

The outcomes reported were:

1. Global symptoms:
   a) Global improvement in symptoms (number of patients) (Blanchard 1993; Forbes 2000; Keefer 2001)
   b) Global symptom score:
      - Global improvement of IBS symptoms (mean Composite Primary Symptom Reduction [CPSR] score; CPSR represents proportional reduction in score from baseline; scale range -1 to +1; Blanchard 1993; Keefer 2001).

2. Individual symptoms:
   a) Pain
      - Pain score (0 to 4 recorded over 28 days where 0=absent to 4=debilitating; i.e. maximum 112) reported by Blanchard (1993).

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in Appendix D.

An adequate method of randomisation was reported in one study (computer-generated random numbers; Forbes 2000); the other studies did not state the method. Allocation concealment was also not reported. The patients were not blinded (because of the type of intervention). No study described an *a-priori* power calculation. The three studies included in the review demonstrated baseline comparability of the groups, although the baseline scores for Blanchard (1993) were higher for the relaxation group on abdominal pain (mean score 31.2 (SD 25.1) compared with 24.4 (21.4) for the symptom monitoring group). This was not a statistically significant difference (Figure 1).

Figure 1: Baseline pain scores

Study or sub-category	N	Relaxation Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Blanchard 1993	8	31.20(25.10)	8	24.40(21.40)		100.00	6.80 [-16.06, 29.66]
Total (95% CI)	8		8			100.00	6.80 [-16.06, 29.66]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.58 (P = 0.56)

-100 -50 0 50 100
Favours treatment   Favours control
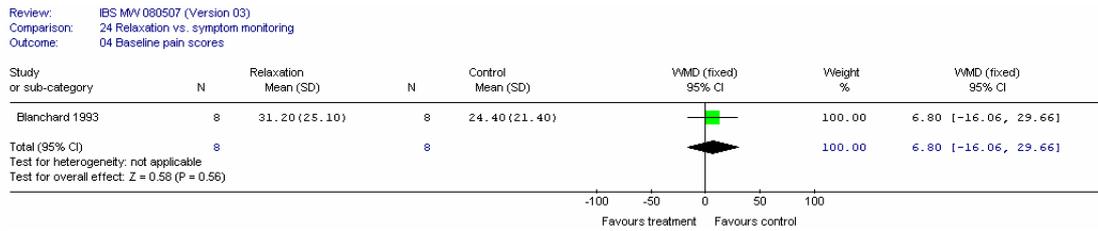
All the participants were followed up in Forbes (2000). There were 20% or fewer drop-outs overall in one study (Keefer 2001): 3/16 dropped out (19%) (2/8 (25%) from the intervention group and 1/8 from the control group). 7/23 dropped out (30.4%) in Blanchard (1993), 6/14 (43%) from the intervention group and 1/9 (11%) from the control group; this study was regarded with caution, as this large and unequal drop-out could have introduced a bias.

RESULTS

A. Relaxation versus symptom monitoring only

There were two studies that compared relaxation with symptom monitoring in people with IBS (Blanchard 1993; Keefer 2001).

1. Global symptoms

a) Number of patients with global improvement in symptoms

This outcome was reported by Blanchard (1993) and Keefer (2001).

Figure 2: Global improvement of symptoms

Study or sub-category	Relaxation n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Keefer 2001	4/6	2/7		64.86	2.33 [0.64, 8.57]
Blanchard 1993	4/8	1/8		35.14	4.00 [0.56, 28.40]
Total (95% CI)	14	15		100.00	2.92 [0.97, 8.80]

Total events: 8 (Relaxation), 3 (Control)
Test for heterogeneity: Chi² = 0.21, df = 1 (P = 0.64), I² = 0%
Test for overall effect: Z = 1.90 (P = 0.06)

0.1 0.2 0.5 1 2 5 10
Favours control   Favours relaxation
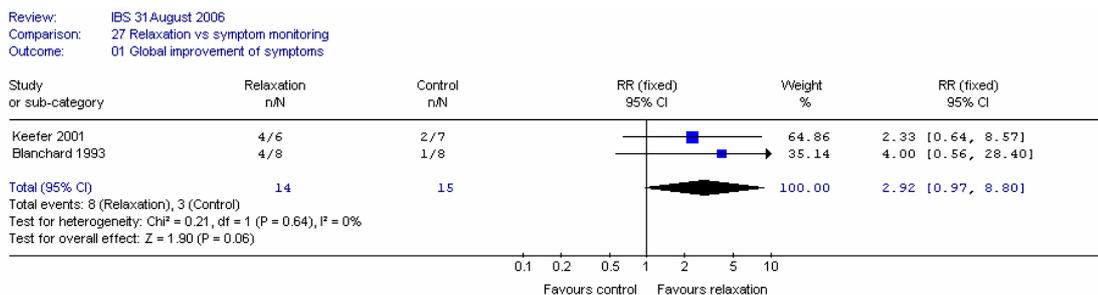
Meta-analysis of two studies in 29 patients showed a large effect, favouring relaxation, but the confidence interval was wide, such that the results are not significant. We noted that there was also attrition bias for the Blanchard (1993) study.

b) Global symptom score

The global improvement of IBS symptoms (mean Composite Primary Symptom Reduction [CPSR] score; CPSR represents proportional reduction in score from baseline; scale -1 to +1) was reported by Blanchard (1993) and Keefer (2001). There was a large statistically significant improvement in symptoms for the relaxation group, but the confidence interval was

also wide. We noted that there was also attrition bias for the Blanchard (1993) study, and the other study, Keefer (2001), was small (13 patients).

Figure 3: Global symptom score (CPSR)

Review: IBS MW 080507 (Version 03)
Comparison: 24 Relaxation vs. symptom monitoring
Outcome: 02 CPSR

Study or sub-category	N	Relaxation Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Blanchard 1993	8	0.52(0.38)	8	-0.14(0.57)		65.51	0.66 [0.19, 1.13]
Keefer 2001	6	0.51(0.60)	7	-0.15(0.60)		34.49	0.66 [0.01, 1.31]
Total (95% CI)	14		15			100.00	0.66 [0.28, 1.04]

Test for heterogeneity: Chi² = 3.20E-31, df = 1 (P = 1.00), I² = 0%
Test for overall effect: Z = 3.37 (P = 0.0008)

-4 -2 0 2 4
Favours control   Favours relaxation
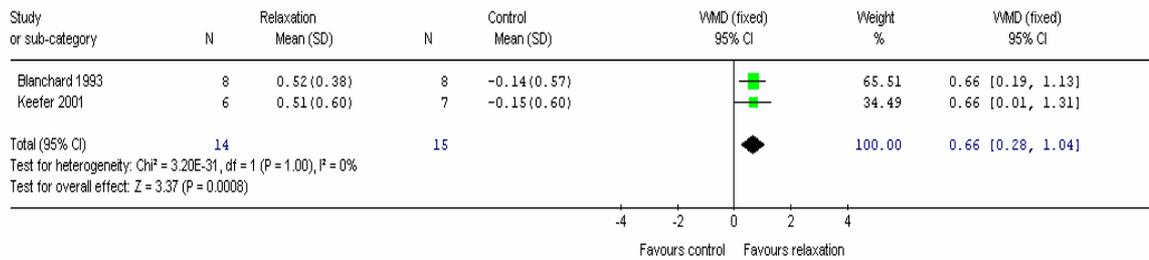
2. Individual symptoms

a) Pain

A pain score (0-4 recorded, over 28 days, where 0=absent and 4=debilitating, i.e. maximum 112) was reported by Blanchard (1993). There was no significant difference between interventions.

Figure 4: Pain score

Review: IBS MW 080507 (Version 03)
Comparison: 24 Relaxation vs. symptom monitoring
Outcome: 03 Pain score
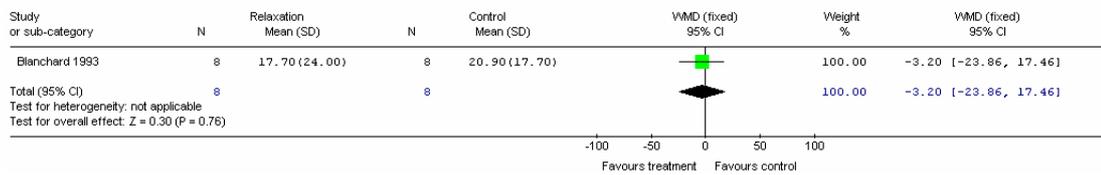
Study or sub-category	N	Relaxation Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Blanchard 1993	8	17.70(24.00)	8	20.90(17.70)		100.00	-3.20 [-23.86, 17.46]
Total (95% CI)	8		8			100.00	-3.20 [-23.86, 17.46]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.30 (P = 0.76)

-100 -50 0 50 100
Favours treatment   Favours control

B. Hypnotherapy versus relaxation

1. Global symptoms

Global improvement in symptoms (number of patients) was reported by Forbes (2000) at 12 weeks. There was no significant difference between interventions.

Figure 5:

Review: IBS 31 August 2006
Comparison: 25 Hypnotherapy versus relaxation
Outcome: 01 Global improvement (number of patients)

Study or sub-category	Hypnotherapy n/N	Relaxation n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Forbes 2000	19/25	16/27		100.00	1.28 [0.87, 1.88]
Total (95% CI)	25	27		100.00	1.28 [0.87, 1.88]

Total events: 19 (Hypnotherapy), 16 (Relaxation)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.27 (P = 0.20)

0.1 0.2 0.5 1 2 5 10
Favours psychotherpy   Favours hypnotherpy
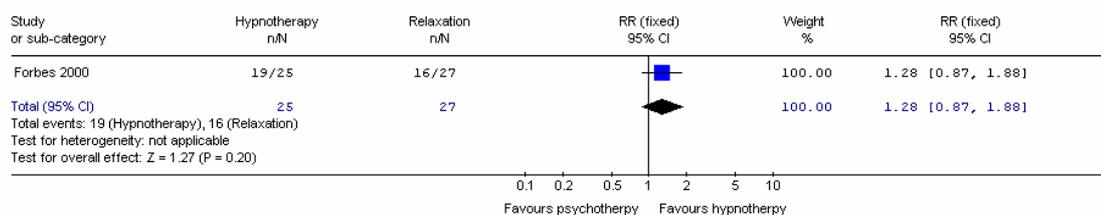
EVIDENCE STATEMENTS

1. There is insufficient evidence to show if there is a difference between relaxation and symptom monitoring, in the number of people with global improvement of symptoms, in people with long term IBS, at least half of whom had psychiatric co-morbidities.

2. There is a limited amount of weak evidence to show a large, significant improvement in global symptom score for people receiving relaxation, compared with symptom monitoring, in people with long term IBS, at least half of whom had psychiatric co-morbidities.

3. There is limited evidence to show no significant difference in pain score between relaxation and symptom monitoring, in people with long term IBS, at least half of whom had psychiatric co-morbidities.

4. There is a fair evidence to show no significant difference between relaxation and hypnotherapy in the number of people with global improvement of symptoms.

Evidence to recommendation

The GDG decided there is insufficient evidence to make a recommendation. This is discussed in the evidence to recommendation statement for relaxation and biofeedback (Section 9.3).

9.2 Biofeedback

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to this review and are reported below.

Types of intervention

Both multiple and single component therapies were eligible for inclusion.

Search strategy for identification of studies

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. Biofeedback, aloe vera and reflexology were combined into one search. The search strategies are listed in Appendix B.

The search strategy identified 560 studies. The titles and abstracts of these studies were assessed. Fifty-four studies were identified as being potentially relevant to the reviews and the papers for these were retrieved in full. Four studies met the inclusion criteria for this review, two of which were reports of the same trial (Blanchard 1992; Meissner 1997). The reference lists of

these studies were inspected for further potential papers, but none were identified. The 17 excluded studies, with reasons for exclusion, are listed in the Appendix E.

DESCRIPTION OF STUDIES

Study Design

There were four randomised trials in this review, reported in three papers (Blanchard 1992; Leahy 1997; Neff 1987); two trials were from the same paper (Blanchard 1992), and one was reported only as an abstract (Leahy 1997). All the studies but one took place in the US. Leahy (1997) was carried out in the UK.

Population

All patients had a diagnosis of IBS and were treated in the secondary care setting in which the study took place. There was a higher proportion of women. The age range was 21 to 76 years.

Patients in the Leahy study were said to be resistant to conventional medical therapy. The other studies did not report whether the IBS was refractory.

Interventions

One study (Leahy 1997) evaluated a single intervention, using a computer-aided gut-directed biofeedback apparatus to teach relaxation for IBS patients when troubled by symptoms. Patients were randomised to biofeedback or counselling. Biofeedback patients received four half-hour sessions.

Three trials (Blanchard 1992 x 2; Neff 1987) evaluated multi-component therapy, which used a combination of educational information, progressive relaxation therapy, thermal biofeedback treatment and training in stress coping strategies. This was offered on an individual basis. The combination treatment consisted of twelve one-hour sessions spread over eight weeks.

In Blanchard (1992a), the same therapist delivered the treatments, but in Blanchard (1992b) eight different therapists took part.

In the two Blanchard 1992 trials, the patients were matched into triads, based on gender, age and predominant GI symptoms, and randomly assigned to multi-component biofeedback, attention placebo or symptom monitoring. Neff (1987) randomly assigned patients to multi-component biofeedback and symptom monitoring.

Comparisons

All the comparative studies used symptom monitoring or attention placebo controls. The latter was used in Blanchard (1992). A combination of two procedures was used: 'pseudo meditation'

(in which patients were asked not to relax) and biofeedback using suppression of alpha-waves in the EEG (this is not associated with the relaxed state).

In the symptom monitoring control group, patients simply monitored their symptoms for the duration of the intervention. The symptom monitoring (control) group were offered treatment at a later stage.

The following comparisons were reported:

- Single component biofeedback versus counselling then both groups had biofeedback (Leahy 1997)
- Multi-component biofeedback versus symptom monitoring (Blanchard 1992 x2; Neff 1987)
- Multi-component biofeedback versus attention control (Blanchard 1992 x2).

METHODOLOGICAL QUALITY OF INCLUDED STUDIES

None of the RCTs reported details of the method of randomisation or allocation concealment. Patients were matched on age, gender and primary GI symptoms before randomisation in the Blanchard (1992) study. Patients in the Neff (1987) study were similar at baseline for age, duration of IBS, and years of education, but there were differences in the number of IBS-D patients: 5/10 in the biofeedback group and 2/9 in the control group.

In Blanchard (1992a) all patients completed the trial, 1/9 dropped out from the control group of the Neff study, and in Blanchard (1992b), 7/31 (22%), 8/30 (27%) and 10/31 (32%) dropped out of the study for multi-component, attention placebo and symptom monitoring respectively. Therefore the second Blanchard trial is at higher risk of bias and will be considered in sensitivity analyses as appropriate.

RESULTS

A. Single component biofeedback

In an abstract, Leahy (1997) reported that counselling had no effect on symptom score, but did not give separate results for the group randomised to biofeedback.

B. Multi-component biofeedback

Three randomised trials (Blanchard 1992a and b; and Neff 1987) in 30 and 115 patients, and 19 patients respectively gave a multi-component therapy as the biofeedback intervention.

1. Global symptoms

All studies reported a composite primary symptom reduction score (CPSR): firstly, each patient recorded in a daily diary a symptom score, comprising abdominal pain, tenderness, diarrhoea, constipation, flatulence, belching and nausea. This was used to calculate a reduction score using the formula:

(baseline symptom score – end of treatment score) / baseline symptom score *100.

a) Global improvement in symptoms (number of patients)

The RCTs reported the number of patients with an improvement in global symptoms; the Blanchard (1992) trials reported rater-assessments, but the patient assessment results were selected for the Neff (1987) study. Meta-analysis of three trials in 101 patients showed a statistically significant improvement in symptoms for biofeedback compared with symptom monitoring; RR 1.85 (95%CI 1.22, 2.79), with insignificant heterogeneity (I^2=32%, p=0.23). This corresponded to a number needed to treat of 4 (95%CI 3, 8), for a control group rate of 0 to 45%. However, there was no significant difference between biofeedback and attention placebo. The comparison of attention placebo versus symptom monitoring was also significant. We noted that the Blanchard (1992b) study had about 30% dropouts. A sensitivity analysis without this study for the comparison of biofeedback with symptom monitoring resulted in more heterogeneity and changed the relative risk to 3.14 (95%CI 1.35, 7.31).

Figure 1

Review:	IBS 31 August 2006 (hyp_other pysch)					
Comparison:	39 Biofeedback vs. symptom monitoring or attention control					
Outcome:	01 Global improvement (number of patients)					
Study or sub-category	Hypnotherapy n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI	
01 Multi-component biofeedback vs symptom monitoring						
Blanchard 1992a	8/10	4/10		21.59	2.00 [0.88, 4.54]	
Blanchard 1992b	20/31	14/31		75.58	1.43 [0.90, 2.28]	
Neff 1987	6/10	0/9		2.83	11.82 [0.76, 184.13]	
Subtotal (95% CI)	51	50		100.00	1.85 [1.22, 2.79]	
Total events: 34 (Hypnotherapy), 18 (Control)						
Test for heterogeneity: Chi² = 2.95, df = 2 (P = 0.23), I² = 32.1%						
Test for overall effect: Z = 2.92 (P = 0.003)						
02 Multi-component biofeedback vs attention control						
Blanchard 1992a	8/10	7/9		25.66	1.03 [0.64, 1.64]	
Blanchard 1992b	20/31	21/30		74.34	0.92 [0.65, 1.31]	
Subtotal (95% CI)	41	39		100.00	0.95 [0.71, 1.26]	
Total events: 28 (Hypnotherapy), 28 (Control)						
Test for heterogeneity: Chi² = 0.14, df = 1 (P = 0.71), I² = 0%						
Test for overall effect: Z = 0.36 (P = 0.72)						
03 Attention placebo vs symptom monitoring						
Blanchard 1992a	7/9	4/10		21.58	1.94 [0.84, 4.48]	
Blanchard 1992b	21/30	14/31		78.42	1.55 [0.99, 2.44]	
Subtotal (95% CI)	39	41		100.00	1.64 [1.10, 2.43]	
Total events: 28 (Hypnotherapy), 18 (Control)						
Test for heterogeneity: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%						
Test for overall effect: Z = 2.42 (P = 0.02)						

0.1 0.2 0.5 1 2 5 10
Favours control   Favours treatment

b) Global improvement of symptoms

The two trials within Blanchard (1992) reported the scores on the CPSR. There was a statistically significant difference, favouring multi-component feedback compared with symptom monitoring, but not in comparison with attention control, although the confidence intervals were fairly wide. We noted that around 30% of the patients in Blanchard (1992b) had missing data and we have assumed the numbers of patients in the analysis are the values for completers only.

Figure 2

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 39 Biofeedback vs. symptom monitoring or attention control
Outcome: 02 CPSR score

Study or sub-category	N	Treatment Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Multi-component biofeedback vs symptom monitoring							
Blanchard 1992a	10	45.20(59.30)	10	9.50(53.30)		33.56	35.70 [-13.72, 85.12]
Blanchard 1992b	24	32.40(63.20)	21	6.40(57.00)		66.44	26.00 [-9.12, 61.12]
Subtotal (95% CI)	34		31			100.00	29.26 [0.63, 57.88]
Test for heterogeneity: Chi² = 0.10, df = 1 (P = 0.75), I² = 0%							
Test for overall effect: Z = 2.00 (P = 0.05)							
02 Multi-component biofeedback vs attention control							
Blanchard 1992a	10	45.20(59.30)	10	38.00(31.60)		43.77	7.20 [-34.45, 48.85]
Blanchard 1992b	24	32.40(63.20)	22	30.20(63.80)		56.23	2.20 [-34.54, 38.94]
Subtotal (95% CI)	34		32			100.00	4.39 [-23.16, 31.94]
Test for heterogeneity: Chi² = 0.03, df = 1 (P = 0.86), I² = 0%							
Test for overall effect: Z = 0.31 (P = 0.75)							
03 Attention control vs symptom monitoring							
Blanchard 1992a	10	38.00(31.60)	10	9.50(53.30)		45.79	28.50 [-9.90, 66.90]
Blanchard 1992b	24	30.20(63.80)	21	6.40(57.00)		54.21	23.80 [-11.50, 59.10]
Subtotal (95% CI)	34		31			100.00	25.95 [-0.04, 51.94]
Test for heterogeneity: Chi² = 0.03, df = 1 (P = 0.86), I² = 0%							
Test for overall effect: Z = 1.96 (P = 0.05)							

-100 -50 0 50 100
Favours control   Favours treatment

2. Individual symptoms

Neff (1987) reported means for these outcomes, but no standard deviations or p-values were given, so the rest of this review uses the results from Blanchard (1992)**.**

a) Pain

The study reported daily abdominal pain and discomfort symptom scores on a scale of 0 to 4, recorded as weekly scores (i.e. maximum of 28). The confidence intervals were too wide to draw conclusions.

Figure 3

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 39 Biofeedback vs. symptom monitoring or attention control
Outcome: 04 Pain score (low = good)

Study or sub-category	N	Intervention Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Biofeedback vs attention control							
Blanchard 1992a	10	9.70(7.00)	10	11.70(10.80)		63.23	-2.00 [-9.98, 5.98]
Blanchard 1992b	24	17.50(19.00)	22	13.50(17.20)		36.77	4.00 [-6.46, 14.46]
Subtotal (95% CI)	34		32			100.00	0.21 [-6.14, 6.55]
Test for heterogeneity: Chi² = 0.80, df = 1 (P = 0.37), I² = 0%							
Test for overall effect: Z = 0.06 (P = 0.95)							
02 Biofeedback vs symptom monitoring							
Blanchard 1992a	10	9.70(7.00)	10	9.10(10.60)		68.36	0.60 [-7.27, 8.47]
Blanchard 1992b	24	17.50(19.00)	21	22.00(20.40)		31.64	-4.50 [-16.07, 7.07]
Subtotal (95% CI)	34		31			100.00	-1.01 [-7.52, 5.50]
Test for heterogeneity: Chi² = 0.51, df = 1 (P = 0.48), I² = 0%							
Test for overall effect: Z = 0.31 (P = 0.76)							
03 Attention control vs symptom monitoring							
Blanchard 1992a	10	11.70(10.80)	10	9.10(10.60)		59.23	2.60 [-6.78, 11.98]
Blanchard 1992b	22	13.50(17.20)	21	22.00(20.40)		40.77	-8.50 [-19.80, 2.80]
Subtotal (95% CI)	32		31			100.00	-1.93 [-9.14, 5.29]
Test for heterogeneity: Chi² = 2.19, df = 1 (P = 0.14), I² = 54.4%							
Test for overall effect: Z = 0.52 (P = 0.60)							

-10 -5 0 5 10
Favours Intervention   Favours control

b) Bloating

The second study reported bloating scores. Generally the confidence intervals were wide, but the attention placebo gave significantly less bloating than the symptom monitoring.

Figure 4
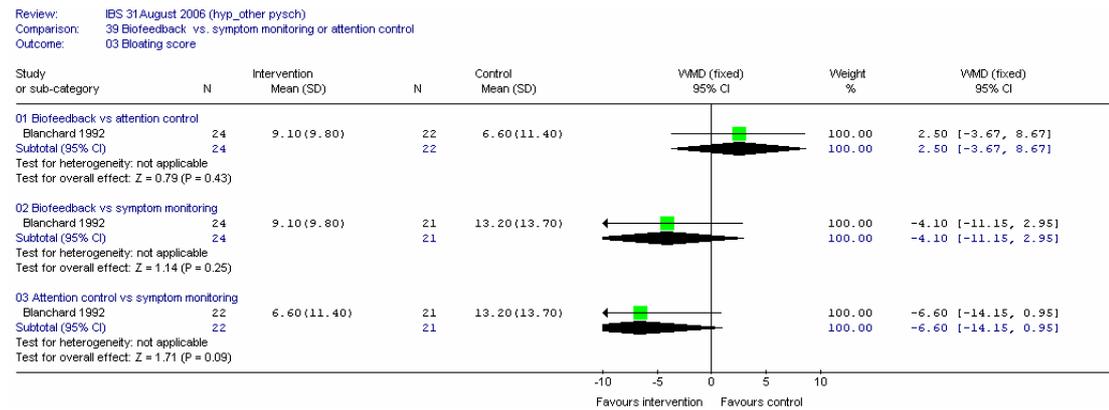
c) Bowel Habit

i. Constipation

The second study reported constipation scores. The scale used was 0 to 4 and the weekly average was used (i.e. maximum of 28). Generally the confidence intervals were too wide to draw conclusions.

Figure 5
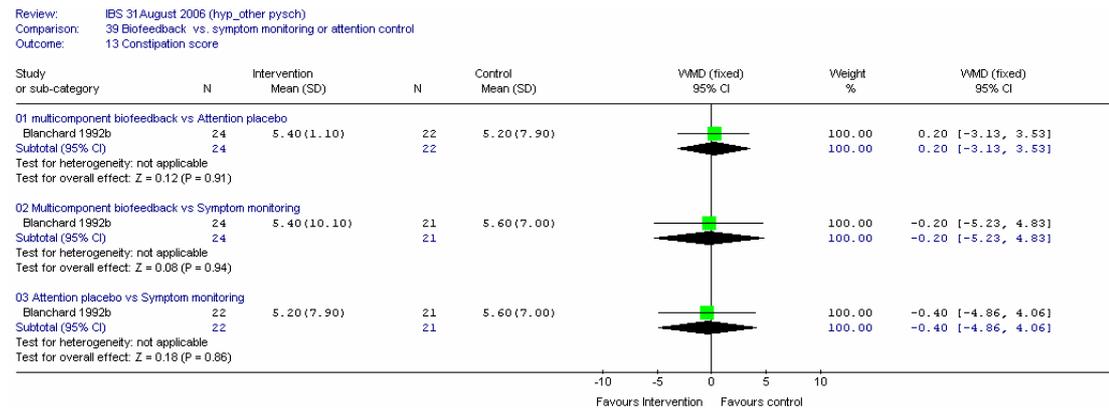
ii. Diarrhoea

The second study reported diarrhoea scores. The scale used was 0 to 4 and the weekly average was used (i.e. maximum of 28). Generally the confidence intervals were wide, but the multi-component biofeedback gave significantly lower diarrhoea score than the symptom monitoring.

Figure 6
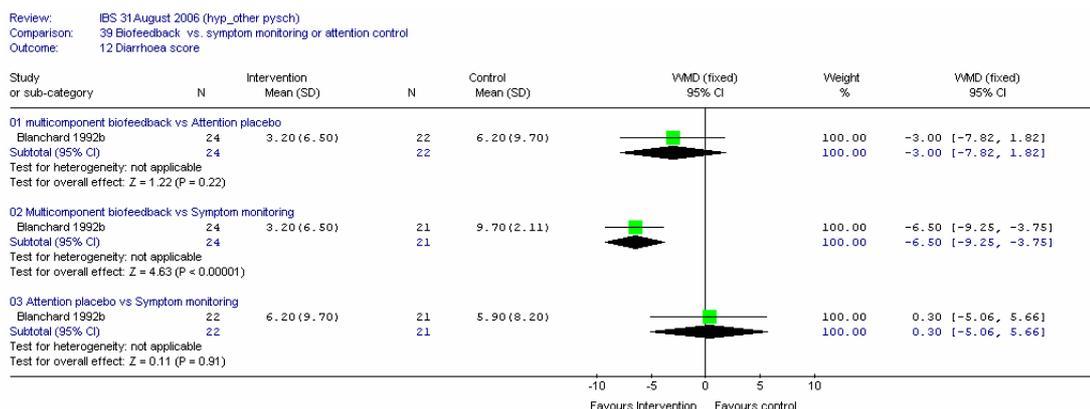
GDG DISCUSSION

The GDG noted that multi-component biofeedback (consisting of education, progressive relaxation therapy, thermal biofeedback treatment and training in stress coping strategies) is effective for people with IBS in comparison with symptom monitoring. However, the attention placebo of pseudo meditation and alpha wave EEG biofeedback had similar efficacy for improving global symptoms. The GDG suggested that actively involving people in the management of their chronic condition may have a beneficial effect. They also noted that spending time with patients and taking them seriously is valuable, and that behavioural treatments are difficult to unravel.

EVIDENCE STATEMENTS

1. There is limited weak evidence to show a statistically significant improvement in global symptoms for biofeedback compared with symptom monitoring, although the confidence interval was fairly wide. There was no difference between biofeedback and attention placebo.

2. There was insufficient evidence to determine the effects of biofeedback on pain, bloating and constipation.

3. There is limited weak evidence to show a statistically significant improvement for reduction in diarrhoea for biofeedback compared with symptom monitoring, although the confidence interval was fairly wide. There was no significant difference between biofeedback and attention placebo and between symptom monitoring and attention placebo but there was much uncertainty due to wide confidence intervals.

EVIDENCE TO RECOMMENDATION

The GDG decided there is insufficient evidence to make a recommendation. This is discussed in the evidence to recommendation statement for relaxation and biofeedback (Section 9.3).

9.3 Evidence to recommendation: relaxation and biofeedback

The GDG took into consideration the limited clinical effectiveness evidence for relaxation and biofeedback. They noted that the relaxation review showed significant improvement in global symptoms for relaxation compared with symptom monitoring, but the trials were small and one of them had a large attrition bias. The biofeedback review compared biofeedback with symptom monitoring and with attention control: there was a significant effect on global symptoms in comparison with the former, but not with the latter and this led the GDG to conclude that attention may be an important factor for biofeedback.

The GDG considered the clinical evidence to be too limited, either to carry out cost effectiveness analyses or to make recommendations for practice. However, they considered these therapies to be worth following up, particularly in view of recent developments in computer-aided biofeedback techniques and positive patient experience within the GDG indicating a preference towards relaxation. Therefore, the GDG proposed a recommendation for research, involving a comparison of computer-aided biofeedback, relaxation and attention control. The research recommendation is given in chapter 12.

9.4 Psychotherapy

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, except that crossover studies were excluded as inappropriate due to the carry-over effect of the interventions.

The following comparisons were to be included:
- Psychotherapy versus waiting list control/symptom monitoring
- Psychotherapy versus usual medical care
- Psychotherapy plus another intervention versus the other intervention alone
- Psychotherapy versus another intervention.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the PSYCINFO database was searched for this review.

The search strategies are the same as those for hypnotherapy, as these were searched together. Details of the search strategies are lisetd in Appendix B. The titles and abstracts of the studies identified by the searches were assessed. Fifteen were identified to be potentially relevant to the reviews and these papers were retrieved in full. The reference lists of the retrieved studies were inspected for further potential papers, but none were identified. Twelve studies were excluded and are listed in Appendix E, along with reasons for exclusion.

CHARACTERISTICS OF INCLUDED STUDIES

Study Design

Three randomised trials were included (Creed 2003; Guthrie 1991; Svedlund 1983), all in secondary care. Two of these were in the UK, the other in Sweden.

Population

The studies included patients between the ages of 17 and 75 years, although each had slightly different inclusion criteria and mean age of participants (Svedlund 1983: mean age was around 34 years, range 17 to 59 years; Guthrie 1991: mean around 48 years, range 20 to 75 years; Creed 2003: mean around 40 years, range 18 to 65 years). All the studies included more women than men.

IBS was stated, or implied, to be refractory in all of the studies. The patients in Guthrie (1991) had had IBS for more than 1 year (median around 4 years, range 1 to 20 years); their symptoms had not improved with medical treatment (bulking agents and/or antispasmodics) over six months. The patients in Creed (2003) had had IBS for more than six months; they had failed to respond to usual medical treatment. The patients in Svedlund (1983) had had IBS for at least a year, range 1 to 40 years, but their response to previous treatment was not stated.

Only one study gave details about IBS characteristics: Creed (2003) reported that the patients had severe abdominal pain (over 60 on a 100 mm visual analogue scale [VAS]); 29% of the patients had diarrhoea-predominant IBS; 23% were constipation-predominant; and 48% had 'general' IBS.

In two of the studies, around half the participants currently had a psychiatric diagnosis. Of the patients in Guthrie (1991), 30% had major depression and a further 18% had anxiety states; in Creed (2003), 47% had a psychiatric diagnosis (mainly anxiety or depression). In Svedlund (1983), 29% had previously had depression and a further 41% had had other mental disorders. It was unclear if this was their current diagnosis.

Svedlund (1983) stated that the patients were less representative of the general IBS population because they had to be prepared to participate in an extended socio-psychological investigation. Creed (2003) reported that the patients all had severe symptoms, but within this group were broadly representative, and Guthrie (1991) recruited patients from a tertiary referral centre, where there is likely to be a higher proportion with abuse and associated psychological problems.

None of the studies reported the number of participants with bloating or whether the symptoms were post-infective. Creed (2003) described IBS as severe. Guthrie (1991) reported that

gastroenterologists assessment of severity was median 5 (range 2 to 8) on scale 0 to 9. Svedlund (1983) did not report severity.

Interventions

All the studies examined the effect of psychotherapy on IBS symptoms. Two studies (Svedlund 1983; Guthrie 1991) compared psychotherapy plus medical therapy versus medical therapy alone; the third (Creed 2003) compared psychotherapy versus medical therapy and also had a second comparison with the SSRI antidepressant, paroxetine.

In Svedlund (1983), all patients received the same medical treatment but those in one group also received dynamically oriented individual psychotherapy in ten hour-long sessions spread over three months. In Guthrie (1991), patients received the same medical treatment but patients in one group also received dynamic psychotherapy in seven interviews over three months, plus a relaxation tape to use at home. At the end of the 3 month period, patients in the control group were given psychotherapy. Creed (2003) randomised patients into three groups: psychodynamic interpersonal therapy (8 sessions over 3 months); 20mg daily of the SSRI antidepressant paroxetine for 3 months; or 'usual care', in which 'patients continued to be seen either by their gastroenterologist and/or general practitioner, using whatever management was deemed appropriate' (Creed 2003). Other medical management was stopped during the trial period in the paroxetine and psychotherapy groups. In the follow-up period, patients were allowed other treatments, principally antidepressants.

Comparisons

The following comparisons were included:

- Psychotherapy plus medical care versus medical care only (Svedlund 1983; Guthrie 1991)
- Psychotherapy versus usual care (Creed 2003)
- Psychotherapy versus another intervention (antidepressant: SSRI – paroxetine) (Creed 2003).

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in the Appendix D.

The method of randomisation was adequate in Creed (2003) (computer generated random numbers) but not stated in either Svedlund (1983) or Guthrie (1991). Allocation concealment was adequate in Creed (2003) (randomisation list held by study administrator) but not stated in either Svedlund (1983) or Guthrie (1991). The patients were not blinded in any study (because of the type of intervention). The GDG did not consider blinding to be important for this review. Creed (2003) reported an *a priori* power calculation but the other two studies did not.

Svedlund (1983) and Creed (2003) demonstrated baseline comparability between the groups. The groups were mainly comparable in Guthrie (1991), except there was a higher proportion of males in the control group than the intervention group (17/49 vs. 8/53, p<0.05). Svedlund (1983) reported significant differences in the assessors rating of pain (however, this was not significantly different for the self-rating, which we used).

There were fewer than 20% drop-outs in all the studies. Svedlund (1983) reported 2 out of 101 participants dropped out. Guthrie (1991) reported that 7 of 53 participants dropped out of the treatment group plus 6 of 49 controls; data were available at 3 months for all but 2 participants (2% drop out), despite the withdrawal from treatment. In Creed (2003) there were missing data: 16% (14/86) in the paroxetine group; 14% (12/85) psychotherapy; 0% in the routine care group did not start the trial. A further 29/86 (34%) in the paroxetine group and 14/85 (16%) in the psychotherapy arm discontinued treatment, but these patients still appear to have been followed. Overall, loss to follow-up at three months was 12/86 (14%) for paroxetine, 11/85 (13%) psychotherapy and 7/86 (8%) usual care arm. At 15 months the authors contacted more of the patients. The authors reported that there were no significant differences at baseline between those who did and did not complete the treatments. For the 3 month pain score and SF36 outcome measures respectively, the patients included in the analysis were 74 and 59 (69%) paroxetine; 74 and 58 (68%) psychotherapy and 79 and 63 (73%) usual care, but some of these patients had discontinued treatment. We decided to include the results from this study, with some reservations, especially about the paroxetine arm and about the SF36 results. The study also recorded the number of patients with an improvement in global symptoms, based on the results from 74, 74 and 80 patients respectively. The GDG decided that this outcome was more representative because patients that dropped out due to side effects would not have rated their global symptoms as improved. The follow-up period allowed the patients to have paroxetine in all arms: 42% in paroxetine group; 19% in psychotherapy, and; 22% in the usual care group, i.e. the follow-up for the comparison of psychotherapy and paroxetine should be considered to be partly confounded. In addition, 10% of the usual care group were given psychological treatment. Therefore we did not report the results for the follow-up period for the comparison psychotherapy versus paroxetine, and the comparison psychotherapy versus usual care was also considered with caution.

Overall, we considered that none of the studies were at high risk of bias. Creed (2003) was treated with caution for the outcomes pain and SF36 because of missing data and non compliance. We also noted there was some confounding in the follow-up period.

RESULTS

A. Psychotherapy plus medical therapy versus medical therapy only

1. Global symptoms

a) Global improvement in symptoms (number of patients)

Two studies reported the numbers of patients with an improvement in global symptoms: Guthrie (1991) gave the numbers as assessed by gastroenterologists at the end of treatment (12 weeks), and; Svedlund (1983) reported the patients' self-assessment at 15 months follow-up (treatment lasted three months). There are significantly more patients with global improvement at both times. The GDG preferred to use the patient assessment in all reviews. At 15 months follow-up, the number needed to treat was 4 (95% 3, 13), for a control group rate of 40%. [The rater assessed result at 12 weeks corresponded to an NNT of 3 (95%CI 2, 4) for a control group rate of 23%].

Figure 1:

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 38 Psychotherapy plus medical care versus medical care
Outcome: 03 Global improvement (number of patients) 12 weeks

Study or sub-category	Psychotherapy n/N	Usual care n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 raters' assessment at 12 weeks					
Guthrie 1991	33/46	10/43		100.00	3.08 [1.74, 5.47]
Subtotal (95% CI)	46	43		100.00	3.08 [1.74, 5.47]
Total events: 33 (Psychotherapy), 10 (Usual care)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.86 (P = 0.0001)					
03 patients' assessment at 15 months follow up					
Svedlund 1983	33/49	20/50		100.00	1.68 [1.14, 2.49]
Subtotal (95% CI)	49	50		100.00	1.68 [1.14, 2.49]
Total events: 33 (Psychotherapy), 20 (Usual care)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.61 (P = 0.009)					

0.1 0.2 0.5 1 2 5 10
Favours control   Favours treatment

b) Global symptom score

Svedlund (1983) reported a patient-rated global symptom score at 12 weeks and 15 months follow-up; this score rated somatic symptoms (19 items; each rated on a 7-point scale 0=absent to high=extremely; i.e. lower is better. It was unclear what the maximum of the scale was but it is assumed to be 6. This gave an overall maximum of 114). There was a small statistically significant difference at 12 weeks, favouring psychotherapy plus medical treatment, which increased to -8.10 (95%CI -12.31, -3.89) at 15 months follow-up. The control group score was 38.0.

Figure 2:

Review: IBS 31 August 2006 (Version workotherpysch)
Comparison: 38 Psychotherapy plus medical care versus medical care
Outcome: 11 Global symptoms score

Study or sub-category	N	Psychotherapy+med Mean (SD)	N	Medical treatment Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 12 weeks							
Svedlund 1983	49	32.91(10.63)	50	37.47(10.74)		100.00	-4.56 [-8.77, -0.35]
Subtotal (95% CI)	49		50			100.00	-4.56 [-8.77, -0.35]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.12 (P = 0.03)							
02 15 months							
Svedlund 1983	49	29.87(10.63)	50	37.97(10.74)		100.00	-8.10 [-12.31, -3.89]
Subtotal (95% CI)	49		50			100.00	-8.10 [-12.31, -3.89]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.77 (P = 0.0002)							

-100 -50 0 50 100
Favours psych+med   Favours med

2. Individual symptoms

a) Abdominal pain

Svedlund (1983) reported a patient-rated pain score at 12 weeks and 15 months follow-up. It was unclear what the maximum of the scale was but the baseline was about 10 units. There was a small statistically significant difference at 12 weeks, favouring psychotherapy plus medical treatment, which increased to -2.30 (95%CI -3.43, -1.17) at 15 months follow-up. The control group score was 8.11.

Figure 3:

Review: IBS 31 August 2006 (Version workotherpysch)
Comparison: 38 Psychotherapy plus medical care versus medical care
Outcome: 05 Pain score - patient rated

Study or sub-category	N	Psychotherapy+ med Mean (SD)	N	Medical care Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 12 weeks							
Svedlund 1983	49	6.76(2.37)	50	7.77(2.39)		100.00	-1.01 [-1.95, -0.07]
Subtotal (95% CI)	49		50			100.00	-1.01 [-1.95, -0.07]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.11 (P = 0.03)							
02 15 months							
Svedlund 1983	49	5.81(1.89)	50	8.11(3.59)		100.00	-2.30 [-3.43, -1.17]
Subtotal (95% CI)	49		50			100.00	-2.30 [-3.43, -1.17]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 4.00 (P < 0.0001)							

-10 -5 0 5 10
Favours psych+med   Favours medical

3. Mental health outcomes (overall mental health; depression; anxiety)

Svedlund (1983) reported the number of patients with improved mental symptoms, as assessed by raters at 12 weeks and by raters and patients at 15-month follow-up. There were significantly more patients improving according to the raters, but the patients' rating showed no significant changes. The authors reported that several patients denied having mental complaints at baseline.

Figure 4:
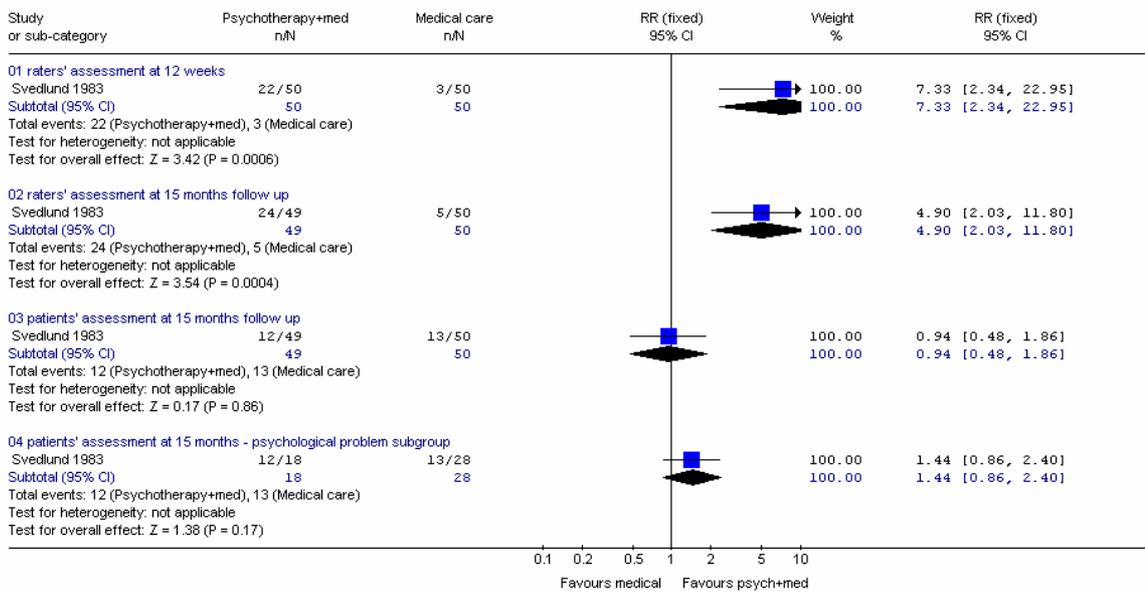
Guthrie (1991) reported a significant improvement, favouring the psychotherapy group, in the median scores on the Hamilton depression scale (p<0.001) and the Clinical anxiety scale (p<0.01), as assessed by the psychiatrist.

B. Psychotherapy versus usual medical therapy

1. Global symptoms

Global improvement in symptoms (number of patients) was reported by Creed (2003) at 12 and 52 weeks. There were significantly more patients with global improvement in the psychotherapy group compared to usual care, at 12 weeks: RR 1.59 (95%CI 1.13, 2.23). This corresponded to an NNT of 5 (95%CI 3, 15), for a control group rate of 38%. However, there was no significant difference at 12 months follow-up. We noted, however, that 10% of the usual therapy patients were given psychotherapy in the follow-up period, which may have reduced the relative effectiveness of the psychotherapy arm.

Figure 5:
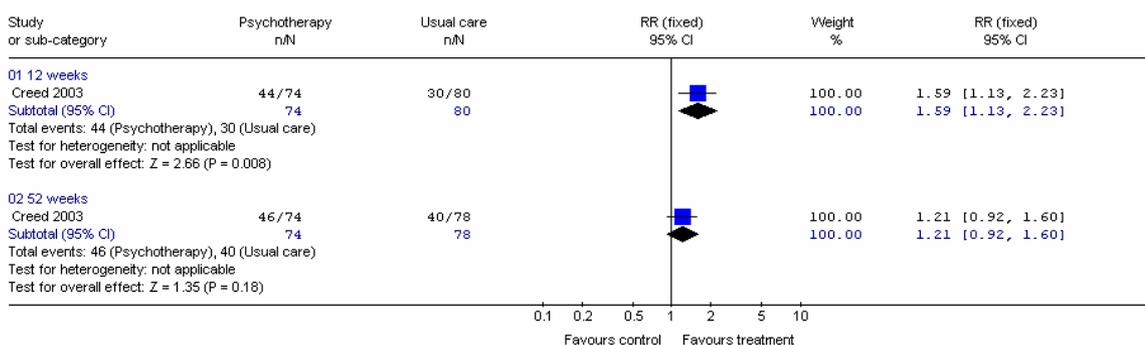
2. Individual symptoms

a) Pain

Pain (change in VAS score on a scale of 100 mm) was reported by Creed (2003) at 12 and 52 weeks. There was no significant difference between interventions at either time.

Figure 6:
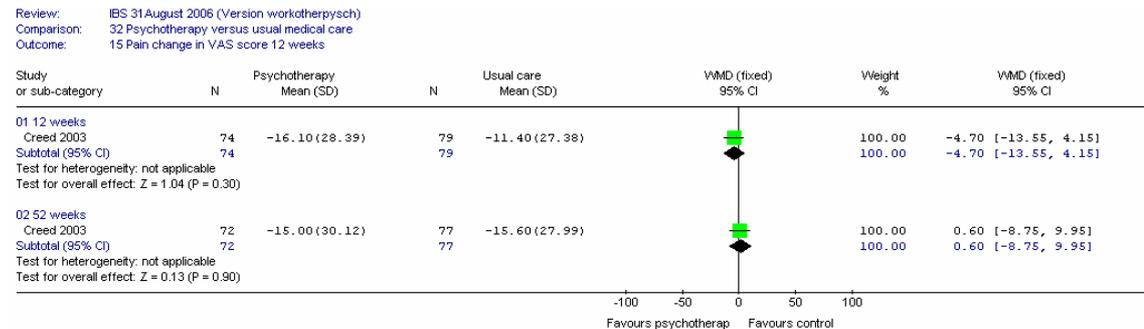
3. Quality of life

Creed (2003) measured the health-related quality of life using SF-36, recording both the physical and mental components. We noted that there was about 30% of missing data for these outcomes at three months follow-up.

a) Physical component

The physical component change score was reported at 12 weeks and at 52 weeks. On this 0 to 100 scale, an increase is a benefit and a negative change score is a worsening. There was a small statistically significant improvement favouring psychotherapy at 12 weeks, which increased slightly at 52 weeks to mean difference 5.50 (95%CI 2.13, 8.87) for a control group value of 38.1.

Figure 7:
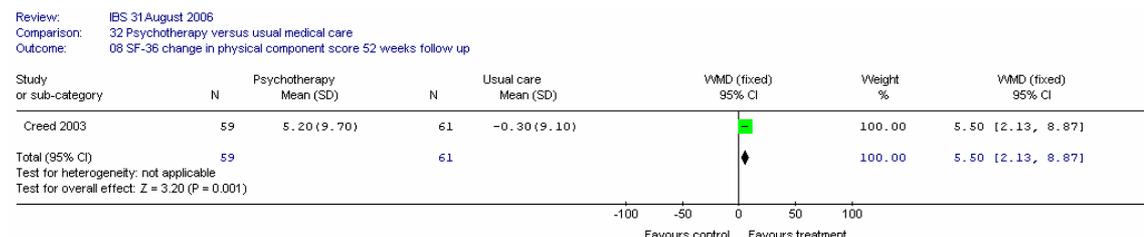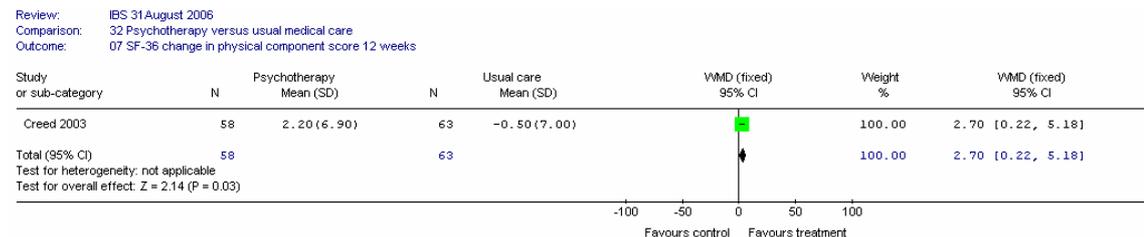
b) Mental component

The SF-36 mental component change score was reported by Creed (2003) at 12 weeks and at 52 weeks. There was a small statistically significant difference at 12 weeks, but no significant difference at one year.

Figure 8:

Review: IBS 31 August 2006
Comparison: 32 Psychotherapy versus usual medical care
Outcome: 09 SF-36 change in mental component score 12 weeks
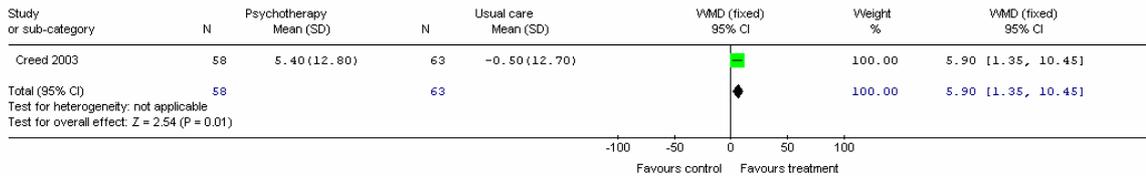
Study or sub-category	N	Psychotherapy Mean (SD)	N	Usual care Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Creed 2003	58	5.40 (12.80)	63	-0.50 (12.70)		100.00	5.90 [1.35, 10.45]
Total (95% CI)	58		63			100.00	5.90 [1.35, 10.45]

Test for heterogeneity: not applicable
Test for overall effect: Z = 2.54 (P = 0.01)

-100 -50 0 50 100
Favours control   Favours treatment

Review: IBS 31 August 2006
Comparison: 32 Psychotherapy versus usual medical care
Outcome: 10 SF-36 change in mental component score 52 weeks follow up
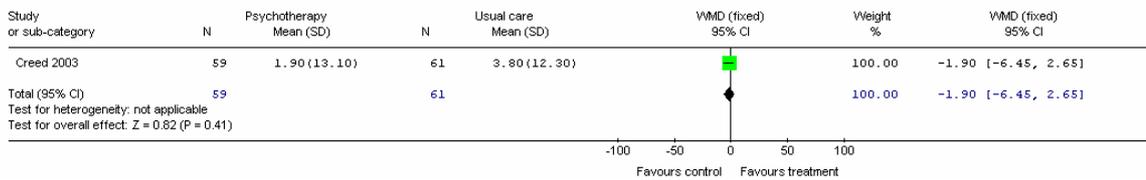
Study or sub-category	N	Psychotherapy Mean (SD)	N	Usual care Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Creed 2003	59	1.90 (13.10)	61	3.80 (12.30)		100.00	-1.90 [-6.45, 2.65]
Total (95% CI)	59		61			100.00	-1.90 [-6.45, 2.65]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.82 (P = 0.41)

-100 -50 0 50 100
Favours control   Favours treatment

4. Number of patients receiving prescriptions for antidepressants in follow up year

Creed (2003) compared the number of patients receiving prescriptions for antidepressants in follow-up year. There was no significant difference between interventions.

Figure 9:

Review: IBS 31 August 2006 (Version workotherpysch)
Comparison: 32 Psychotherapy versus usual medical care
Outcome: 13 Number of patients requiring prescriptions over 12 m
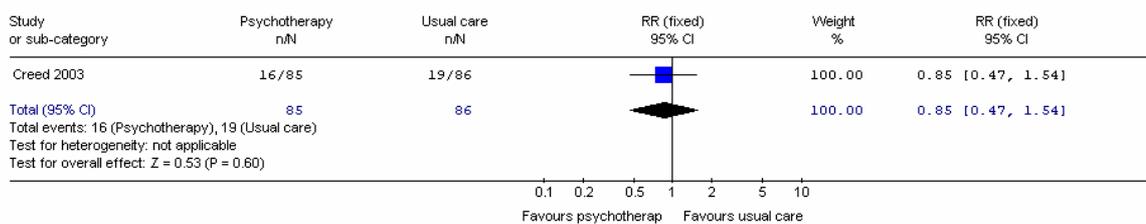
Study or sub-category	Psychotherapy n/N	Usual care n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Creed 2003	16/85	19/86		100.00	0.85 [0.47, 1.54]
Total (95% CI)	85	86		100.00	0.85 [0.47, 1.54]

Total events: 16 (Psychotherapy), 19 (Usual care)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.53 (P = 0.60)

0.1 0.2 0.5 1 2 5 10
Favours psychotherap   Favours usual care

5. Number of patients discontinuing treatment

Creed (2003) also reported the number of patients in each group who discontinued treatment. There was a large, statistically significant difference between interventions, favouring usual care.

Figure 10:
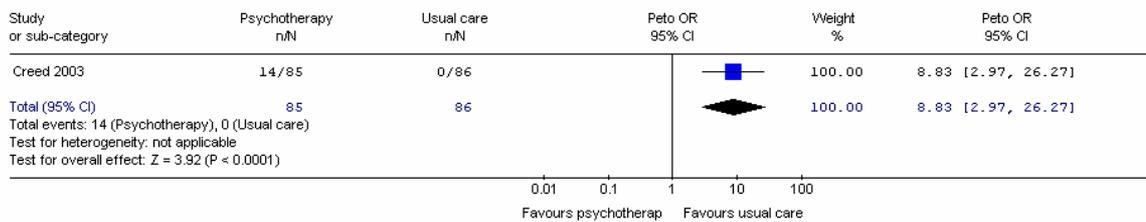
C. Psychotherapy versus antidepressant medication (SSRI - paroxetine)

1. Global symptoms

a) Global improvement in symptoms (number of patients)

The numbers of patients improved were reported by Creed (2003) at 12 weeks and at 52 weeks. There was no significant difference at either time, although we noted that the 52 week results are likely to be confounded by the use of antidepressants in both arms.

Figure 11:
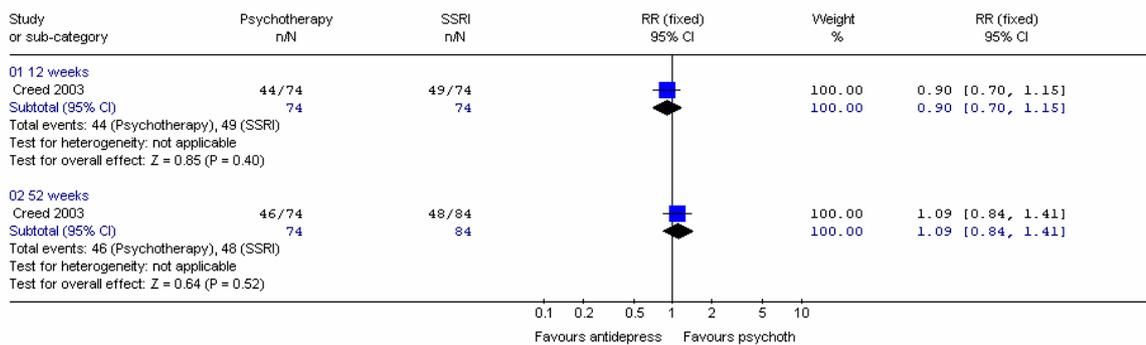
2. Individual symptoms

a) Pain

Pain (change in VAS score) was reported by Creed (2003) at 12 weeks. There was no significant difference between interventions.

Figure 12:
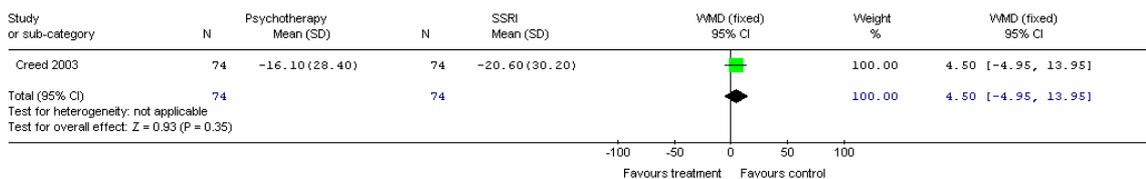
3. Quality of life

Creed (2003) measured the health-related quality of life using SF-36, recording both the physical and mental components. We noted that there was about 30% of missing data for these outcomes at three months follow-up.

a) Physical component

The physical component change score was reported at 12 weeks. There was no significant difference between interventions.

Figure 13:
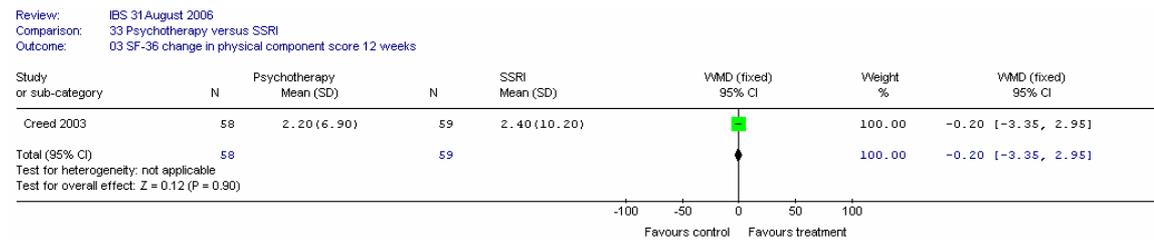
b) Mental health outcomes

The SF-36 mental component change score was reported by Creed (2003) at 12 weeks. There was no significant difference between interventions.

Figure 14:
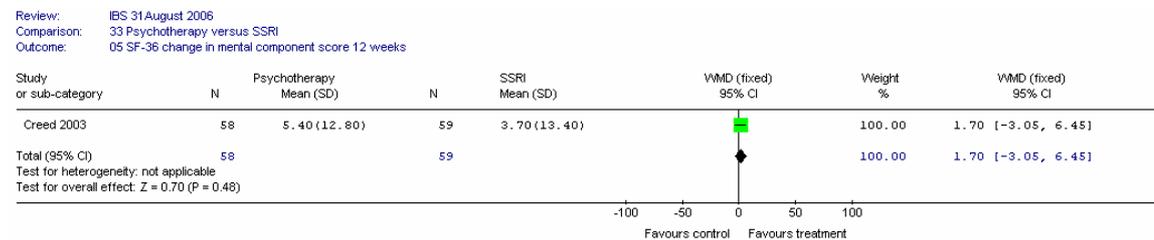
4. Number of patients receiving prescriptions for antidepressants in follow up year

Creed (2003) compared the number of patients receiving prescriptions for antidepressants in follow-up year. There was a statistically significant difference between interventions, favouring psychotherapy: RR 0.45 (95%CI 0.27, 0.75). This corresponded to a number needed to harm of 5 (95%CI 3, 10), for an antidepressant group rate of 42%.

Figure 15:

Review: IBS 31 August 2006 (Version workotherpysch)
Comparison: 33 Psychotherapy versus SSRI
Outcome: 11 Number of patients requiring prescriptions over 12 m
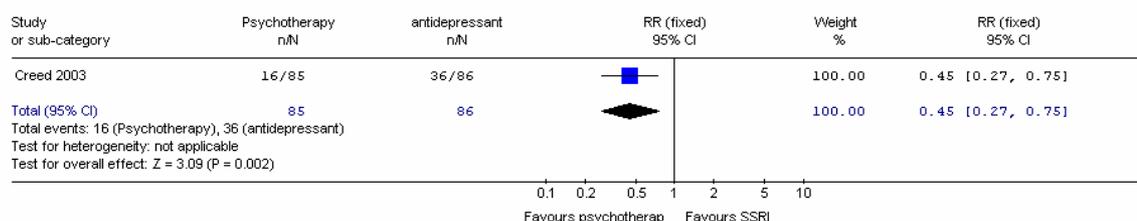
Study or sub-category	Psychotherapy n/N	antidepressant n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Creed 2003	16/85	36/86		100.00	0.45 [0.27, 0.75]
Total (95% CI)	85	86		100.00	0.45 [0.27, 0.75]

Total events: 16 (Psychotherapy), 36 (antidepressant)
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.09 (P = 0.002)

0.1 0.2 0.5 1 2 5 10
Favours psychotherap   Favours SSRI

5. Number of patients discontinuing treatment

Creed (2003) also reported the number of patients in each group who discontinued treatment. There was a statistically significant difference between interventions, favouring psychotherapy. This gave an NNH of 6 (95%CI 4, 20), for an antidepressant group rate of 34%.

Figure 16:

Review: IBS 31 August 2006 (Version workotherpysch)
Comparison: 33 Psychotherapy versus SSRI
Outcome: 10 Number of patients discontinuing treatment

Study or sub-category	Psychotherapy n/N	Antidepressant n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Creed 2003	14/85	29/86		100.00	0.49 [0.28, 0.86]
Total (95% CI)	85	86		100.00	0.49 [0.28, 0.86]

Total events: 14 (Psychotherapy), 29 (Antidepressant)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.49 (P = 0.01)

0.1 0.2 0.5 1 2 5 10
Favours psychotherap   Favours SSRI
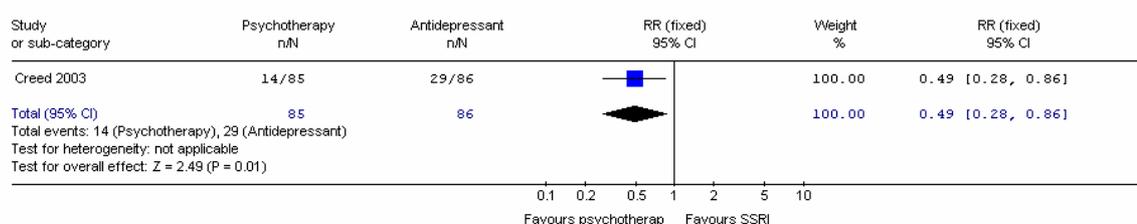
ECONOMIC LITERATURE FOR PSYCHOTHERAPY

One relevant health economic analysis was identified on the cost-effectiveness of psychotherapy in the treatment of IBS. Creed (2003) was a trial based economic evaluations conducted in the UK which recruited patients from secondary and tertiary care with severe IBS. This study aimed to assess whether psychotherapy would be superior to usual care in reducing abdominal pain and improving quality of life and whether these improvements could be achieved at no additional cost due to treatment costs being offset by reduced health care costs. It also included a comparison of SSRIs with usual care. The patient population considered were secondary and tertiary care patients with severe IBS who had not responded to usual treatment. The included patients had a mean duration of IBS of 8 years. This study was considered to be relevant to patients with refractory IBS only. The psychotherapy intervention consisted of 8 sessions over 3 months delivered by trained therapists. After three months, patients in the psychotherapy arm returned to their GP and received usual care for one year during which time they were followed-up. In the comparator arm patients received usual care from either their gastroenterologist or their GP for the three month treatment period and the following year of follow-up. The primary outcome was abdominal pain measured on a VAS of severity with secondary outcomes considering days with pain, overall change in symptoms and HRQofL measured by the SF-36. Direct health care costs per week were estimated for the intervention

and follow-up periods. This included hospital costs (inpatient days, outpatient, day-patient and A&E attendances), primary care costs (GP surgery and home visits, practice nurse and practice based counsellor visits), domiciliary care services (NHS and PSS) and day centres, use of alternative therapies and prescribed medications. These were adjusted to allow for any differences before baseline and bias corrected 95% confidence intervals were presented. The cost of providing psychotherapy was not presented separately from the other direct health care costs.

The number of patients with an improvement in global symptoms was significantly higher for psychotherapy at the end of treatment compared to usual care. The clinical outcomes from this trial have been summarised in detail in the clinical effectiveness review. Direct health care costs were significantly increased for psychotherapy compared to usual care during the intervention period and were significantly decreased during the following year. There was no significant increase in direct health care costs over the 15 month trial period.

This study was a partial economic evaluation as it did not assess the incremental cost of any benefit achieved in the form of a cost-effectiveness ratio. The evidence provided by this study was considered to be indirect as the patients were recruited from secondary and tertiary care and costs may differ for refractory patients managed in primary care. No potential areas of significant bias were identified. Direct health care costs were significantly increased by psychotherapy during the intervention period and they were significantly increased during the follow-up period. However, the study was powered to detect a specific change in clinical rather than cost outcomes. As this study did not provide an estimate of the cost per QALY for psychotherapy compared to usual care, it was not particularly useful in determining whether recommending psychotherapy for use in the NHS would result in the efficient use of NHS resources.

COST-EFFECTIVENESS ANALYSIS FOR PSYCHOTHERAPY

The section describes the health economic analysis undertaken to inform recommendations on the use of psychotherapy as one-off intervention in the management of IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

- The effectiveness of psychotherapy in addition to usual care compared to usual care alone in people with refractory IBS was based on the number of patients with an improvement in global symptoms (at the end of treatment) for psychotherapy vs usual care (RR = 3.08, 95%CI 1.74 – 5.47, based on Guthrie 1991).
- We assumed that the relative risk of response had fallen to 1.68 (95% CI 1.14 – 2.49) by 15 months based on follow-up data from Svedlund (1983).

- The evidence included in the clinical effectiveness review did not allow a subtype specific estimate of clinical effectiveness to be estimated. Therefore it was assumed that psychotherapy is equally effective in all IBS subtypes although it should be noted that the trials were carried out in patients with refractory IBS and a significant proportion of the trial population had a current or previous psychiatric diagnosis.
- It was assumed that psychotherapy is given over 12 weeks as this was the duration used in RCTs.
- The costs of psychotherapy were estimated from the number and duration of sessions and the unit cost for face-to-face time with a psychotherapist. The mean duration of psychotherapy from the three RCTs (Svedlund 1983; Guthrie 1991; Creed 2003) was 9.9 hours. The cost of face-to-face time with a psychotherapist was based on the reference cost for counselling services in primary care (£48 per hour) on advice from GDG members that psychotherapy is provided by counsellors and nurses in the NHS at present. This gave a total cost of £471 (range £348 - £672) over 12 weeks.

In the Creed (2003) economic analysis, there was a statistically significant lower cost per week for psychotherapy vs usual care (-£8.11 to -£0.04 95% CI) in the year following the intervention period (costs have been uplifted to reflect current prices). We took the mid-point of this interval which gave a cost per week of £-4.08 for psychotherapy compared to usual care in the year following treatment. This was applied in the basecase analysis for psychotherapy resulting in cost savings of £212 over the year following intervention.

The results of the Creed (2003) study also suggested that there was a reduction in service use during the intervention period, as whilst the weekly costs during this period were significantly higher for psychotherapy compared to usual care (£1.23 to £15.30), these were not comparable to the cost of providing psychotherapy which would be expected to have a mean cost of £40 per week based on the number and duration of sessions provided in this study and the unit costs presented above. This suggested that there was a significant reduction in other health care use during the intervention period. As it was not possible to separate the costs of psychotherapy from the costs offsets due to lower resource use during the intervention period of the Creed (2003) study, we have based the intervention cost on the estimate described above. However, as it seems reasonable that patients will access NHS services less whilst they are receiving an effective intervention, we have applied the reduction in resource use seen in the follow-up period as a cost off-set during the intervention period. The direct costs measured by Creed (2003) during the intervention period have been considered in a sensitivity analysis.

Modelled response rates

In the basecase scenario the response rate of 25% in the no treatment arm was taken from the mean placebo arm response rate from the behavioural therapy trials. This represents the group of patients whose symptoms improve under usual care. The RR for an improvement in global

symptoms for psychotherapy vs no treatment at the end of treatment is 3.08; therefore the response rate in the intervention arm is 78% after 12 weeks. The response rate has fallen to 43% by 15 months based on the 15 month follow-up data from Svedlund (1983). The time-frame of the analysis was limited to 15 months, with no further costs or benefits accrued beyond this time-point, as this was the longest follow-up available for psychotherapy in IBS.

Figure 17: Response rate in the basecase analysis
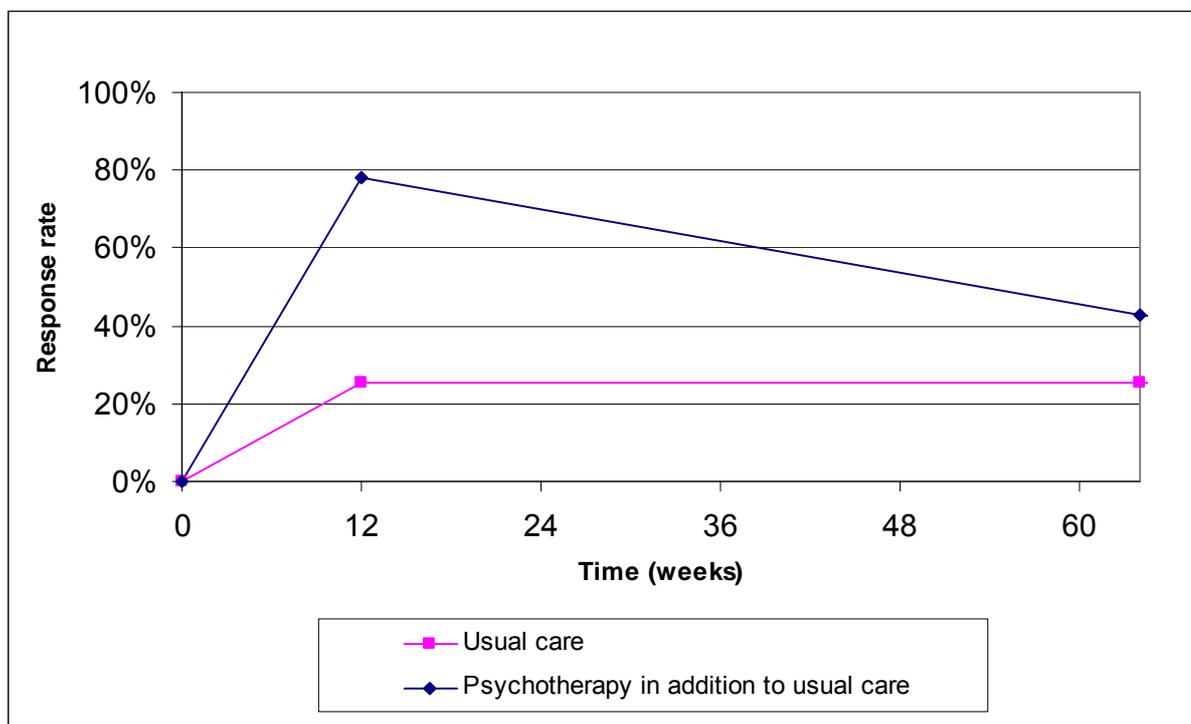
Table 1: Intervention specific parameters – Psychotherapy

Description	Value	Evidence
RR of response for intervention vs placebo (at end of treatment)	3.08 (1.74 – 5.47)	Guthrie (1991)
RR of response for intervention vs placebo (at 15 months)	1.68 (1.14 – 2.49)	Svedlund (1983)
Cost for psychotherapy	£471 (range £348 - £672) over 12 weeks	Weighted mean duration across studies and unit cost form Netten (2006)
Cost-offset due to reduced resource use during intervention and 1 year after	-$4.70 (95%CI -$9.34 to -$0.05) per week Equiv to -£4.08 (95% CI -£8.11 to -£0.04) at current UK prices*	Creed (2003)

* Uplifted to 05/06 prices using Hospital and Community Health Services Pay and Prices

Index, Netten (2006)

Psychotherapy in addition to usual care for 100 patients with refractory IBS is estimated to gain an additional 2.94 QALYs for an additional cost of £21,035 compared to usual care alone under the basecase assumptions. The incremental cost per QALY is therefore £7,160. The probabilistic sensitivity analysis considers the uncertainty in this basecase estimate due to the uncertainty in the parameters used to estimate the cost-effectiveness. The CEAC in Figure 18 shows that given the parameter uncertainty, psychotherapy in additional to usual care has a 84% probability of having a cost per QALY under £20,000 and a 92% probability of having a cost per QALY under £30,000, compared to usual care alone.

Figure 18: CEAC for psychotherapy in addition to usual care compared to usual care alone in patients with refractory IBS
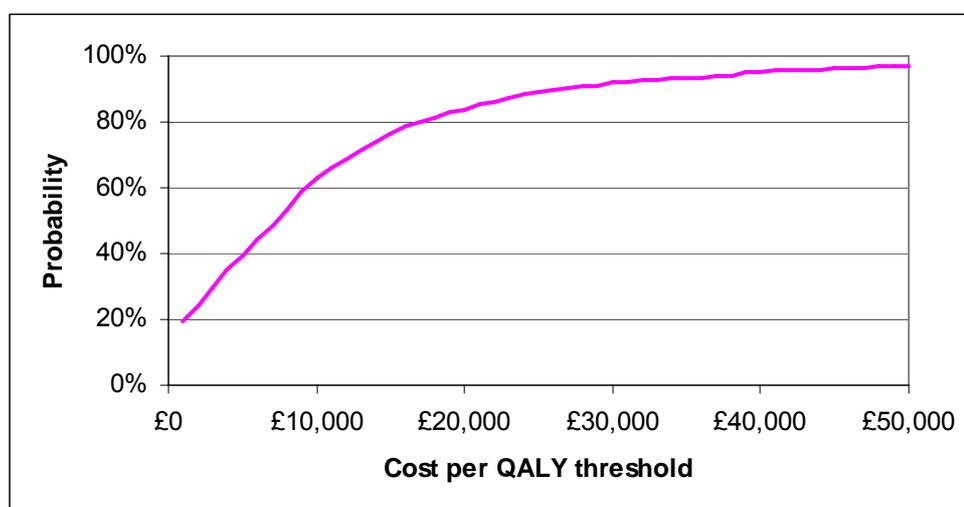
The incremental cost-effectiveness is dependent on the probability of an improvement for patients who receive usual care. When we applied a lower response rate of 9% in the usual care arm, the cost per QALY was increased to £14,629. As this sensitivity analysis significantly increased the cost per QALY estimate, the probabilistic sensitivity analysis was re-run using this lower response rate for the comparator arm. The mean cost per QALY from the 1000 samples was £17,577 and the cost per QALY had a 51% probability of being under £20,000 per QALY and a 62% probability of being under £30,000 per QALY.

The threshold analysis showed that a response to treatment would need to provide more than 0.026 QALYs per annum to give a cost per QALY of under £20,000 in the basecase analysis. When the utility gain associated with a response to treatment was increased to 0.135 (equivalent to the QALY gain expected for a complete remission of symptoms) the cost per QALY was significantly lower at £3,782.

When we assumed no fall-off in response up to 52 weeks post-intervention, the cost per QALY was £5,000. This would be further reduced by any continued response beyond 52 weeks. When

we assumed that there was no significant difference between psychotherapy and usual care at 12 months, the cost per QALY increased to £9,062.

When we estimated the incremental cost during the intervention and follow-up period directly from those measured in Creed (2003), psychotherapy was cost saving compared to usual care and resulted in greater QALY gain, dominating usual care. When we excluded the reduction in resource use observed in the Creed (2003) study from the analysis, the incremental cost of psychotherapy increased significantly to £47,154 and the cost per QALY increased to £16,051. As this is a significant increase in the cost per QALY, the probabilistic analysis was re-run under this assumption. In this conservative estimate the cost per QALY for psychotherapy in addition to usual care compared to usual care alone had a 57% probability that of being under £20K and a 77% probability of being under £30K.

Table 2: Sensitivity results for psychotherapy in addition to usual care compared to usual are alone for 100 patients with refractory IBS (all subtypes)

Scenario	Usual care		Behavioural intervention and usual care		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	2.02	£21,035	4.96	£7,160
Lower response rate in comparator arm (9%)	£0	1.30	£21,035	2.74	£14,629
No fall-off in effect for 1 year	£0	2.02	£21,035	6.23	£5,000
Effect falls off over 12 months	£0	2.02	£21,035	4.34	£9,062
No resource use reduction	£0	2.02	£47,154	4.96	£16,051
Costs as measured in Creed (2003)	£0	2.02	-£11,307	4.96	-£3,849 Psychotherapy dominates
High utility gain of 0.135	£0.00	3.83	£21,035	9.39	£3,782
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.025 and 0.026.				

Further analyses on the cost-effectiveness of psychotherapy compared to other behavioural interventions are given in section 9.7.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. There is fair evidence to show a significant global improvement in symptoms after 12 weeks and after 15 months, for dynamic psychotherapy in addition to medical therapy compared with medical therapy alone, given in secondary care to patients with long term or refractory IBS, approximately half of whom had psychological co-morbidities.

2. There is fair evidence to show a significant global improvement in symptoms after 12 weeks, for psychodynamic interpersonal therapy compared with medical treatment, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

3. There is weak evidence to show no significant global improvement in IBS symptoms after 12 months follow up, for psychodynamic interpersonal therapy compared with medical treatment, given in secondary care to patients with long term IBS, many of whom had had psychological co-morbidities. This fall-off in effect may have been caused by confounding in the control arm.

4. There is moderately good evidence to show a significant decrease in pain after 12 weeks and 15 months, for dynamic psychotherapy in addition to medical therapy compared with medical therapy alone, given in secondary care to patients with long term IBS, many of whom had had psychological co-morbidities.

5. There is fair evidence to show no significant reduction in pain after 12 weeks and 12 months follow up, for psychodynamic interpersonal therapy compared with medical treatment, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

6. There is weak evidence to show no significant difference in the patients' assessment of their mental health at 15 months, between patients given dynamic psychotherapy in addition to medical therapy compared with medical therapy alone, either for all patients or for a subgroup with a psychological co-morbidity history.

7. There is weak evidence to show a small, significant improvement in the physical component of the SF36 quality of life score after 12 weeks and 12 months follow up, and a small significant difference in the mental health score after 12 weeks, for psychodynamic interpersonal therapy compared with medical treatment, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities. There was no significant difference in the mental health score at 12 months follow up.

8. There is weak evidence to show no significant difference in the number of patients requiring a prescription for antidepressants over 12 months, for psychodynamic interpersonal therapy

compared with medical treatment, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

9. There is weak evidence to show significantly more patients discontinued treatment for psychodynamic interpersonal therapy compared with medical treatment, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

10. There is fair evidence to show no significant difference in global improvement in symptoms, pain and quality of life (physical and mental) after 12 weeks, between psychodynamic interpersonal therapy compared with an SSRI, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

11. There is weak evidence to show that significantly fewer patients required a prescription for antidepressants over 12 months, for psychodynamic interpersonal therapy compared with an SSRI, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

12. There is weak evidence to show significantly fewer patients discontinued treatment for psychodynamic interpersonal therapy compared with an SSRI, given in secondary care to patients with refractory IBS, approximately half of whom had psychological co-morbidities.

HEALTH ECONOMIC STATEMENT

Evidence from a trial based economic evaluation in secondary and tertiary care patients with a high level of NHS service use at baseline showed that direct health care costs are lower in the year following treatment for 3months of psychotherapy compared to 3 months of usual care. Health care costs were significantly higher during the intervention period for psychotherapy compared to usual care. This evidence is unlikely to be applicable to primary care patients except those with refractory IBS.

Evidence from a decision analytic model showed that the addition of psychotherapy to usual care is cost-effective in individuals with refractory IBS although the cost-effectiveness was sensitive to uncertainty around the proportion of patients experiencing an improvement in global symptom score with usual care alone. It was also sensitive to the whether or not there was any reduction in health care service use during and following treatment

GDG DISCUSSION

Despite the prevalence of psychiatric comorbidities in the trials for this review, the GDG considered that dynamic psychotherapy could be a useful option for all people with refractory IBS, and had potential to be a first line therapy. The GDG therefore decided to include

pyschotherapy in one of its top five research recommendations, with the potential for this intervention to be considered as a first line therapy option.

EVIDENCE TO RECOMMENDATION

The evidence to recommendation statement for psychotherapy, CBT and hypnotherapy is detailed in section 9.8.

The combined guideline recommendation for psychotherapy, CBT and hypnotherapy is also stated in section 9.8.

9.5 Cognitive behavioural therapy

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, except that crossover studies were excluded as inappropriate due to the carry-over effect of the CBT interventions.

The following comparisons were to be included:
- CBT versus waiting list control or symptom monitoring only
- CBT type 1 versus type 2 (e.g. stress management versus contingency management)
- CBT individual versus CBT group
- CBT + another intervention (e.g. medical therapy) versus the other intervention alone
- CBT versus medical treatment.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the PSYCINFO database was searched for this review. The search strategies are listed in Appendix B.

Study Design

Seventeen parallel randomised trials were included (Bennett 1985; Bergeron 1983; Blanchard 1993; Bogalo 2006; Boyce 2003; Corney 1991; Drossman 2003; Fernandez 1998; Gong 2002; Greene 1994; Heymann-Monnikes 2000; Kennedy 2005; Lynch 1989; Payne 1995; Tkachuk 2003; Toner 1998; Vollmer 1998). Further details are given in the included studies table.

Six studies had more than one arm: Bergeron (1983) (stress management, relaxation, biofeedback); Boyce (2003) (CBT, relaxation, routine medical care); Drossman (2003) (half of the patients were randomised to CBT or attention control; the other half to desipramine or placebo tablet; the two halves were divided randomly initially); Fernandez (1998) (stress management, contingency management, conventional medical treatment or placebo); Payne

(1995) (CBT, self help group or waiting list control); Vollmer (1998) (CBT in a group, CBT individually or symptom monitoring waiting list control). Thus, there were 28 possible comparisons for this review. We did not include the comparison with desipramine because this drug is not licensed for use in the UK.

Two studies were reported only as abstracts (Bergeron 1983; Gong 2002). The former had no data reported. Three studies were RCTs with limited or incomplete data (Bennett 1985; Toner 1998; Bogalo 2006). The Bennett study gave little detail (e.g. it was unclear how many patients were assigned to each group; no primary data were given; only p values for ANOVAs). The Toner (1998) study was briefly reported as part of a larger article. This study enrolled 101 individuals with IBS, who were randomly allocated to group CBT, a psychoeducational group treatment (education about the condition and support) or usual medical treatment. No primary outcome data were reported. Bogalo (2006) appeared to use data from a randomised trial. However, this paper only reported outcomes for the intervention group not the controls. For the rest of this review, only the 14 studies with sufficient data are reported.

The studies were carried out in the UK (Corney 1991; Kennedy 2005), Europe, Canada, the USA and China. Trials lasted between 6 and 12 weeks (the duration of the Gong (2002) study was not stated). One study was conducted in primary care (Kennedy 2005); seven were in secondary care (Blanchard 1993; Corney 1991; Drossman 2003; Fernandez 1998; Gong 2002; Heymann-Monnikes 2000; Lynch 1989); one study (Boyce 2003) recruited equal numbers of patients through gastroenterology clinics and through newspaper advertisements, and the others did not report the setting. The total number of patients in the studies ranged from 20 to 431, with three studies including more than 100 patients (Boyce 2003; Drossman 2003; Kennedy 2005). All but four studies (Boyce 2003; Drossman 2003; Gong 2002; Kennedy 2005) had fewer than 25 patients in the treatment arm.

Population

All the studies included only patients with IBS, apart from Drossman (2003), for which 78% of the patients had IBS. None of the studies reported the number of patients with bloating or whether the symptoms were post-infective. The mean age of patients across studies was around 40 years, with those aged between 16 and 70 years included. All the studies included more women than men.

Four studies reported or implied that the patients had refractory IBS (Greene 1994; Heymann-Monnikes 2000; Lynch 1989; Tkachuk 2003): Tkachuk (2003) stated that the patients had refractory IBS; Lynch (1989) had patients referred from gastroenterology clinics and had duration of IBS of around 9 years; Heymann-Monnikes (2000) had tertiary referral patients. The patients in Greene (1994) had had IBS for 14.5 years. Boyce (2003) reported that about 50% of

patients had received medication for IBS. The other studies did not give any information about IBS duration.

Most studies reported whether the patients had an Axis I diagnosis (e.g. major depression; schizoaffective disorder; paranoid state) or other psychiatric or psychological co-morbidities:

- Heymann-Monnikes (2000) stated that patients were excluded if they had any mental disorder.
- Drossman (2003) reported that no patients had psychiatric disorders, but almost half had a history of physical or sexual abuse.
- Boyce (2003) excluded patients if they had major current psychotic illness, used psychological treatment, or antidepressants or antipsychotics.
- Two studies excluded patients with serious psychiatric disorders (Lynch 1989 (implied); Vollmer 1998).
- Five studies stated that the majority of patients had an Axis I diagnosis (Blanchard 1993, 50-73%; Greene 1994, 90%; Payne 1995, 80-92%; Tkachuk 2003, 68%; Vollmer 1998, 82-90%).
- Three studies mentioned that the patients had received psychiatric treatments: in Fernandez (1998) 49% had had psychiatric treatment; in Kennedy (2005) 43% had consulted a doctor because of a psychological problem; in Lynch (1989) 6/21 (29%) patients used psychotropic drugs.
- Corney (1991) had over 50% of the patients with one or more social problems.
- The remaining studies (Bergeron 1983; Gong 2002) did not mention Axis I diagnoses or psychiatric complaints/treatments.

Overall, therefore, eight studies can be considered to be representative or partly representative of the IBS population with concurrent psychiatric or psychological illnesses (Blanchard 1993; Fernandez 1998; Greene 1994; Kennedy 2005; Lynch 1989; Payne 1995; Tkachuk 2003; Vollmer 1998). Three studies can be considered to be in patients with IBS who do not have concurrent psychiatric or psychological illnesses (Heymann-Monnikes 2000; Drossman 2003; Boyce 2003). We noted that 15 to 20% of patients in primary care have a co-existing psychiatric condition and approximately half of those in secondary care.

Interventions

Twelve studies described some form of CBT: Boyce (2003); Corney (1991); Drossman (2003); Fernandez (1998) (one group had stress management sessions involving progressive muscle relaxation; another had contingency management sessions involving contingency contract for new behaviours; self-observation, shaping, stimulus control, neutralising inadequate habits and breaking learned illness behaviour, social skills training); Gong (2002); Greene (1994); Heymann-Monnikes (2000); Kennedy (2005); Lynch (1989); Payne (1995); Tkachuk (2003); Vollmer (1998). Two studies gave relaxation training (Bergeron 1983; Blanchard 1993).

The CBT interventions differed in the methods used (group or individual; administered by nurse practitioner or other professional); the number of sessions given (6 to 12), and; the duration of the study (6 to 12 weeks in the studies that stated this).

CBT was defined in different ways in different papers. Two studies had therapy delivered by nurses (Corney 1991; Kennedy 2005). Details are given in the included studies table (Appendix D).

It was decided to combine all CBT, behavioural therapy (BT) and cognitive therapy (CT) interventions under the general heading of CBT.

CBT was compared with relaxation training; symptom monitoring; self-help support groups; medical therapy or placebo. Two studies had placebo as a comparator (Drossman 2003; Fernandez 1998). Drossman (2003) randomised the patients into two groups and then randomised one group to CBT and attention control and the other group to desipramine and placebo desipramine. We decided to classify the CBT-attention control comparison as placebo and treat the desipramine placebo as no treatment. Fernandez (1998) had a placebo condition consisting of giving the patient exercises to visualise bowel function; and prompting their own capacity for self regulation through thought.

For the purposes of this review, we combined the comparators waiting list control, symptom monitoring, no treatment and placebo condition. The following comparisons were included:

1. CBT versus a waiting list control group, symptom monitoring only or placebo: nine studies, 12 comparisons (Blanchard 1993; Drossman 2003 x2; Fernandez 1998 x2; Gong 2002; Greene 1994; Lynch 1989; Payne 1995; Tkachuk 2003; Vollmer 1998 x2):
   o CBT versus waiting list control (Lynch 1989)
   o CBT versus symptom monitoring (Blanchard 1993; Greene 1994)
   o CBT versus waiting list control including symptom monitoring (Payne 1995; Tkachuk 2003; Vollmer 1998 x2)
   o CBT versus an attention control condition involving symptom monitoring plus education plus access to a therapist (Drossman 2003)
   o CBT versus placebo condition (Fernandez 1998 x2)
   o CBT versus no treatment (Drossman 2003; Gong 2002);
2. CBT + another intervention versus the other intervention alone:
   o CBT + mebeverine versus mebeverine (Kennedy 2005)
   o CBT (multicomponent behavioural therapy) + optimised medical treatment versus optimised medical treatment alone (Heymann-Monnikes 2000);
3. CBT 1 versus CBT 2 (three studies):

- o Stress management versus relaxation (Bergeron 1983 – no data)
- o CBT versus relaxation (Boyce 2003)
- o Stress management versus contingency management (Fernandez 1998);

4. CBT individual therapy versus CBT group therapy:
   - o Vollmer (1998);

5. CBT versus another intervention:
   - o CBT versus biofeedback (Bergeron 1983 – no data)
   - o CBT versus support group (Payne 1995)
   - o CBT versus psychoeducational group (Toner 1998 – no data);

6. CBT versus routine medical treatment (five studies, six comparisons):
   - o CBT versus fluphenazine (anti-anxiety), mebeverine and fybogel (Bennett 1985)
   - o CBT vs 'routine medical care' (Boyce 2003 x2; Corney 1991; Fernandez 1998 x2).

Six studies stated that the patients were allowed to continue their IBS medical treatment: Fernandez (1998), although 50% did not take the medication properly; in Heymann-Monnikes (2000), 9/12 in the BT group and 11/12 in symptom monitoring had concurrent medication for IBS; Kennedy (2005); Lynch (1989), 10/21 had analgesics at recruitment and 6 used Metamucil or similar bulking agents; Tkachuk (2003); Vollmer (1998), no patients were excluded because of their ongoing IBS drug treatment. Gong (2002) reported that all patients received selective gastrointestinal calcium antagonists.

One study (Boyce 2003) stated that the 'vast majority' were not taking concurrent medication. The rest did not state the concurrent medications for IBS.

In view of the use of other IBS medication in both arms of the CBT versus placebo/symptom monitoring comparisons, we decided to combine interventions 1 and 2 as subgroups because they each could be considered to be variations of CBT versus no treatment/symptom monitoring. Furthermore, the GDG advised that CBT treatment would not interact with other medical treatments. We decided that if subgroup analysis were to be used, the attention control and placebo condition (Fernandez 1998; Drossman 2003) should be considered as a separate group to the other comparators.

Outcomes
The outcomes examined were:
1. Global symptoms:
   a) Global improvement in symptoms (number of patients)
   b) Global symptom score.
2. Individual symptoms:
   a) Pain
   b) Bloating

c)   Bowel habits.

3.   Adverse events

4.   Quality of life

5.   Number of patients withdrawing from the study

6.   Mental health outcomes (overall mental health; depression; anxiety)

Five studies recorded longer term follow up: Vollmer (1998) at 12 weeks; Corney (1991) at 40 weeks; Boyce (2003) and Kennedy (2005) both at 26 and 52 weeks; Fernandez (1998) at 52 weeks.

The following outcome measures were recorded:

Table 1.

Outcome measure	Measured at treatment end (weeks)	Measured at follow up (weeks)
Global scores		
Global IBS symptom score (Bowel symptom severity scale 0-48 scale; high = severe)	Boyce 2003 (8)	Boyce 2003 (26) Boyce 2003 (52)
Global IBS symptom score (satisfaction, global wellbeing, diary pain scores, health related quality of life); high score = good; maximum unclear	Drossman 2003 (12)	
Global IBS symptom score (7 symptoms daily for 2 weeks each scored 0-4 from not a problem to debilitating symptom intensity; high score = bad)	Greene 1994 (8)	
Global IBS symptom score (20 items over 14 days); high score = bad	Heymann-Monnikes 2000 (10)	
Global IBS symptom score (<75=normal; 75-174 mild; 175-299 moderate; 300-500 severe symptoms)	Kennedy 2005 (6)	Kennedy 2005 (26) Kennedy 2005 (52)
Global IBS symptom score (pain, discomfort, diarrhoea, constipation each rated 0=no symptoms to 6=very severe symptoms)	Lynch 1989 (8)	
Global improvement of IBS symptoms (mean Composite Primary Symptom Reduction [CPSR] score; CPSR represents % reduction in score from baseline); i.e. high = bad	Blanchard 1993 (8) Greene 1994 (8) Payne 1995 (8) Vollmer 1998 (10)	Vollmer 1998 (12)
Global improvement of IBS symptoms (mean score) VAS (1=very much better to 7=very much worse; 4=no change)	Heymann-Monnikes 2000 (10)	
Global improvement of IBS	Blanchard 1993 (8)	Fernandez 1998 (52)

symptoms (no. patients)	Fernandez 1998 (10) Gong 2002 (duration of treatment not stated) Greene 1994 (8) Lynch 1989 (8) Payne 1995 (8) Tkachuk 2003 (9) Vollmer 1998 (10)	
Pain scores		
Pain score (0=not a problem; 4=debilitating symptoms) daily for 4 weeks	Blanchard 1993 (8)	
Pain score (VAS score); high = bad.	Corney 1991 (16)	Corney 1991 (40)
Pain score (McGill average daily pain); high score = bad	Drossman 2003 (12)	
Pain score (0=not a problem; 1=mild; 2=moderate; 3=severe; 4=debilitating; scores totalled for 1 week); high score = bad	Fernandez 1998 (10)	
Pain score (scored 0-4 daily for 2 weeks); high score = bad	Greene 1994 (8)	
Pain score (0=no pain; 6=very severe pain)	Lynch 1989 (8)	
Pain score (0=not a problem; 4=intense and incapacitating)	Tkachuk 2003 (9) Vollmer 1998 (10)	
Pain score (0=not a problem; 4=debilitating symptoms) daily for 4 weeks	Blanchard 1993 (8)	
Bloating		
Bloating score (0=no symptom; 6=very severe symptom)	Greene 1994 (8) Lynch 1989 (8)	
Bloating score (0=not a problem; 4=intense and incapacitating)	Tkachuk 2003 (9)	
Bloating score (0=not a problem; 4=debilitating symptoms)	Vollmer 1998 (10)	
Bowel habits		
Constipation (VAS score); high = bad	Corney 1991 (16)	Corney 1991 (40)
Constipation (0=not a problem; 4=debilitating symptoms)	Greene 1994 (8)	
Constipation (0=no symptom; 6=very severe symptom)	Lynch 1989 (8)	
Diarrhoea (VAS score); high=bad	Corney 1991 (16)	Corney 1991 (40)
Diarrhoea (0=not a problem; 4=debilitating symptoms)	Greene 1994 (8)	
Diarrhoea (0=no symptoms; 6=very severe symptoms)	Lynch 1989 (8)	
Mental health		
Dysfunctional cognitions (Cognitive Scale for Functional Bowel Disorders); scale 25-175; high=bad	Tkachuk 2003 (9)	
Mental health (max = 100)	Tkachuk 2003 (9)	
Psychological distress (Clinical		Corney 1991 (40)

Interview Schedule); 0-48; high=bad		
Overall anxiety and psychological distress (Anxiety, State-Trait Anxiety Inventory [STAI]); Scale range 20-80; high = bad	Greene 1994 (8) Heymann-Monnikes 2000 (10) Payne 1995 (8) Tkachuk 2003 (9)	
Psychological distress (HADS); range 0-56; high=bad	Boyce 2003 (8) Kennedy 2005 (8)	Boyce 2003 (26) Boyce 2003 (52) Kennedy 2005 (26) Kennedy 2005 (52)
Beck depression inventory (scale maximum 63; high=bad)	Greene 1994 (8) Heymann-Monnikes 2000 (10) Lynch 1989 (8) Payne 1995 (8) Tkachuk 2003 (9)	
Quality of life		
Disruption of daily life (0=no symptom; 6=very severe symptom)	Lynch 1989 (8)	
Physical health (SF-36 Physical Composite Scale); high = bad; scale max 100	Tkachuk 2003 (9)	
Quality of life (IBS-QOL) ; high score = good; max 84	Drossman 2003 (12)	
Quality of life (GLQI score); max score 144; high=good	Heymann-Monnikes 2000 (10)	
Work and social adjustment score (handicap in work, home, leisure and relationships; 0=not affected; 8 severely affected for each area; maximum total 40)	Kennedy 2005 (6)	Kennedy 2005 (26) Kennedy 2005 (52)
Other		
Satisfaction (responder = satisfaction 3 or more on a scale where each of 8 items were rated 1=strongly disagree to 5=strongly agree)	Drossman 2003 (12)	

Where necessary, linear scales that had a maximum value corresponding to a good outcome were inverted by subtracting the mean value from the maximum and using the same standard deviation. Studies could then be combined in a meta-analysis.

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D.

An adequate method of randomisation was reported in two studies (Drossman 2003 – computer generated; Kennedy 2005 – random number tables); the other studies did not state the method.

Allocation concealment was adequate in two studies (Drossman 2003 and Kennedy 2005 – independent third party); the other studies did not report allocation concealment.

Because of the type of intervention, the patients were not blinded. However, the GDG decided that blinding was not important for this review.

Two studies (Drossman 2003; Kennedy 2005) described an *a-priori* power calculation. All studies included in the review demonstrated baseline comparability of the groups apart from the two that were reported only as abstracts (Bergeron 1983; Gong 2002).

All the patients were followed up in four studies (Gong 2002; Greene 1994; Payne 1995; Tkachuk 2003). There were fewer than 20% missing data in five studies (Corney 1991: 2%; Drossman 2003: 7%; Heymann-Monnikes 2000: 7%; Kennedy 2005: 11%; Vollmer 1998: 6%).

There was more than 20% missing data at the end of treatment in five studies (Bennett 1985; Blanchard 1993; Boyce 2003; Fernandez 1998; Lynch 1989):

- Bennett (1985) reported 28% missing data

- Blanchard (1993) reported 6 drop-outs from relaxation group; 1 from controls; the 7 drop-outs out of the original 16 participants (44%) were replaced to give 16 participants (so not an ITT analysis); dropouts were unequal between the groups

- Boyce (2003) reported that 66 of the original 105 participants were left at the end of treatment (8 weeks; 63% left; 37% drop-out); within groups drop outs were 13/35 (37%) in the CBT group, 16/35 (46%) in the relaxation arm and 9/34 (26%) in the medical therapy group

- Fernandez (1998) reported that 33 patients dropped out at the end of treatment (16/23 (70%) from the placebo group; 6/21 (29%) for stress management; 7/23 (30%) for contingency management and 4/23 (17%) on medical treatment), i.e. 48% drop-out overall

- Lynch (1989) reported 6/27 missing data (22%; not stated from which group) and dropouts were replaced to achieve 21 patients in all.

- Drop-out was unclear in the remaining study (Bergeron 1983).

Of the five studies that reported longer term follow-up:

- Boyce (2003) reported that 52/105 patients were followed-up at one year (50% missing data)

- Fernandez (1998) had 53% missing data at 12 months

- Kennedy (2005) reported 28% and 36% loss to follow up in the CBT+ mebeverine and mebeverine groups respectively at 3 months; 26% and 24% at 6 months, and 27% and 25% at 12 months for the primary outcome (IBS symptom score)

- Corney (1991) appeared to lose one patient to follow-up at 9 months

- Vollmer (1998) had 45% loss to follow up at 3 months.

Overall, we regarded the comparisons with placebo in Fernandez (1998) as having high potential for bias and these were not included in the analysis. The other comparisons in Fernandez (1998), Blanchard (1993), Boyce (2003) and Lynch (1989) were treated with caution, especially the relaxation arm of the Boyce study. Sensitivity analyses were carried out as appropriate. We also noted that Drossman (2003) had a population in which only 78% of patients had IBS. This study was similarly treated with caution. At follow-up, Fernandez (1998), Boyce (2003) and Vollmer (1998) were treated as having high potential for bias, and Kennedy (2005) had some potential for bias.

RESULTS

A. CBT versus waiting list control group, placebo or symptom monitoring only; and CBT + IBS medication versus IBS medication alone

There were nine studies included in the analysis that compared CBT with a waiting list control group; symptom monitoring only, or; placebo condition in patients with IBS (Blanchard 1993; Drossman 2003; Fernandez 1998; Gong 2002; Greene 1994; Lynch 1989; Payne 1995; Tkachuk 2003; Vollmer 1998). Two studies compared the combination of IBS medication and CBT with IBS medication alone (Heymann-Monnikes 2000; Kennedy 2005). Heymann-Monnikes (2000) compared CBT plus optimised medical treatment versus optimised medical treatment alone. Kennedy (2005) investigated the addition of CBT to mebeverine in each arm.

1. Global symptoms

a) Number of patients with global improvement of symptoms

Eight studies with 217 patients compared CBT with symptom monitoring; no treatment, or; attention control (Blanchard 1993; Fernandez 1998; Gong 2002; Greene 1994; Lynch 1989; Payne 1995; Tkachuk 2003; Vollmer 1998 [individual and group CBT interventions]) for the outcome of global improvement in terms of the number of patients improved at the end of treatment. For this outcome measure the two CBT groups were combined in the Vollmer (1998) study.

Figure 1

Review: IBS (CBT)
Comparison: 17 CBT versus placebo
Outcome: 01 Number of patients improved

Study or sub-category	CBT n/N	control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Blanchard 1993	4/14	0/9		1.96	6.00 [0.36, 99.66]
Gong 2002	33/35	25/35		81.73	1.32 [1.05, 1.65]
Greene 1994	8/10	1/10		3.27	8.00 [1.21, 52.69]
Lynch 1989	7/11	0/10		1.71	13.75 [0.88, 213.65]
Payne 1995	9/12	1/10		3.57	7.50 [1.14, 49.54]
Tkachuk 2003	3/14	1/14		3.27	3.00 [0.35, 25.46]
VollmerB CBTgp vs pl	13/22	1/10		4.50	5.91 [0.89, 39.18]
Total (95% CI)	**118**	**98**		**100.00**	**2.32 [1.71, 3.15]**

Total events: 77 (CBT), 29 (control)
Test for heterogeneity: Chi² = 30.48, df = 6 (P < 0.0001), I² = 80.3%
Test for overall effect: Z = 5.44 (P < 0.00001)

0.01   0.1   1   10   100
Favours control   Favours CBT

The relative risk analysis (Figure 1) showed significant heterogeneity (I^2=80%, p<0.0001), attributable to Gong (2002). This was a larger study in which nearly all the patients improved with CBT and many with no treatment. A sensitivity analysis of the relative risk without Gong (2002), gave no heterogeneity (I^2=0%, p=0.99) and the result was statistically and clinically significant (Figure 2). It was not clear why Gong (2002) should be so different, however, we noted that, whilst the majority of studies had patients with a psychiatric diagnosis or treatment, the exception was Gong (2002) (not stated). In addition, Gong (2002) was an abstract and all patients received selective GI calcium antagonists; the comparator was no treatment.

The RR was 6.82 (95%CI 2.87, 16.18), for the meta-analysis of 6 studies in 146 patients. This corresponded to a number needed to treat (NNT) of 3 (95%CI 2, 3), for a baseline risk of 0 to 10%.

Figure 2

Review: IBS (CBT)
Comparison: 17 CBT versus placebo
Outcome: 01 Number of patients improved

Study or sub-category	CBT n/N	control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Blanchard 1993	4/14	0/9		10.74	6.00 [0.36, 99.66]
Greene 1994	8/10	1/10		17.90	8.00 [1.21, 52.69]
Lynch 1989	7/11	0/10		9.34	13.75 [0.88, 213.65]
Payne 1995	9/12	1/10		19.52	7.50 [1.14, 49.54]
Tkachuk 2003	3/14	1/14		17.90	3.00 [0.35, 25.46]
VollmerB CBTgp vs pl	13/22	1/10		24.61	5.91 [0.89, 39.18]
Total (95% CI)	83	63		100.00	6.82 [2.87, 16.18]

Total events: 44 (CBT), 4 (control)
Test for heterogeneity: Chi² = 0.88, df = 5 (P = 0.97), I² = 0%
Test for overall effect: Z = 4.35 (P < 0.0001)

0.01   0.1   1   10   100
Favours control   Favours CBT

Due to the high drop-out rates, a further sensitivity analysis was conducted excluding data from Blanchard (1993) and Lynch (1989). This made a slight difference (Figure 3) and therefore this figure was reported in the GRADE tables and used in the health economic modelling. This gave an NNT of 3 (95%CI 2, 4) for a control group risk of 7 to 10%.

Figure 3

Review: IBS (CBT)
Comparison: 17 CBT versus placebo
Outcome: 01 Number of patients improved

Study or sub-category	CBT n/N	control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Greene 1994	8/10	1/10		22.39	8.00 [1.21, 52.69]
Payne 1995	9/12	1/10		24.43	7.50 [1.14, 49.54]
Tkachuk 2003	3/14	1/14		22.39	3.00 [0.35, 25.46]
VollmerB CBTgp vs pl	13/22	1/10		30.79	5.91 [0.89, 39.18]
Total (95% CI)	58	44		100.00	6.11 [2.33, 16.07]

Total events: 33 (CBT), 4 (control)
Test for heterogeneity: Chi² = 0.55, df = 3 (P = 0.91), I² = 0%
Test for overall effect: Z = 3.67 (P = 0.0002)

0.01   0.1   1   10   100
Favours control   Favours CBT

One study (Fernandez 1998) reported global improvement in terms of the number of patients improved at 1 year follow-up. This was likely to be flawed because of the high drop-out rate for the no treatment group and is not reported here.

b) Global symptom improvement score (CPSR)

Four studies (Blanchard 1993; Greene 1994; Payne 1995; Vollmer 1998) reported outcomes in terms of a global IBS change score, the 'Composite Primary Symptom Reduction (CPSR) score' (including measurements of pain, tenderness, diarrhoea, constipation). The scale ranged from -1 to +1, so for example -0.66 represented a 66% improvement from baseline; 0.04 represented a 4% worsening. There was a statistically significant difference in symptom score of: -0.57 (95%CI -0.73, -0.42), which is clinically significant.

Figure 4


```
Review:       IBS 31 August 2006 (Version workotherpysch)
Comparison:   17 CBT versus placebo
Outcome:      02 CPSR score

Study                    CBT              Placebo          WMD (fixed)         Weight      WMD (fixed)
or sub-category    N     Mean (SD)    N   Mean (SD)        95% CI              %           95% CI

Greene 1994        10   -0.66(0.16)   10  -0.02(0.56)                         17.88       -0.64 [-1.00, -0.28]
Payne 1995         12   -0.67(0.24)   10  -0.10(0.23)                         60.09       -0.57 [-0.77, -0.37]
VollmerB CBTgp vs pl  11 -0.52(0.40)   5   0.04(0.45)                         11.02       -0.56 [-1.02, -0.10]
VollmerC CBTin vs pl  11 -0.46(0.40)   5   0.04(0.45)                         11.02       -0.50 [-0.96, -0.04]

Total (95% CI)     44                 30                                      100.00      -0.57 [-0.73, -0.42]
Test for heterogeneity: Chi² = 0.23, df = 3 (P = 0.97), I² = 0%
Test for overall effect: Z = 7.37 (P < 0.00001)
                                                      -1   -0.5   0   0.5   1
                                                   Favours treatment  Favours control
```

c) Global symptom score (scale: high is good)

One large study included a global IBS score based on 'satisfaction, global wellbeing, diary pain scores, and health related quality of life' (Drossman 2003). For this study, CBT was significantly better than attention control, MD 0.09 (95%CI 0.04, 0.14), but was not significantly different from the no treatment placebo for desipramine; MD 0.04 (95%CI -0.01, 0.09). However, the range for the composite scale was unclear, so the clinical significance could not be assessed.

Figure 5


```
Review:       IBS 31 August 2006
Comparison:   17 CBT versus placebo
Outcome:      13 Global IBS score

Study                    CBT              Control          WMD (fixed)         Weight      WMD (fixed)
or sub-category    N     Mean (SD)    N   Mean (SD)        95% CI              %           95% CI

01 CBT versus attention control
Drossman A CBT vs at  135  0.49(0.23)  66  0.40(0.16)                         100.00      0.09 [0.04, 0.14]
Subtotal (95% CI)  135                66                                      100.00      0.09 [0.04, 0.14]
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.22 (P = 0.001)

02 CBT versus placebo
Drossman C CBT vs pl  135  0.49(0.23)  66  0.45(0.16)                         100.00      0.04 [-0.01, 0.09]
Subtotal (95% CI)  135                66                                      100.00      0.04 [-0.01, 0.09]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.43 (P = 0.15)
                                                      -1   -0.5   0   0.5   1
                                                   Favours Control  Favours CBT
```

d) Global change in symptom score

Heymann-Monnikes (2000) also reported the change in overall wellbeing compared to pre-treatment on a visual analogue scale ranging from: 1=very much better, to; 7=very much worse, with; 4 indicating no change. This was a statistically significant improvement for the CBT + medical treatment compared with medical treatment alone.

Figure 6


```
Review:      IBS 31 August 2006
Comparison:  24 CBT versus placebo/medical tmt
Outcome:     19 Global change in IBS score; scale high score = bad

Study                           CBT                Control              WMD (fixed)        Weight      WMD (fixed)
or sub-category          N      Mean (SD)     N    Mean (SD)            95% CI             %           95% CI

01 CBT + medical treatment vs medical treatment
  Heymann-Monnikes '00   12     1.92(0.67)    12   3.80(0.42)             ▬              100.00      -1.88 [-2.33, -1.43]
Subtotal (95% CI)        12                   12                          ◆              100.00      -1.88 [-2.33, -1.43]
Test for heterogeneity: not applicable
Test for overall effect: Z = 8.24 (P < 0.00001)

Total (95% CI)           12                   12                          ◆              100.00      -1.88 [-2.33, -1.43]
Test for heterogeneity: not applicable
Test for overall effect: Z = 8.24 (P < 0.00001)
                                                           -10    -5    0    5    10
                                                              Favours CBT  Favours control
```

e) Global symptom score (scale: high = severe)

Another study (Lynch 1989) described an IBS score in terms of 'pain, discomfort, diarrhoea, and constipation, each rated 0 (no symptoms) to 6 (very severe symptoms)'. The treated patients had a reduction of 2.16 points compared with 0.36 points in the waiting list group (no SDs reported, p<0.05).

A further study (Greene 1994) of CBT versus symptom monitoring reported an IBS score based on 7 symptoms daily for 2 weeks each scored 0 to 4 (severe). The Heymann-Monnikes (2000) comparison of CBT+ medical treatment vs medical treatment alone reported a global symptom score derived from 20 items scored over 14 days. Kennedy (2005) reported the global IBS symptom score for patients receiving CBT plus mebeverine versus mebeverine alone. On the scale used, a score of <75 represented normal; 75-174 mild symptoms; 175-299 moderate symptoms and 300-500 severe symptoms.

Meta-analysis of these studies using the standardised mean difference showed a statistically significantly lower (better) global symptom scores for CBT compared with placebo.

Figure 7


```
Review:      IBS 31 August 2006
Comparison:  24 CBT versus placebo/medical tmt
Outcome:     14 Global IBS score; scale high score = bad

Study                           CBT                 Control               SMD (fixed)        Weight      SMD (fixed)
or sub-category          N      Mean (SD)      N    Mean (SD)             95% CI             %           95% CI

03 CBT vs symptom monitoring
  Greene 1994            10     52.80(38.10)   10   76.10(61.10)            ▬              11.93      -0.44 [-1.33, 0.45]
Subtotal (95% CI)        10                    10                          ◆              11.93      -0.44 [-1.33, 0.45]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.97 (P = 0.33)

04 CBT + medical treatment vs medical treatment
  Heymann-Monnikes '00   12     0.40(0.54)     12   0.74(0.43)              ▬              13.81      -0.67 [-1.50, 0.15]
Subtotal (95% CI)        12                    12                          ◆              13.81      -0.67 [-1.50, 0.15]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.59 (P = 0.11)

05 CBT+mebeverine vs mebeverine alone
  Kennedy 2005           58     171.92(93.69)  71   242.84(115.86)          ▬              74.27      -0.66 [-1.02, -0.31]
Subtotal (95% CI)        58                    71                          ◆              74.27      -0.66 [-1.02, -0.31]
Test for heterogeneity: not applicable
Test for overall effect: Z = 3.64 (P = 0.0003)

Total (95% CI)           80                    93                          ◆             100.00      -0.64 [-0.94, -0.33]
Test for heterogeneity: Chi² = 0.22, df = 2 (P = 0.90), I² = 0%
Test for overall effect: Z = 4.06 (P < 0.0001)
                                                           -4    -2    0    2    4
                                                              Favours CBT  Favours Control
```

For the specific comparison of CBT plus mebeverine versus mebeverine alone, there was a statistically significant improvement; mean difference: -71 (95%CI -107, -35) on a scale of 0 to 500.

Follow-up in global symptom scores

Kennedy (2005) (CBT + mebeverine versus mebeverine alone) also reported this outcome at follow-up at 13; 26, and; 52 weeks. At 13 weeks there was a statistically significant difference in favour of CBT + mebeverine; MD: -82 (95%CI -123, -42); at 26 weeks there was a borderline significant difference between interventions; MD: -40 (95%CI -80, 0.4; p=0.05), and there was no significant effect at 52 weeks; MD: -26 (95%CI -66, 15). The data were extracted from a graph, but the latter two results do not agree with the 'estimated treatment effects' (-14 and 3 respectively) reported in table 10 of the paper. There was close agreement between the graph and the table for the effect size at the end of treatment and fairly close agreement at 13 weeks (table: -68 and -71 respectively). We have used the results from the graph because the table was said to be 'estimated'.

Figure 8

Review: IBS 31 August 2006
Comparison: 20 CBT plus medical therapy versus medical therapy alone
Outcome: 04 Global IBS score follow up

Study or sub-category	N	CBT + medical Mean (SD)	N	Medical therapy Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 6 weeks (end of intervention)							
Kennedy 2005	58	171.92(93.69)	71	242.84(115.86)		100.00	-70.92 [-107.08, -34.76]
Subtotal (95% CI)	58		71			100.00	-70.92 [-107.08, -34.76]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.84 (P = 0.0001)							
02 13 weeks							
Kennedy 2005	52	169.08(93.93)	49	251.35(111.44)		100.00	-82.27 [-122.59, -41.95]
Subtotal (95% CI)	52		49			100.00	-82.27 [-122.59, -41.95]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 4.00 (P < 0.0001)							
03 26 weeks							
Kennedy 2005	53	188.94(105.37)	58	228.65(110.22)		100.00	-39.71 [-79.83, 0.41]
Subtotal (95% CI)	53		58			100.00	-39.71 [-79.83, 0.41]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.94 (P = 0.05)							
04 52 weeks							
Kennedy 2005	52	200.29(93.93)	58	225.82(121.25)		100.00	-25.53 [-65.85, 14.79]
Subtotal (95% CI)	52		58			100.00	-25.53 [-65.85, 14.79]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.24 (P = 0.21)							

-1000   -500   0   500   1000
Favours treatment   Favours control

2. Individual symptoms

a) Pain

Seven studies (Blanchard 1993; Drossman 2003; Fernandez 1998; Greene 1994; Lynch 1989; Tkachuk 2003; Vollmer 1998) reported pain score outcomes. However, different pain scoring scales were used: Blanchard (1993): 0-4 scale daily added up over 4 weeks; Drossman (2003) used the McGill pain score (items scored 0-3, averaged, and added over 2 weeks); Fernandez (1998): 0 to 4 scale daily added up over 1 week; Greene (1994): 0 to 4 scale daily added up over 4 weeks; Lynch (1989) used a score ranging from 0 to 6; Tkachuk (2003) and Vollmer (1998) used a score ranging from 0 to 4 daily. In all cases, the maximum of the scale corresponded to severe pain.

The studies were combined in a meta-analysis (omitting Fernandez (1998) due to high drop-out rates) using standardised mean differences. There was no significant difference between CBT and control for pain score.

Figure 9

Review: IBS 31 August 2006
Comparison: 17 CBT versus placebo
Outcome: 12 Pain

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
Drossman A CBT vs at	67	9.52(6.30)	66	10.13(7.20)		38.90	-0.09 [-0.43, 0.25]
Drossman C CBT vs pl	68	9.52(6.30)	66	10.73(6.70)		39.05	-0.19 [-0.52, 0.15]
Greene 1994	10	20.50(17.20)	10	24.90(22.60)		5.82	-0.21 [-1.09, 0.67]
Lynch 1989	11	0.57(0.00)	10	0.15(0.00)			Not estimable
Tkachuk 2003	14	1.07(0.86)	14	1.08(0.91)		8.20	-0.01 [-0.75, 0.73]
VollmerB CBTgp vs pl	11	0.56(0.58)	5	0.61(0.43)		4.02	-0.09 [-1.15, 0.97]
VollmerC CBTin vs pl	11	0.74(0.81)	5	0.61(0.43)		4.01	0.17 [-0.89, 1.23]
Blanchard 1993	8	13.50(0.00)	8	3.50(0.00)			Not estimable
Total (95% CI)	200		184			100.00	-0.12 [-0.33, 0.10]

Test for heterogeneity: Chi² = 0.59, df = 5 (P = 0.99), I² = 0%
Test for overall effect: Z = 1.08 (P = 0.28)

-4 -2 0 2 4
Favours CBT   Favours control

b) Bloating

Bloating was reported by Greene (1994) on a 0-4 daily scale added up over 4 weeks, i.e. maximum 112; Lynch (1989) on a 0-6 scale; Tkachuk (2003) on a 0-4 scale daily, and; Vollmer (1998) (group and individual CBT; on a 0-4 scale daily). In each case, the maximum of the scale corresponded to severe bloating. The studies were combined using standardised mean differences.

Figure 10

Review: IBS 31 August 2006
Comparison: 24 CBT versus placebo/medical tmt
Outcome: 03 Bloating

Study or sub-category	N	CBT Mean (SD)	N	Placebo Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
Greene 1994	10	8.90(8.90)	10	11.40(10.00)		27.10	-0.25 [-1.13, 0.63]
Lynch 1989	11	0.40(0.00)	10	-0.13(0.00)			Not estimable
Tkachuk 2003	14	1.32(0.77)	14	1.28(0.89)		38.30	0.05 [-0.69, 0.79]
VollmerB CBTgp vs pl	11	0.69(0.73)	5	0.69(0.53)		18.81	0.00 [-1.06, 1.06]
VollmerC CBTin vs pl	11	0.25(0.26)	5	0.69(0.53)		15.79	-1.16 [-2.31, -0.01]
Total (95% CI)	57		44			100.00	-0.23 [-0.69, 0.22]

Test for heterogeneity: Chi² = 3.22, df = 3 (P = 0.36), I² = 6.7%
Test for overall effect: Z = 1.00 (P = 0.32)

-4 -2 0 2 4
Favours treatment   Favours control

There was no significant difference between interventions in the bloating score.

c) Bowel habits

Ratings of constipation and diarrhoea were reported by Greene (1994) on a 0-4 scale daily added up over 4 weeks, i.e. maximum 112, and; Lynch (1989) on a 0-6 scale. In each case, the maximum of the scale corresponded to severe symptoms.

Both studies reported mean scores at baseline and after the intervention period. However, Lynch (1989) did not report standard deviations and was not analysed further.

In Greene (1994), baseline constipation scores were 10.3 (SD 7.7) in the CBT group compared with 8.8 (SD 14.2) in the placebo group. Baseline diarrhoea scores were 13.9 (SD 7.6) compared with 10.7 (SD 10.4). These are not significant differences.

Final scores at eight weeks showed a fairly small, statistically significant difference in the diarrhoea score (Figure 11). There was no significant difference in the constipation score.

Figure 11


```
Review:        IBS 31 August 2006
Comparison:    24 CBT versus placebo/medical tmt
Outcome:       18 Bowel scores

Study                                      CBT              Placebo          WMD (fixed)      Weight      WMD (fixed)
or sub-category            N          Mean (SD)      N      Mean (SD)          95% CI           %          95% CI

01 Diarrhoea scale 0-4 daily scale added over 4 weeks (severe)
  Greene 1994              10          2.00(2.80)    10     7.70(8.40)                        100.00      -5.70 [-11.19, -0.21]
Subtotal (95% CI)         10                         10                                      100.00      -5.70 [-11.19, -0.21]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.04 (P = 0.04)

02 Constipation 0-4 daily scale added over 4 weeks (severe)
  Greene 1994              10          4.90(4.80)    10     7.80(9.00)                        100.00      -2.90 [-9.22, 3.42]
Subtotal (95% CI)         10                         10                                      100.00      -2.90 [-9.22, 3.42]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.90 (P = 0.37)

03 Diarrhoea 0-4 daily scale added over 4 weeks (severe) baseline
  Greene 1994              10          13.90(7.60)   10     10.70(10.40)                      100.00       3.20 [-4.78, 11.18]
Subtotal (95% CI)         10                         10                                      100.00       3.20 [-4.78, 11.18]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

04 Constipation 0-4 daily scale added over 4 weeks (severe) (baseline)
  Greene 1994              10          10.30(7.70)   10     8.80(14.20)                       100.00       1.50 [-8.51, 11.51]
Subtotal (95% CI)         10                         10                                      100.00       1.50 [-8.51, 11.51]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.29 (P = 0.77)

                    -100   -50    0    50   100
                   Favours treatment  Favours control
```

3. Adverse events

These were not reported in any of the studies.

4. Quality of life

The IBS-QOL scale (maximum 84; high is good) was used to report outcomes in Drossman (2003), whilst a scale of 'disruption of daily life' was used in Lynch (1989), although no standard deviations were reported.

In Drossman (2003), there was a small statistically significant difference between CBT and no treatment on the IBS QoL scale. There was no significant difference between CBT and attention control.

Figure 12


```
Review:        IBS 31 August 2006
Comparison:    24 CBT versus placebo/medical tmt
Outcome:       21 IBS QoL

Study                                      CBT              control          WMD (fixed)      Weight      WMD (fixed)
or sub-category            N          Mean (SD)      N      Mean (SD)          95% CI           %          95% CI

02 IBS QoL (CBT vs attention control)
  Drossman A CBT vs at    144         76.55(13.50)   71    73.60(14.00)                       100.00       2.95 [-0.98, 6.88]
Subtotal (95% CI)        144                         71                                      100.00       2.95 [-0.98, 6.88]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.47 (P = 0.14)

03 IBS QoL (CBT vs no treatment)
  Drossman C CBT vs pl    144         76.55(13.50)   72    72.02(16.30)                       100.00       4.53 [0.17, 8.89]
Subtotal (95% CI)        144                         72                                      100.00       4.53 [0.17, 8.89]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.03 (P = 0.04)

                    -100   -50    0    50   100
                   Favours control   Favours CBT
```

Kennedy (2005) (CBT+mebeverine versus mebeverine alone) also reported on a global 'Work and social adjustment score' (handicap in work, home, leisure and relationships; 0=not affected; 8=severely affected for each area; maximum total 40). This outcome was also

reported at 26 and 52 weeks follow-up. There was a statistically significant improvement in this score that was maintained over the 52 weeks of follow-up. The difference was fairly small though: MD at 6 weeks -4.70 (95%CI -7.43, -1.97) on a scale of 0 to 40.

Figure 13

Review: IBS 31 August 2006
Comparison: 20 CBT plus medical therapy versus medical therapy alone
Outcome: 05 Work and Social Adjustment Score follow up

Study or sub-category	N	CBT+medical Mean (SD)	N	Medical therapy Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 6 weeks (end of intervention)							
Kennedy 2005	72	9.10(8.20)	77	13.80(8.80)		100.00	-4.70 [-7.43, -1.97]
Subtotal (95% CI)	72		77			100.00	-4.70 [-7.43, -1.97]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.37 (P = 0.0007)							
02 26 weeks							
Kennedy 2005	54	9.30(7.90)	58	12.50(9.30)		100.00	-3.20 [-6.39, -0.01]
Subtotal (95% CI)	54		58			100.00	-3.20 [-6.39, -0.01]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.97 (P = 0.05)							
03 52 weeks							
Kennedy 2005	51	9.30(8.30)	58	13.10(9.70)		100.00	-3.80 [-7.18, -0.42]
Subtotal (95% CI)	51		58			100.00	-3.80 [-7.18, -0.42]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.20 (P = 0.03)							

-100 -50 0 50 100
Favours treatment   Favours control

Heymann-Monnikes (2000) (CBT + medical treatment versus medical treatment) reported scores on the GI quality of life instrument (scale maximum 144; high is good). There was a statistically and clinically significant improvement in quality of life for the CBT group.

Figure 14

Review: IBS 31 August 2006
Comparison: 20 CBT plus medical therapy versus medical therapy alone
Outcome: 09 Quality of life

Study or sub-category	N	CBT+medical Mean (SD)	N	Medical therapy Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Heymann-Monnikes '00	12	104.83(17.60)	12	83.10(13.90)		100.00	21.73 [9.04, 34.42]
Total (95% CI)	12		12			100.00	21.73 [9.04, 34.42]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 3.36 (P = 0.0008)							

-100 -50 0 50 100
Favours control   Favours CBT

5. Number of patients withdrawing from study

Blanchard (1993) and Fernandez (1998) reported withdrawals from the study.

Figure 15

Review: IBS 31 August 2006
Comparison: 17 CBT versus placebo
Outcome: 04 Withdrawals

Study or sub-category	CBT n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Fernandez 1998	22/44	16/23		94.52	0.72 [0.48, 1.07]
Blanchard 1993	6/14	1/9		5.48	3.86 [0.55, 26.96]
Total (95% CI)	58	32		100.00	0.89 [0.59, 1.34]
Total events: 28 (CBT), 17 (Placebo)					
Test for heterogeneity: Chi² = 3.28, df = 1 (P = 0.07), I² = 69.5%					
Test for overall effect: Z = 0.56 (P = 0.58)					

0.01 0.1 1 10 100
Favours treatment   Favours control

6. Mental health outcomes

Five studies (Greene 1994; Heymann-Monnikes 2000; Lynch 1989; Payne 1995; Tkachuk 2003) reported Beck depression inventory scores.

a) Beck depression inventory score

There was a statistically significant improvement in Beck Depression scores (scale maximum 63; high=bad), favouring CBT; WMD -4.68 (95%CI -6.79, -2.57), but the difference was fairly small. There was no heterogeneity between studies (I^2=0%, p=0.82).

Figure 16
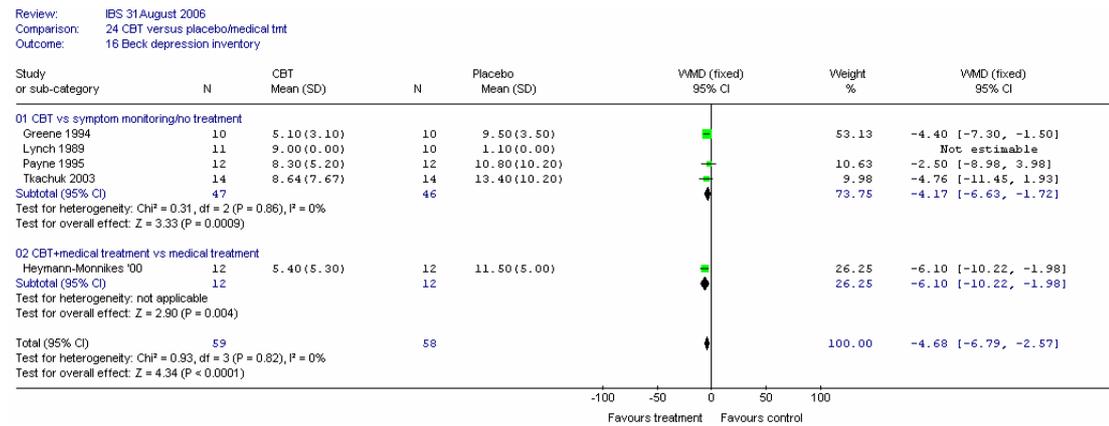
b) Overall anxiety and psychological distress: State-Trait Anxiety Inventory (STAI).

Four studies (Greene 1994; Heymann-Monnikes 2000; Payne 1995; Tkachuk 2003) reported anxiety using the State-Trait Anxiety Inventory (STAI) (scale range 20-80; high = bad). There was no significant difference between CBT and control.

Figure 17
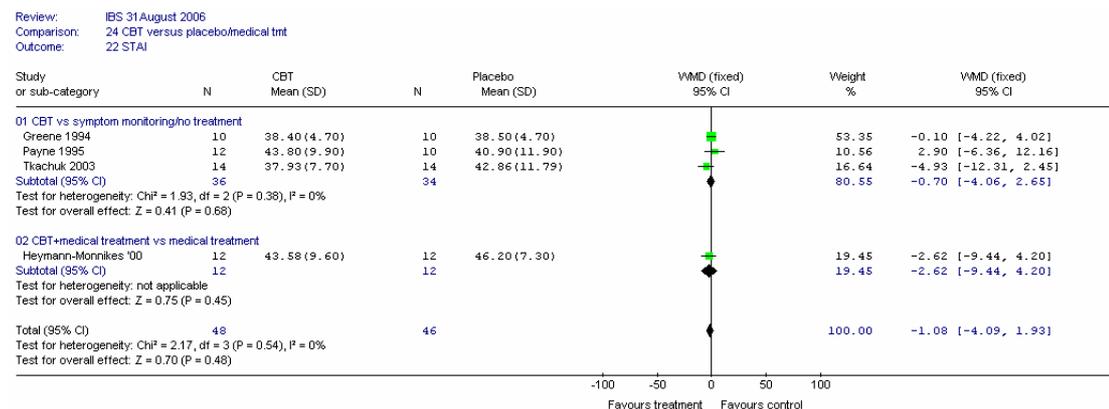
c) Psychological distress (HADS)

Kennedy (2005) (CBT plus mebeverine versus mebeverine) reported the HADS score (range 0 to 56; high=bad) at the end of treatment (6 weeks) and at follow-up at 26 and 52 weeks.

There was a small statistically significant difference favouring CBT which was maintained over 52 weeks; MD at 6 weeks: -2.80 (95%CI -5.31, -0.29).

Figure 18
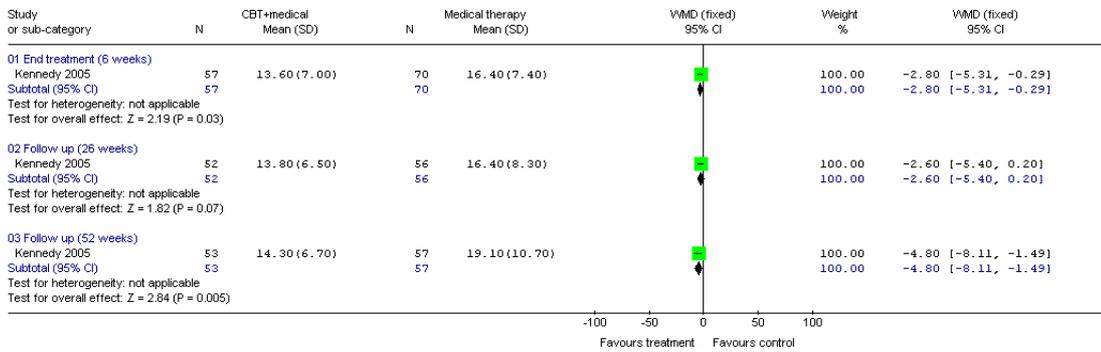
d) SF-36 mental health composite

One study (Tkachuk 2003) reported the SF-36 mental health composite score (maximum 100, high = bad). There was no significant difference between interventions.

Figure 19
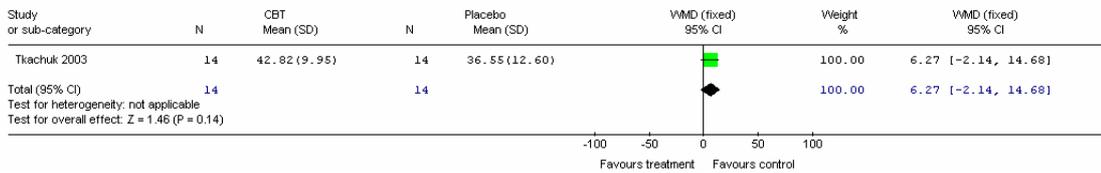
e) Dysfunctional cognitions score

One study (Tkachuk 2003) reported dysfunctional cognitions (scale 25-175; high=bad). This study showed a statistically significant improvement for the CBT patients compared with waiting list control with symptom monitoring. We noted that this study had the majority of patients with an Axis I diagnosis.

Figure 20
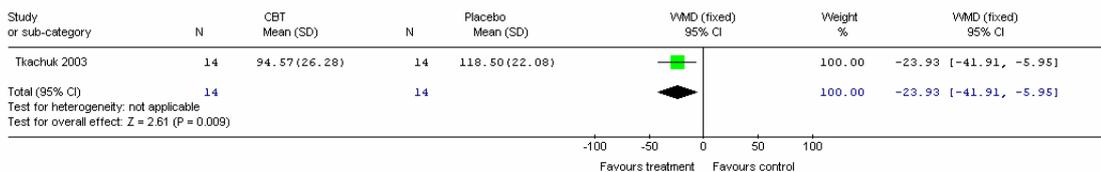
B. CBT type 1 versus type 2 (e.g. stress management versus contingency management)

Boyce (2003) compared CBT with relaxation. Fernandez (1998) compared stress management with contingency management. Payne (1995) compared CBT with self-help groups.

B1. CBT versus relaxation

1. Global symptoms

Boyce (2003) reported Global IBS symptom score (symptom severity on a 0 to 48 scale) at the end of treatment (8 weeks) and at follow-up at 26 and 52 weeks. There was no significant difference between interventions at any time. We noted that there were large numbers of drop-outs, especially in the relaxation arm.

Figure 21
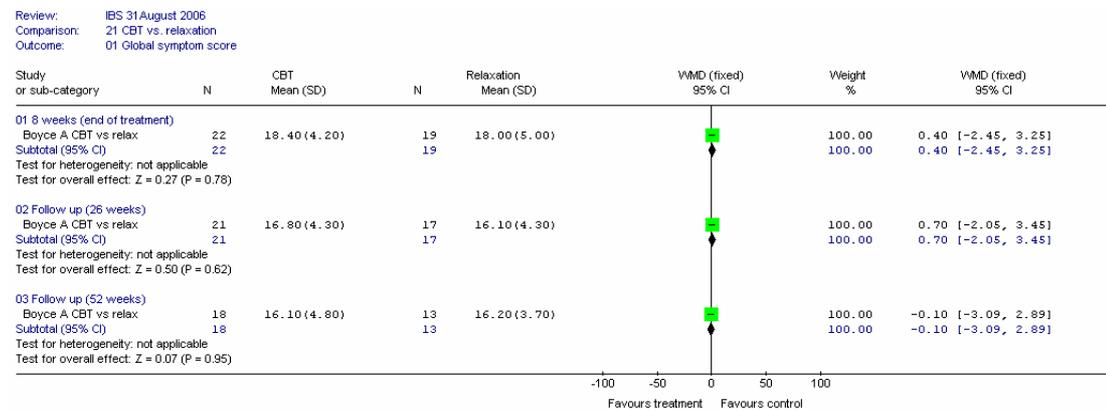
2. Mental health outcomes

Boyce (2003) reported the HADS score (0-56; high=bad) at the end of treatment (8 weeks) and at follow-up at 26 and 52 weeks. There was no significant difference between interventions.

Figure 22
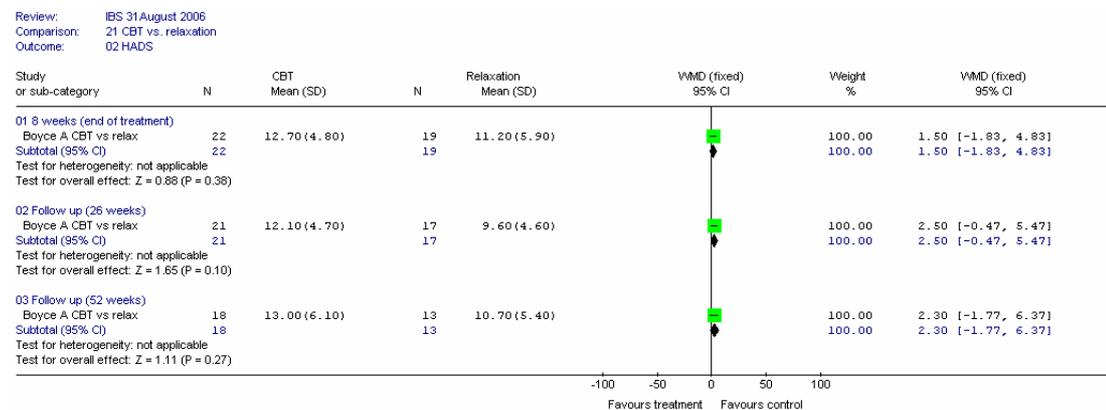
B2. Stress management versus contingency management

1. Global symptoms

Fernandez (1998) found that 8/15 of the stress management group and 14/16 of the contingency management group improved at the end of treatment (10 weeks). This study also reported global improvement in terms of the number of patients improved at 1 year follow-up: 8 of the 13 remaining patients in the stress management group versus 7 of 11 patients in the contingency management group. At both durations the confidence intervals were fairly wide and there was no statistically significant difference between interventions, however, at the end of treatment stress management was favoured.

Figure 23
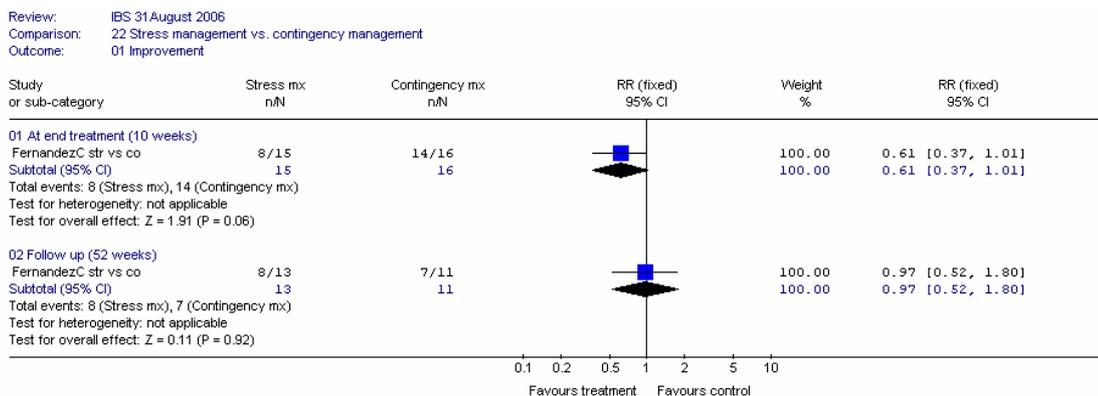
2. Individual symptoms

a) Pain

Fernandez (1998) reported mean pain scores on a scale of 0 to 4 (scores totalled for 1 week). There was no significant difference between interventions. No other individual symptoms were reported.

Figure 24
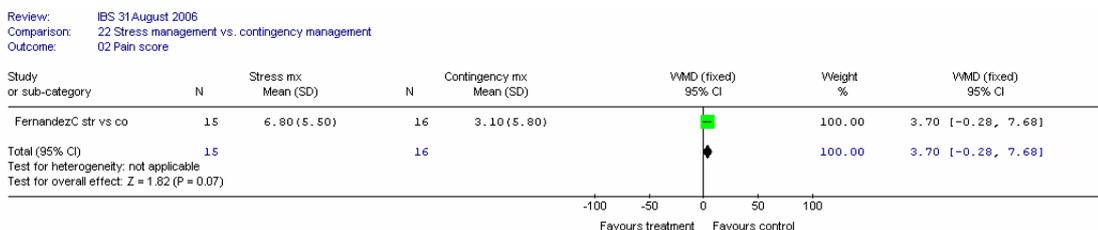
3. Number of patients withdrawing from study

Fernandez (1998) reported 6/21 patients withdrawing from the stress management group compared with 7/23 for contingency management.

B3. CBT versus self help groups

1. Global symptoms

At the end of treatment (8 weeks), Payne (1995) reported that 9/12 CBT patients and 3/12 self-help group patients were improved after treatment. At follow-up (12 weeks) 10/12 CBT patients and 2/11 self-help group patients were improved. At both times there was a statistically significantly better result for the CBT group, but confidence intervals were wide.

Figure 25
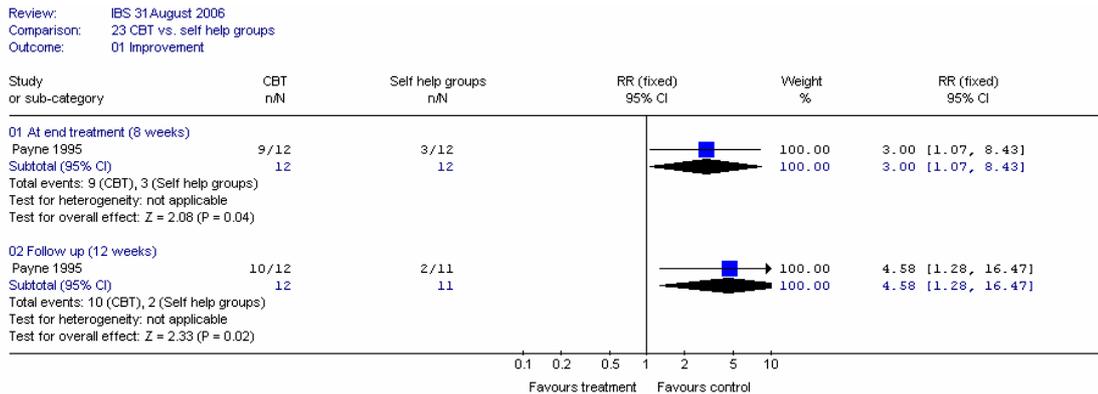
2. Mental health outcomes

a) Beck Depression Inventory

Payne (1995) reported Beck Depression Inventory scores (scale maximum 63; high=bad) at the end of treatment (8 weeks) and at follow-up at 12 weeks. There was no significant difference at either time.

Figure 26
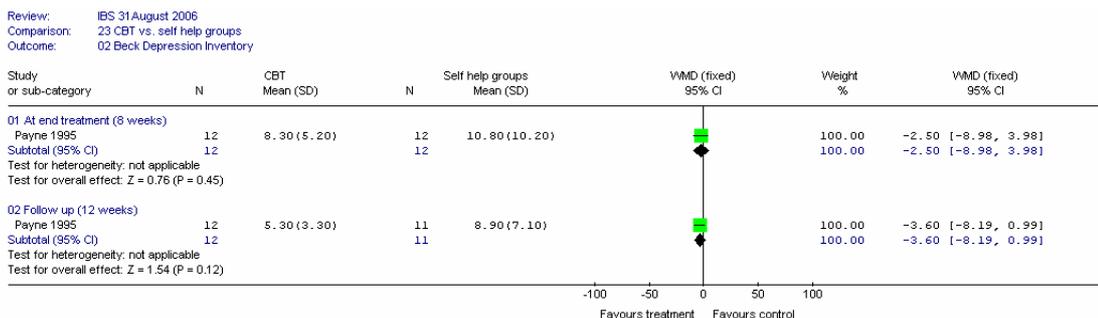
b) Overall anxiety and psychological distress: State-Trait Anxiety Inventory (STAI)

Payne (1995) reported STAI scores (20 to 80; high = bad) at the end of treatment (8 weeks) and at follow-up at 12 weeks. There was no significant difference between interventions.

Figure 27

```
Review:      IBS 31 August 2006
Comparison:  23 CBT vs. self help groups
Outcome:     03 STAI

Study                                  CBT              Self help groups      WMD (fixed)        Weight        WMD (fixed)
or sub-category           N         Mean (SD)      N       Mean (SD)           95% CI              %            95% CI

01 At end treatment (8 weeks)
  Payne 1995              12       48.30(9.90)     12      41.60(12.90)                          100.00      6.70 [-2.50, 15.90]
Subtotal (95% CI)         12                       12                                            100.00      6.70 [-2.50, 15.90]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.43 (P = 0.15)

02 Follow up (12 weeks)
  Payne 1995              12       38.00(7.70)     12      41.60(12.90)                          100.00     -3.60 [-12.10, 4.90]
Subtotal (95% CI)         12                       12                                            100.00     -3.60 [-12.10, 4.90]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.83 (P = 0.41)

                                                          -100   -50    0    50   100
                                                          Favours treatment   Favours control
```

C. CBT individual versus group

One study (Vollmer 1998) compared CBT on an individual basis with CBT on a group basis. Outcomes reported were global improvement in symptoms (mean score); global improvement in symptoms (number of patients); pain score, and; bloating score (0=not a problem; 4=debilitating symptoms).

1. Global outcomes

a) Global improvement in symptoms score

Vollmer (1998) reported the mean global improvement in symptoms score (CPSR; scale -1 to +1) for CBT group patients compared with individual CBT at the end of treatment. There was no significant difference between intervention, but the confidence interval was fairly wide, leading to uncertainty. No standard deviations were given for the follow-up scores.

Figure 28

```
Review:      IBS 31 August 2006
Comparison:  19 CBT group versus CBT individual
Outcome:     01 Global improvement of symptoms (mean score)

Study                               CBT group          CBT individual        WMD (fixed)        Weight        WMD (fixed)
or sub-category           N        Mean (SD)       N      Mean (SD)            95% CI              %            95% CI

  VollmerA CBTgp vs in    11       0.52(0.40)      11     0.46(0.40)                             100.00      0.06 [-0.27, 0.39]

Total (95% CI)            11                       11                                            100.00      0.06 [-0.27, 0.39]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.35 (P = 0.73)

                                                          -1    -0.5    0    0.5    1
                                                          Favours treatment   Favours control
```

b) Global improvement in symptoms (number of patients)

Vollmer (1998) reported 7/11 patients improved with group CBT compared with 6/11 for individual CBT. There was no significant difference between intervention, but the confidence interval was fairly wide, leading to uncertainty.

Figure 29

Review: IBS 31 August 2006
Comparison: 19 CBT group versus CBT individual
Outcome: 02 Global improvement in symptoms (number of patients)

Study or sub-category	CBT group n/N	CBT individual n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
VollmerA CBTgp vs in	7/11	6/11		100.00	1.17 [0.58, 2.35]
Total (95% CI)	11	11		100.00	1.17 [0.58, 2.35]

Total events: 7 (CBT group), 6 (CBT individual)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.43 (P = 0.67)

0.1 0.2 0.5 1 2 5 10
Favours treatment   Favours control

2. Individual symptoms

a) Pain

The mean pain score (scale 0 to 4) showed no significant difference between interventions.

Figure 30

Review: IBS 31 August 2006
Comparison: 19 CBT group versus CBT individual
Outcome: 03 Pain score

Study or sub-category	N	CBT group Mean (SD)	N	CBT individual Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
VollmerA CBTgp vs in	11	0.74(0.81)	11	0.56(0.58)		100.00	0.18 [-0.41, 0.77]
Total (95% CI)	11		11			100.00	0.18 [-0.41, 0.77]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.60 (P = 0.55)

-4 -2 0 2 4
Favours treatment   Favours control
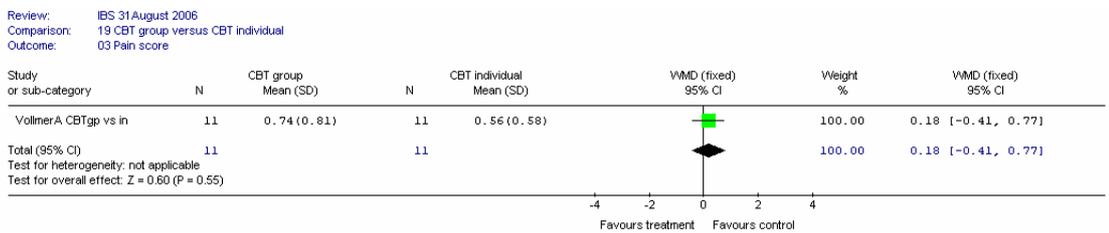
b) Bloating

The mean bloating score showed no significant difference between CBT group therapy and individual therapy, although individual therapy was favoured.

Figure 31

Review: IBS 31 August 2006
Comparison: 19 CBT group versus CBT individual
Outcome: 04 Bloating score

Study or sub-category	N	CBT group Mean (SD)	N	CBT individual Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
VollmerA CBTgp vs in	11	0.69(0.73)	11	0.25(0.26)		100.00	0.44 [-0.02, 0.90]
Total (95% CI)	11		11			100.00	0.44 [-0.02, 0.90]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.88 (P = 0.06)

-4 -2 0 2 4
Favours treatment   Favours control

D. CBT versus medical therapy

Boyce (2003) compared CBT or relaxation with routine medical care. Corney (1991) compared behavioural psychotherapy (nurse behaviour therapist 6-15 one-hour sessions) with 'conventional medical treatment'. Fernandez (1998) compared stress management or contingency management with 'conventional medical treatment'.

1. Global outcomes

a) Number of patients with improvement in global symptoms

One study (Fernandez 1998) reported the number of patients improved. Meta-analysis showed a statistically significant improvement for CBT compared with medical treatment, but

the confidence intervals were wide. At 52 week follow-up there were too many withdrawals for this to be reliable.

Figure 32

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 01 Number of patients improved
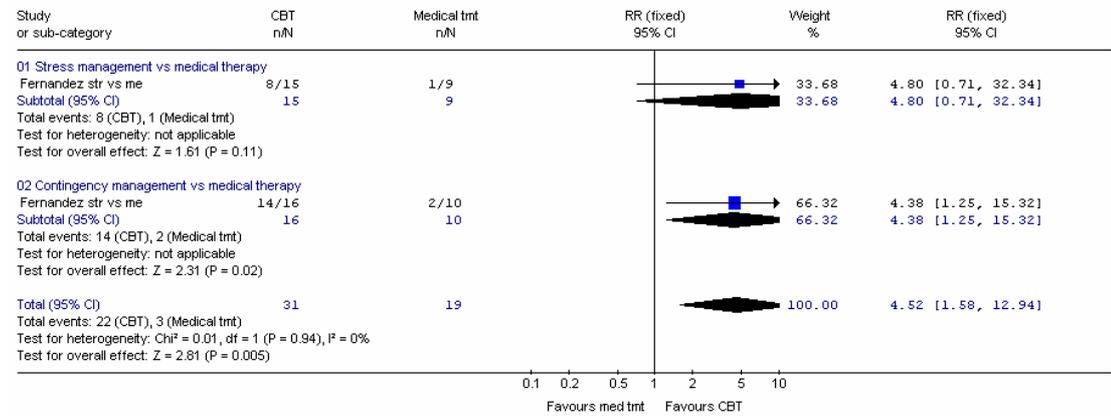
b) **Global symptom score**

Boyce (2003) (CBT or relaxation versus medical care) reported Global IBS symptom score (symptom severity on a 0 to 48 scale) in 66 patients at the end of treatment (8 weeks) and at follow-up at 26 and 52 weeks. At 8 weeks, there was no significant difference between interventions, although we noted that the drop-out rates were fairly high, especially for the relaxation arm of the study.

Figure 33

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 02 Global IBS score; scale high score = bad
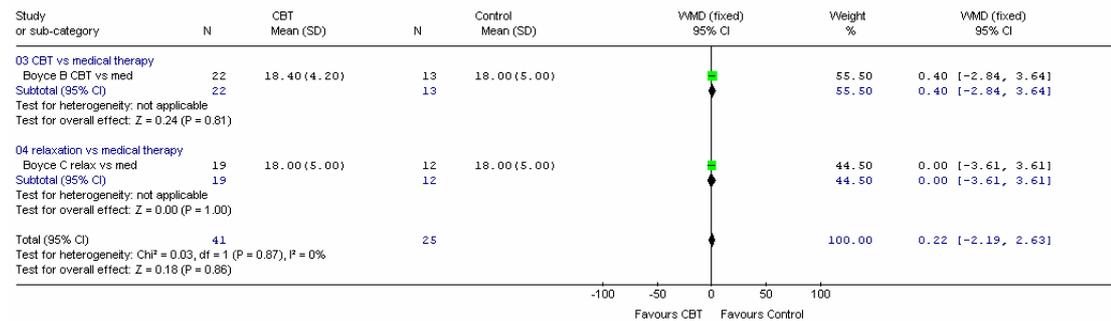
2. **Individual symptoms**

a) **Pain**

Corney (1991), in 42 patients, reported pain on a visual analogue scale (VAS) (the scale was unclear, but may be 0 to 8 scale in 3 dimensions) at the end of treatment (16 weeks) and at follow-up (40 weeks). There was no significant difference between interventions at either time, but the confidence intervals were fairly wide.

Figure 34

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 04 Pain score; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Pain scores at end of treatment							
Corney 1991	22	8.53(6.48)	20	8.80(7.52)		100.00	-0.27 [-4.54, 4.00]
Subtotal (95% CI)	22		20			100.00	-0.27 [-4.54, 4.00]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.12 (P = 0.90)							
02 Pain scores at 40 weeks follow up							
Corney 1991	21	6.33(6.38)	20	9.20(7.56)		100.00	-2.87 [-7.16, 1.42]
Subtotal (95% CI)	21		20			100.00	-2.87 [-7.16, 1.42]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.31 (P = 0.19)							

-10 -5 0 5 10
Favours CBT   Favours Control

Fernandez (1998), in 50 patients, reported a pain score rated 0 (not a problem) to 4 (debilitating). There were significant differences between both active treatment groups and the medical treatment control group (p<0.001 for the contingency management group and p<0.022 for the stress management group). We noted that around 30% of each CBT group was missing data and 26% of the control group.

We combined these studies in a meta-analysis using standardised mean differences and found significant heterogeneity (I^2=79%, p=0.009). The source of heterogeneity was not clear, although one difference is that the CBT intervention in Corney (1991) was behavioural psychotherapy.

Figure 35

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 03 Pain score; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
01 Pain scores							
Corney 1991	22	8.53(6.48)	20	8.80(7.52)		53.18	-0.04 [-0.64, 0.57]
Fernandez co vs me	16	3.10(5.80)	10	15.80(9.90)		22.89	-1.62 [-2.54, -0.69]
Fernandez str vs me	15	6.80(5.50)	9	15.80(9.90)		23.94	-1.17 [-2.08, -0.27]
Subtotal (95% CI)	53		39			100.00	-0.67 [-1.11, -0.23]
Test for heterogeneity: Chi² = 9.43, df = 2 (P = 0.009), I² = 78.8%							
Test for overall effect: Z = 2.98 (P = 0.003)							
Total (95% CI)	53		39			100.00	-0.67 [-1.11, -0.23]
Test for heterogeneity: Chi² = 9.43, df = 2 (P = 0.009), I² = 78.8%							
Test for overall effect: Z = 2.98 (P = 0.003)							

-10 -5 0 5 10
Favours CBT   Favours Control

b) Bowel habits

Corney (1991) reported constipation and diarrhoea on a visual analogue scale (VAS) at end of treatment (16 weeks) and at follow-up (40 weeks). The scale was unclear. There was no significant difference between interventions, although the confidence interval may have been wide.

Figure 36

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 05 Constipation score; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Constipation scores at end of treatment							
Corney 1991	22	5.14(6.39)	20	4.70(6.21)		100.00	0.44 [-3.37, 4.25]
Subtotal (95% CI)	22		20			100.00	0.44 [-3.37, 4.25]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.23 (P = 0.82)							
02 Constipation scores at 40 weeks follow up							
Corney 1991	21	3.76(6.32)	20	4.90(6.90)		100.00	-1.14 [-5.20, 2.92]
Subtotal (95% CI)	21		20			100.00	-1.14 [-5.20, 2.92]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.55 (P = 0.58)							

-10   -5   0   5   10
Favours CBT   Favours Control

For diarrhoea scores there was no significant difference, but the confidence interval was probably wide, depending on the scale.

Figure 37

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 06 Diarrhoea score; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Diarrhoea scores at end of treatment							
Corney 1991	22	4.00(6.16)	20	7.45(13.44)		100.00	-3.45 [-9.88, 2.98]
Subtotal (95% CI)	22		20			100.00	-3.45 [-9.88, 2.98]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.05 (P = 0.29)							
02 Diarrhoea scores at 40 weeks follow up							
Corney 1991	21	3.14(4.76)	20	4.10(6.29)		100.00	-0.96 [-4.39, 2.47]
Subtotal (95% CI)	21		20			100.00	-0.96 [-4.39, 2.47]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.55 (P = 0.58)							

-10   -5   0   5   10
Favours CBT   Favours Control

c) Number of patients withdrawing from study

Fernandez (1998) reported 6/21 withdrawals from the stress management group, 7/23 from the contingency management group and 4/23 from the medical treatment group.

d) Mental health outcomes

Boyce (2003) (CBT versus medical care) reported the psychological distress on the HADS scale (0-56, high is bad) at the end of treatment (8 weeks) and at follow-up at 26 and 52 weeks. At 8 weeks, there was no significant difference between interventions.

Figure 38

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 07 HADS score; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 CBT vs medical care at end of treatment							
Boyce B CBT vs med	22	12.70(4.80)	25	11.00(6.50)		100.00	1.70 [-1.54, 4.94]
Subtotal (95% CI)	22		25			100.00	1.70 [-1.54, 4.94]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.03 (P = 0.30)							
02 relaxation vs medical care at end of treatment							
Boyce C relax vs med	19	11.20(5.90)	25	11.00(6.50)		100.00	0.20 [-3.48, 3.88]
Subtotal (95% CI)	19		25			100.00	0.20 [-3.48, 3.88]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.11 (P = 0.92)							

-100   -50   0   50   100
Favours CBT   Favours Control

Corney (1991) reported psychological distress at 9 months follow up using the Clinical Interview Schedule (0-48, high=bad). There was no significant difference between interventions.

Figure 39

Review: IBS 31 August 2006
Comparison: 18 CBT versus medical therapy
Outcome: 08 Psychological distress Clinical interview schedule; scale high score = bad

Study or sub-category	N	CBT Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 BT vs medical care at 9m follow up							
Corney 1991	21	12.60(11.50)	20	18.10(16.50)		100.00	-5.50 [-14.25, 3.25]
Subtotal (95% CI)	21		20			100.00	-5.50 [-14.25, 3.25]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.23 (P = 0.22)							

ECONOMIC LITERATURE FOR CBT

One relevant health economic analysis was identified on the cost-effectiveness of CBT in the management of IBS. Kennedy (2006) was a trial based economic evaluation conducted in the UK, which recruited patients from primary care with IBS symptoms of moderate or greater severity following 2 weeks of GP care and 4 weeks of mebeverine. Patients were randomised to receive either CBT plus mebeverine or mebeverine alone. CBT consisted of six 50-minute sessions delivered by face-to-face contact with a trained nurse. The primary effectiveness measure was global IBS symptom score. Direct and indirect costs were measured and cost-effectiveness acceptability curves were used to describe the probability that CBT plus mebeverine would be more cost-effective than mebeverine alone for various willingness to pay thresholds for a 10 unit change in global IBS symptom score.

The addition of CBT produced significantly better global IBS symptom scores at 3 months but the effect had diminished and was no longer significant after 6 and 12 months of follow-up. Similar results were observed in the secondary effectiveness measures and further details of the clinical outcomes have been given in the clinical effectiveness review. The mean intervention cost of CBT was £308 per patient. The addition of CBT did not reduce service costs at 3, 6 or 12 months. It is unclear whether service costs included intervention costs but given that the mean service costs at 3 months were less than £308 in both arms it is likely that they excluded the intervention cost. The CEACs presented show that CBT has a low probability (<25%) of being cost-effective when willingness to pay thresholds for a 10 unit change in global IBS symptom score are below £100. As this study did not provide an estimate of the cost per QALY for the addition of CBT to antispasmodic therapy, it was not particularly useful in determining whether recommending CBT for use in the NHS would result in the efficient use of NHS resources.

COST-EFFECTIVENESS ANALYSIS FOR CBT

This section describes the health economic analysis undertaken to inform recommendations on the use of CBT as a one-off intervention in the management of IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5

and the model inputs and assumptions relevant to this particular intervention are described below.

- The effectiveness of CBT in addition to usual care compared to usual care alone in people with refractory IBS was based on the number of patients with an improvement in global symptoms (at the end of treatment) for CBT vs no treatment, symptom monitoring or attention control. (RR 6.11 (95%CI 2.33, 16.07) for baseline rate of 9%, from Figure 3 of CBT review).
- We assumed that the response rate for CBT fell by 55% over the first 6 months and that there was no significant difference between CBT and usual care by the end of 12 months. This assumption was based on follow-up data for global symptom score from the study by Kennedy (2005). The mean difference, after adjustment for baseline difference, was used to scale the response rate after the end of treatment.
- The evidence included in the clinical effectiveness review did not allow a subtype specific estimate of clinical effectiveness to be estimated. Therefore it was assumed that CBT is equally effective in all IBS subtypes.
- CBT was assumed to be given over 12 weeks with alternative durations of 6 and 8 weeks considered in a sensitivity analyses, in which the costs are assumed to remain constant, but the effect is achieved over a shorter intervention period (i.e. the same number of sessions given at the same cost over a shorter time-frame).
- A 15 month time-frame was used so that the cost-effectiveness could be compared to against other psychological interventions for which there was 15 month efficacy data.

Modelled response rates

In the basecase scenario the response rate of 25% in the no treatment arm was taken from the mean placebo arm response rate from the behavioural therapy trials. This represents the group of patients whose symptoms improve without any specific intervention. The RR for an improvement in global symptoms for CBT vs no treatment at the end of treatment is 6.11; therefore the modelled response rate in the intervention arm is 100% at the end of treatment (12 weeks). As shown in Figure 40, the response rate in the CBT arm decreases to 59% by 6 months and 25% by one year, based on the assumptions regarding fall-off in effectiveness described above.

We have also considered a maintained benefits scenario in which the response to CBT was assumed to be maintained for the one year after the end of treatment but we assumed no further benefit beyond that point.

The cost of CBT was calculated using the mean number and duration of sessions from the studies used to calculate the RR, giving a mean duration of 6.6 hours of CBT (excluding Payne (1995) which did not provide sufficient information). To this we applied the unit cost for face-to-

face time with a Cognitive Behavioural Therapist of £57 per hour (PSSRU 2006), which gave a total mean cost of £375 (95%CI £167 - £582).

The study by Kennedy (2005) found no significant difference in direct health care costs at 3, 6 or 12 months for CBT plus mebeverine compared to mebeverine alone in patients with severe symptoms after 3 months of mebeverine. It is likely that significant reductions in resource use would only be observed in patients who are high service users at baseline and who then experience a reduction in symptoms as a result of therapy. It is likely that the population included in the Kennedy (2005) study were not high service users since they were recruited from primary care after a failure to respond to only one intervention. This is in contrast with the Creed (2003) study which recruited patients from secondary and tertiary gastroenterology clinics who had a median of 6 visits to the doctor in the past 6 months, median symptom duration of 8 years and a median of 30 days with pain in the past 30 days. If an indirect comparison is made between CBT and psychotherapy, the odds ratios at the end of treatment suggest that CBT is at least as effective as the psychotherapy delivered in the Creed (2003) study (CBT compared to usual care OR=13.54, 95%CI 4.12-44.48, and psychotherapy vs usual care OR=2.44, 95%CI 1.28 – 4.67). It is therefore possible that similar reductions in resource use would be observed for CBT if the population were restricted to high service users. However, as there is no direct evidence for this we have excluded any reduced resource use for CBT in the basecase analysis. It was included in a sensitivity analysis by applying the reduction in resource use observed during the follow-up period of the Creed (2003) study for psychotherapy compared to usual care (£-4.08 per week, 95%CI-£8.11 to -£0.04) indirectly to CBT.

Figure 40: Response rate in the basecase analysis
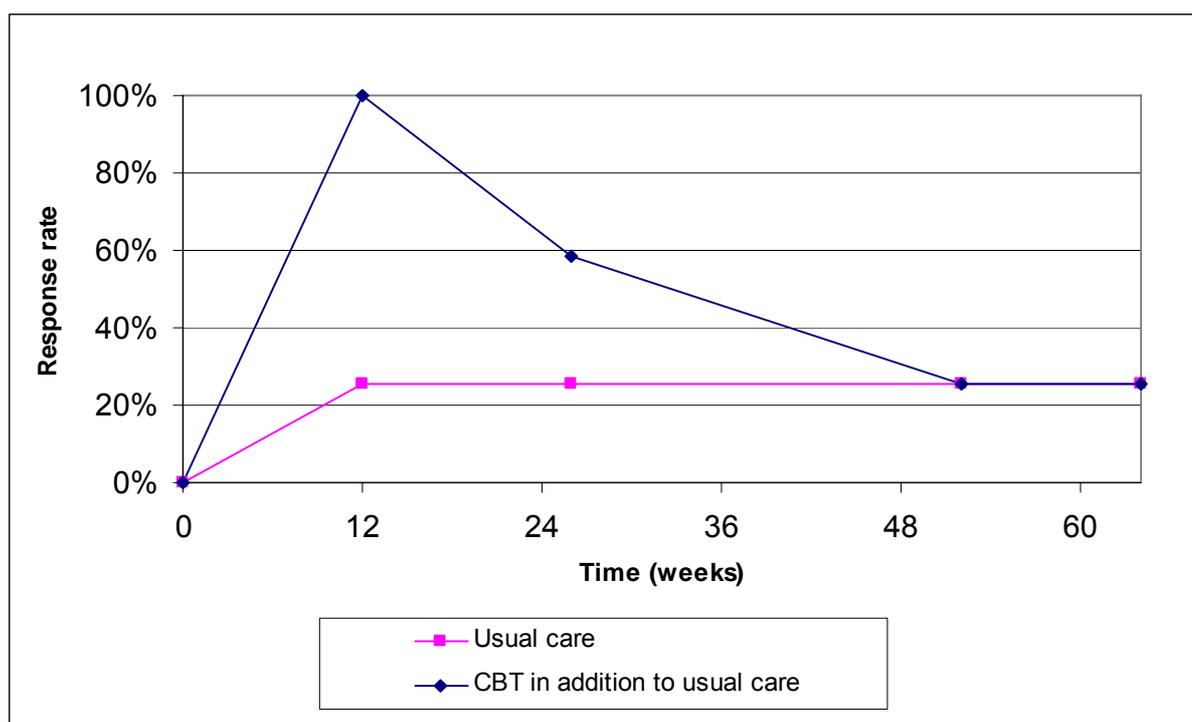
Table 2: Intervention specific parameters – CBT

Description	Value	Evidence
RR of response for intervention vs placebo (at end of treatment)	6.11 (2.33 – 16.07)	Meta-analysis of RCT evidence for improvement in global symptoms
Fall-off in effect at 6 months compared to end of treatment	56% (47% to 66%)	Kennedy (2005), global symptom score
Fall-off in effect at 12 months compared to end of treatment	100% (Fixed)	Kennedy (2005), global symptom score
CBT cost: equiv to 6.6 hours per patient	£375 (£167 - £582).	Weighted mean duration across studies and unit cost from Netten (2006)

CBT in addition to usual care for 100 patients with refractory IBS is estimated to gain an additional 2.24 QALYs for an additional cost of £37,460 compared to usual care alone under the basecase assumptions. The incremental cost per QALY is therefore £16,732. The probabilistic sensitivity analysis considers the uncertainty in this basecase estimate due to the uncertainty in the parameters used to estimate the cost-effectiveness. The CEAC in Figure 41 shows that given the parameter uncertainty, CBT in additional to usual care has a 55% probability of having a cost per QALY under £20,000 and a 79% probability of having a cost per QALY under £30,000, compared to usual care alone.

Figure 41: CEAC for CBT in addition to usual care compared to usual care alone in patients with refractory IBS
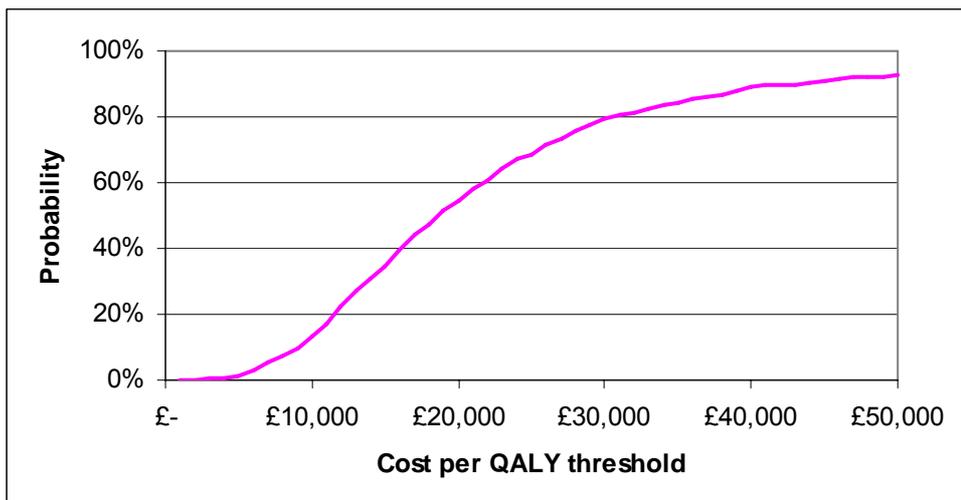
When we assumed that CBT is given over 6 or 8 weeks but at the same cost as in the basecase, the cost per QALY was £15, 771 or £16,079 respectively as the QALY gain was marginally increased (2.38 for 6 weeks and 2.33 for 12 weeks, compared to 2.24 in the

basecase). However, it should be noted that in this sensitivity analysis it was assumed that the response rate in the placebo arm was also achieved over a shorter duration, so that the RR was applied at the end of the intervention period to the same baseline response rate of 25%. These results suggest that assuming a 12 week intervention period in the basecase may have slightly underestimated the cost-effectiveness of CBT but it didn't significantly bias the cost per QALY estimate.

The RR for an improvement in global symptoms for CBT has been applied to a 25% response rate in the comparator arm giving a 100% response rate at 12 weeks for CBT. However, in the CBT trials, the mean response rate in the control arm was 9%. The sensitivity analysis using this lower response rate in the comparator arm gave a cost per QALY of £27,129. As this sensitivity analysis significantly increased the cost per QALY estimate, the probabilistic sensitivity analysis was re-run using this lower response rate for the comparator arm. The mean cost per QALY from the 1000 samples was £25,940 and the cost per QALY had a 31% probability of being under £20,000 per QALY and a 48% probability of being under £30,000 per QALY.

The threshold analysis on utility gain showed that the response to treatment would need to provide more than 0.059 QALYs per annum to give a cost per QALY of under £20,000 in the basecase analysis. When the utility gain associated with a response to treatment was increased to 0.135 (equivalent to the QALY gain expected for a complete remission of symptoms) the cost per QALY was significantly lower at £8,837.

When we assumed no fall-off in response up to 52 weeks post-intervention the cost per QALY was £6,317. This estimate would be further reduced by any continued response beyond 52 weeks. When we assumed that there was no significant difference between CBT and usual care from 6 months, the cost per QALY increased to £28,184. Whilst these two scenarios represent extreme possibilities for the estimated fall-off in response, they demonstrate that the cost-effectiveness is sensitive to the rate of fall-off in response after the end of intervention.

When we assumed that the reduction in resource use observed in the one year after psychotherapy from the Creed (2003) study could also be expected in patients receiving CBT, the incremental cost for CBT reduced to £11,342 and the cost per QALY reduced to £5,066.

Table 3: Sensitivity results for CBT in addition to usual care compared to usual care alone for 100 patients with refractory IBS (all subtypes)

Scenario	Usual care		Behavioural intervention and usual care		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	2.02	£37,460	4.26	£16,732
Intervention given over 6 weeks	£0	2.13	£37,460	4.50	£15,771
Intervention given over 8 weeks	£0	2.09	£37,460	4.42	£16,079
Lower response rate in comparator arm (9%)	£0	0.72	£37,460	2.10	£27,129
No fall-off in effect for 1 year	£0	2.02	£37,460	7.95	£6,317
Effect falls off over first 6 months	£0	2.02	£37,460	3.35	£28,184
Resource use reduction from Creed (2003) study	£0	2.02	£11,342	4.26	£5,066
High utility gain of 0.135	£0.00	3.83	£37,460	8.07	£8,837
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.059 and 0.060.				

Further analyses on the cost-effectiveness of CBT compared to other behavioural interventions are given in section 9.7.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. There is good evidence to show a significant global improvement in symptoms for CBT when compared with no treatment/symptom monitoring, mainly in patients with psychiatric co-morbidities and refractory IBS.

2. There is moderately good evidence to show a borderline global improvement in symptom score for CBT in addition to mebeverine compared with mebeverine alone, at 26 weeks follow up, but there is no significant difference at 52 weeks, in primary care patients with about 50% psychiatric co-morbidities and IBS that did not respond to three months treatment with mebeverine.

3. There is moderately good evidence to show no significant difference in pain and bloating for CBT when compared with no treatment/symptom monitoring in patients, most of whom had psychiatric co-morbidities

4. There is limited evidence to show no significant effect on constipation, but a small, significant improvement in diarrhoea for CBT when compared with no treatment/symptom monitoring in patients, most of whom had psychiatric co-morbidities

5. There is weak evidence to show no significant difference in quality of life (IBS QoL) for CBT when compared with no treatment/symptom monitoring in patients with psychiatric co-morbidities

6. There is moderately good evidence to show a significant global improvement in symptom score when CBT is added to mebeverine when compared with mebeverine alone.

HEALTH ECONOMIC STATEMENT

Evidence from a trial based economic evaluation showed that the addition of CBT to antispasmodic therapy does not result in lower service costs at 3, 6 or 12 months in individuals with symptoms of moderate or greater severity after 2 weeks of GP care and 4 weeks of mebeverine.

Evidence from a decision analytic model showed that the addition of CBT to usual care is cost-effective in individual with refractory IBS The ICER is sensitive to the proportion of patients experiencing an improvement in global symptom score with usual care alone and the efficacy in the 9 months after intervention, although none of the sensitivity analyses generated ICERs

above £30,000 per QALY. A threshold analysis showed that an improvement in global symptoms must result in a utility gain of at least 0.06 for the cost per QALY to remain below £20,000.

GDG DISCUSSION

The majority of people in the randomised trials had psychiatric co-morbidities and it is the view of the GDG that these could have skewed data when seeking to apply trial findings to the IBS population as a whole.

Generally, CBT has a positive benefit in improving global symptom scores for people with IBS in the trials. Meta-analysis demonstrates the benefit of CBT in producing an initial big treatment effect. The GDG view is that people with IBS are likely to feel that they are coping better with their symptoms, whilst recognising the potential for a treatment tail off. Even though there is some evidence that there is sustainable treatment effect, tail-off is usually addressed by a top up session.

CBT has not generally been used as a first line therapy for the management of IBS, but the GDG agreed that this needs to be investigated further. The GDG therefore decided to include CBT in one of its top five research recommendations.

EVIDENCE TO RECOMMENDATION

The evidence to recommendation statement for psychotherapy, CBT and hypnotherapy is detailed in section 9.8.

The combined guideline recommendation for psychotherapy, CBT and hypnotherapy is also stated in section 9.8.

9.6 Hypnotherapy

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, except that crossover studies were excluded as inappropriate due to the carry-over effect of the hypnotherapy interventions.

The following comparisons were included:
- Hypnotherapy versus waiting list control, or symptom monitoring only
- Hypnotherapy versus usual medical care
- Hypnotherapy individual versus hypnotherapy group
- Hypnotherapy versus another intervention (e.g. psychotherapy or relaxation).

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the PSYCINFO database was searched for this review. The search strategies are listed in Appendix B.

The titles and abstracts from the search strategy were assessed. Nineteen were identified to be potentially relevant to the review and these papers were retrieved in full. The reference lists of the retrieved studies were inspected for further potential papers, but none were identified. The 13 excluded studies are listed in Appendix E, along with reasons for exclusion.

Study Design

Six parallel group design randomised trials were included (Forbes 2000; Galovski 1998; Harvey 1989; Palsson 2002; Roberts 2006; Whorwell 1984). Further details are given in the included studies table.

Four of the studies were carried out in the UK (Forbes 2000; Harvey 1989; Roberts 2006; Whorwell 1984). The remaining studies were carried out in the USA (Galovski 1998; Palsson 2002). Trials lasted between 6 and 12 weeks. One study was conducted among patients recruited from primary care (Roberts 2006); the others were in secondary care.

The total number of patients in the studies ranged from 12 to 81. Only two studies included more than 25 patients in a treatment arm (Forbes 2000; Roberts 2006). Forbes included 25 and 27 patients in the two treatment arms respectively. In Roberts (2006) a power calculation was done, which suggested that 50 patients per group would be needed; however, the study only recruited 40 patients in one arm and 41 patients in the other, so it was underpowered. On the basis of this power calculation, it is likely that all the studies are underpowered.

Population

All the studies included only patients with IBS. None of the studies reported the number of patients with bloating or whether the symptoms were post-infective, and it was unclear if the patients had pain at baseline. The mean age of patients was around 40 years, with those aged between 18 and 65 years included. All the studies included more women than men.

IBS was stated, or implied, to be refractory all of the studies. The patients in Forbes (2000) had had IBS for more than six months and the inclusion criteria required that they had failed on conventional treatments, with the exception of antidepressants. Palsson (2002) stated that the patients all had symptoms refractory to standard medical management. Galovski (1998) had patients with a mean duration of IBS symptoms of six years (range 0.5 to 17years). Harvey (1989) did not report the duration of symptoms, but stated that the patients had refractory IBS.

Roberts (2006) included primary care patients with IBS for more than six weeks, who were said to have failed conventional management. Whorwell (1984) included patients with severe refractory IBS who had not responded to any therapy over at least one year (the mean number of therapies previously tried per patient was six).

The patients in Forbes (2000) and Roberts (2006) were allowed to continue pre-existing therapy for IBS, including antispasmodics and antidepressants; those in Palsson (2002) discontinued their IBS medication. Continued medication use was not stated in Galovski (1998), Harvey (1989), or Whorwell (1984).

In Galovski (1998), 67% of patients had an Axis I diagnosis; one patient with bipolar disorder with a current manic state was excluded. Forbes (2000) stated that 19/52 (37%) of patients were considered to be psychiatric cases according to the GHQ. Harvey (1989) reported that 8/22 (36%) had psychological problems (GHQ≥5). Patients requiring psychotropic medications were excluded from Palsson (2002). Psychiatric co-morbidities were not stated in the other two studies (Roberts 2006; Whorwell 1984).

Interventions

The hypnotherapy interventions were all 'gut-directed hypnotherapy' based on the methods described by Whorwell in 1984. All the trials assessed individual therapy; in one trial the comparator was group hypnotherapy (Harvey 1989). The studies varied in how hypnotherapy was delivered: Roberts (2006) had 5 weekly half-hour sessions and follow-up data were available at 3, 6 and 12 months (not end of therapy). Harvey (1989) had four 40-minute sessions over 7 weeks; Palsson (2002) had seven 45-minute sessions over 12 weeks; Forbes (2000) had six 30-minute sessions over 12 weeks, and; Whorwell (1984) had seven 30-minute sessions over 3 months.

Hypnotherapy was compared with relaxation training, psychotherapy, symptom monitoring, waiting list control, and usual care. In one trial (Harvey 1989), hypnotherapy in groups was compared with individual therapy. In the Whorwell (1984) trial, hypnotherapy was compared with 'psychotherapy', but author communication described this as 'supportive listening', more akin to attention control than psychotherapy. In Roberts (2006), IBS medication was continued in both groups, so that their comparison, hypnotherapy versus usual management, was, in reality, hypnotherapy versus no treatment. It was agreed to combine the comparators, waiting list control, attention control, symptom monitoring and no treatment / usual care.

The following comparisons were included:

- Hypnotherapy versus: a waiting list control group; attention control; symptom monitoring only, or; usual care (four studies: Galovski (1998); Palsson (2002); Roberts (2006); Whorwell (1984)):

- o Hypnotherapy versus waiting list control (Palsson 2002)
- o Hypnotherapy versus symptom monitoring (Galovski 1998)
- o Hypnotherapy versus usual management (conventional medication in primary care) (Roberts 2006)
- o Hypnotherapy versus attention control + placebo tablet (both therapies delivered by same therapist) (Whorwell 1984);
- Group hypnotherapy versus individual hypnotherapy:
  - o Harvey (1989);
- Hypnotherapy versus another intervention:
  - o Hypnotherapy versus audiotape on relaxation (tape produced by same therapist) (Forbes 2000).

METHODOLOGICAL QUALITY

The results of the quality assessment for included trials are shown in Appendix D.

One study (Forbes 2000) reported an adequate method of randomisation (computer-generated random numbers); the other studies did not state the method. Allocation concealment was partially adequate in one study (Roberts 2006, sealed envelopes); the other studies were unclear. The patients were not blinded (because of the type of intervention). However, the GDG did not consider this to be important for the behavioural interventions. One study (Roberts 2006) described an *a-priori* power calculation, but did not meet the required number of patients during recruitment.

The comparability of groups at baseline varied amongst the studies:
- Two studies demonstrated baseline comparability of the groups (Forbes 2000; Galovski 1998)
- Two were mainly comparable:
  - o Roberts (2006) reported that there were more males in the intervention group (8/40 versus 4/41); and there were some differences in baseline quality of life scores (on three of eight subscales, p value not given)
  - o Whorwell (1984) reported that bowel habit was more severely disordered in patients receiving hypnotherapy than in control patients (intervention group baseline score 17.2 versus controls 12.8; where abnormality of bowel habit was scored 0=none, 1=mild, 2=moderate or 3=severe, and scores totalled over 7 days, i.e. scale from 0 to 21, p=0.005 for baseline difference);
- Two did not state the comparability for the randomised population (Harvey 1989; Palsson 2002).
  - o However, Palsson (2002) gave baseline pain and bloating scores and the proportion of hard/loose stools only for completers, and these were not comparable across

groups (more severe pain and bloating and lower proportion of hard/loose stools for the intervention group).

All the patients were followed up in two studies (Galovski 1998; Whorwell 1984). There were 20% or fewer drop-outs overall in two studies (Harvey 1989; Forbes 2000). In Harvey (1989), 3 out of 36 [8%] were missing and Forbes (2000) had 7/52 missing data for symptom diaries, but only 25/52 (48%) complied with the follow up for psychological outcomes. One study (Palsson 2002) had more than 20% missing data in the control group: the 6 drop-outs in the study were all from the control group (i.e. 40% drop-out in this group), however, the study stated that the drop outs were related to non-treatment related causes such as relocation, scheduling difficulties and unrelated medical problems. Nevertheless we regarded this study with caution because this unequal drop-out could still have introduced a bias. In the other study (Roberts 2006), data were missing for 18% of patients at 3 months; 17% at 6 months and 35% at 12 month follow-up. However, the study stated that analysis indicated that the missing data were 'missing completely at random', so that the results for the missing data would not be significantly different from those that completed the study.

Overall, there is a risk of bias in the Palsson (2002) study for the pain and bloating outcomes, and the uneven drop-out rates between the groups should be taken into consideration (40% drop-out among controls versus none from the intervention group) and the differences at baseline. Forbes (2000) was considered at high risk for the psychological outcomes. The Roberts (2006) 12 month follow-up data should be regarded with caution, also due to the fairly high drop-out rate (35%). The difference in baseline for bowel habit should be taken into consideration in the Whorwell (1984) study.

RESULTS

A. Hypnotherapy versus waiting list control group, attention control, symptom monitoring only or usual management

Four studies compared hypnotherapy with a waiting list control group; attention control; symptom monitoring only, or; usual management in patients with IBS (Galovski 1998; Palsson 2002; Roberts 2006; Whorwell 1984).

1. Global symptoms

a) Number of patients with global improvement in symptoms

This outcome was reported by Galovski 1998 at 6 weeks for hypnotherapy versus symptom monitoring in 11 patients.

Figure 1


```
Review:       IBS 31 August 2006 (hyp_other pysch)
Comparison:   24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome:      01 Global improvement (number of patients)
```

Study or sub-category	Hypnotherapy n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Galovski 1998	4/5	0/6		100.00	10.50 [0.70, 157.91]
Total (95% CI)	5	6		100.00	10.50 [0.70, 157.91]

Total events: 4 (Hypnotherapy), 0 (Control)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.70 (P = 0.09)

```
        0.01    0.1      1      10     100
        Favours control    Favours treatment
```

The confidence interval was too wide to determine if there is an effect (but see next section for further evidence on global improvement of symptoms – number of patients).

b) Global improvement of symptoms score

The Whorwell (1984) study in 30 patients reported the overall improvement of symptoms and general wellbeing, scored weekly on a scale of 0 to 3 (where 0 is no improvement and 3 is maximum improvement). From a baseline score of 0 in both groups, patients in the hypnotherapy group increased to a mean weekly value of 2.95 and those in the psychotherapy group increased to 0.52, i.e. a difference of 2.43. This was reported to be statistically significant (p<0.0001), i.e. a large effect.

Chinn (2000) introduced a statistical approach that re-expresses standardised mean differences as odds ratios, according to the following simple formula:

$$\log OR = (\pi/\sqrt{3})\, SMD$$

The standard error of the standardised mean difference can be converted to the standard error of the log odds ratio by multiplying by $\pi/\sqrt{3} = 1.8140$. We carried out this procedure for Whorwell (1984) in order to combine the data with those of Galovski (1998). This involved calculation of the standard error from the p value, conversion of the mean difference to standardised mean difference by dividing both MD and standard error by the standard deviation and then converting to log OR.

Meta-analysis of the two studies, in 41 patients, gave a pooled odds ratio of 3.85 (95%CI 2.03, 7.29), with non-significant heterogeneity (I^2=45%, p=0.16). This was statistically significant, in favour of hypnotherapy.

Figure 2: Meta-analysis for global improvement of symptoms (number of patients)

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome: 13 Global improvement of symptoms

Study or sub-category	log[OR] (SE)	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 Whorwell (OR calculated from mean difference and p value)				
Whorwell 1984	1.2646 (0.3312)		96.52	3.54 [1.85, 6.78]
Subtotal (95% CI)			96.52	3.54 [1.85, 6.78]
Test for heterogeneity: not applicable				
Test for overall effect: Z = 3.82 (P = 0.0001)				
02 Galovski (OR calculated directly)				
Galovski 1998	3.6636 (1.7438)		3.48	39.00 [1.28, 1189.62]
Subtotal (95% CI)			3.48	39.00 [1.28, 1189.62]
Test for heterogeneity: not applicable				
Test for overall effect: Z = 2.10 (P = 0.04)				
Total (95% CI)			100.00	3.85 [2.03, 7.29]
Test for heterogeneity: Chi² = 1.83, df = 1 (P = 0.18), I² = 45.3%				
Test for overall effect: Z = 4.14 (P < 0.0001)				

0.001 0.01 0.1 1 10 100 1000
Favours control   Favours hypnotherapy

Galovski (1998) reported the global improvement of IBS symptoms at six weeks in 11 patients using the Composite Primary Symptom Reduction (CPSR) score; CPSR represents the proportional reduction in the score from baseline; scale -1 to +1. The confidence interval was too wide to determine if there was a difference between interventions.

Figure 3

Review: IBS 31 August 2006 (hyp_other pysch)
Comparison: 24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome: 02 CPSR score

Study or sub-category	N	Hypnotherapy Mean (SD)	N	Control Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Galovski 1998	5	-0.55 (0.53)	6	-0.32 (0.49)		100.00	-0.23 [-0.84, 0.38]
Total (95% CI)	5		6			100.00	-0.23 [-0.84, 0.38]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.74 (P = 0.46)							

-1 -0.5 0 0.5 1
Favours treatment   Favours control
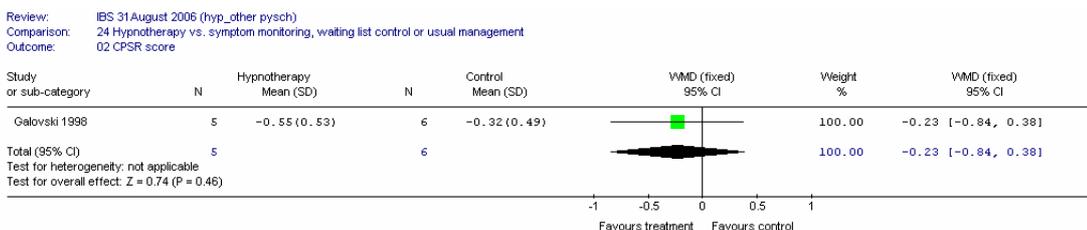
c) Global symptom score

The change over baseline in global symptom score was reported by Roberts (2006) at 12 weeks (about 7 weeks after the end of treatment) for hypnotherapy versus usual IBS care in 81primary care, refractory patients and at 26 and 52 weeks follow-up. 26 week standard deviations were not given, although the means were, and we noted that there was 35% missing data at 52 weeks (although the authors showed this to be missing-at-random, which made the results more acceptable). There was a statistically significant improvement in symptom score at 12 weeks, favouring hypnotherapy. The scale was not given, but reference was made to a questionnaire using 22 items each rated at 1-7 (7=high) (Wiklund 2003). This would have meant a maximum score of 154, but this was not entirely clear. The baseline scores were about 40, so a change of 8.5 units seems a reasonable effect size. At six months, the decreases in symptom score were 10 and 8 for the intervention and control groups respectively, i.e. a change of -2 units. At 12 months (follow-up) the change in symptom score was -2.70 (95%CI -10.48, 5.09), i.e., no longer significant.

Figure 4


```
Review:      IBS 31 August 2006 (hyp_other pysch)
Comparison:  24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome:     10 Global symptom score

Study                      Hypnotherapy        Control          WMD (fixed)    Weight    WMD (fixed)
or sub-category      N     Mean (SD)       N    Mean (SD)         95% CI          %        95% CI

01 Change from baseline 12 weeks
  Roberts 2006      33   -13.00(10.30)    34  -4.50(14.60)                      100.00   -8.50 [-14.54, -2.46]
Subtotal (95% CI)   33                    34                                    100.00   -8.50 [-14.54, -2.46]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.76 (P = 0.006)

02 Change from baseline 52 weeks follow up
  Roberts 2006      29    -9.10(14.00)    24  -6.40(14.70)                      100.00   -2.70 [-10.48, 5.08]
Subtotal (95% CI)   29                    24                                    100.00   -2.70 [-10.48, 5.08]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.68 (P = 0.50)

                                            -100   -50    0    50   100
                                         Favours hypnotherapy  Favours control
```

2. Individual symptoms

a) Pain

A pain score was reported by three studies (Palsson 2002; Roberts 2006; Whorwell 1984), all at 12 weeks. Palsson (2002) compared hypnotherapy versus waiting list control in 30 patients; neither had concurrent IBS medical treatments, and recorded pain score on a scale of 0 to 4 recorded over 14 days, where 0=none, 1=mild, 2=moderate, 3=severe and 4=incapacitating, i.e. maximum 56. We noted that this study had 40% missing data in the control group, in addition, there was a significant difference in the baseline pain score for completers in the intervention group was of 7.9 units, which was large compared to the difference in effect size (11.8 units). Therefore the results from this study were considered to be potentially biased and are therefore not reported here.

Roberts (2006) showed a significant difference in the pain score of -14.40 (95%CI -24.69, -4.11) at 3 months, but this was no longer significant at 12 months. The baseline scores were 53-55. Again there was 35% missing data, said to be missing-at-random.

Whorwell (1984) recorded a pain score (0-3 recorded over 7 days, where 0=none, 1=mild, 2=moderate, 3=severe, i.e. maximum 21). From a baseline score of 13 in both groups, patients receiving hypnotherapy reduced their mean score to 2.2 (i.e. a fall of 10.8), while those on psychotherapy had a mean score of 11.6 at 12 weeks (i.e. a fall of only 1.4); no standard deviations were given, but the difference between groups of -9.4 was reported to be statistically significant (p<0.0001).

Figure 5


```
Review:      IBS 31 August 2006 (hyp_other pysch)
Comparison:  24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome:     11 Pain score

Study                      Hypnotherapy        Control          WMD (fixed)    Weight    WMD (fixed)
or sub-category      N     Mean (SD)       N    Mean (SD)         95% CI          %        95% CI

01 Change from baseline 12 weeks
  Roberts 2006      33   -21.20(19.53)    34  -6.80(23.32)                      100.00   -14.40 [-24.69, -4.11]
Subtotal (95% CI)   33                    34                                    100.00   -14.40 [-24.69, -4.11]
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.74 (P = 0.006)

02 Change from baseline 52 weeks follow up
  Roberts 2006      29   -16.30(22.08)    24 -15.70(24.49)                      100.00   -0.60 [-13.27, 12.07]
Subtotal (95% CI)   29                    24                                    100.00   -0.60 [-13.27, 12.07]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.09 (P = 0.93)

                                            -100   -50    0    50   100
                                         Favours hypnotherapy  Favours control
```

b) Bloating

Two studies reported a bloating score at 12 weeks (Palsson 2002; Whorwell 1984). Palsson (2002) used a scale of 0 to 4 recorded over 14 days, where 0=none, 1=mild, 2=moderate, 3=severe and 4=incapacitating. However, the baseline values for bloating were much lower for the control group (data given for completers only, mean 13.6 at baseline) than the intervention group (mean 20.3 at baseline), and there was 40% missing data in the control group. The study was therefore considered to be confounded for this outcome and was not considered further.

Whorwell (1984) measured a bloating score (0-3 recorded over 7 days, where 0=none, 1=mild, 2=moderate, 3=severe). From a baseline score of around 16, patients receiving hypnotherapy reduced their mean score to 3.2 (i.e. a fall of 12.8), while those receiving supportive listening ('psychotherapy') had a mean score of 13.2 at 12 weeks (i.e. a fall of only 2.8); no standard deviations were given, but the difference between groups of -10.0 was reported to be statistically significant (p<0.0001).

c) Bowel habit

Roberts (2006) reported scores for constipation and diarrhoea. There was a non-significant difference (Figure 6), favouring hypnotherapy, in the diarrhoea score of -7.90 (95%CI -16.29, 0.49) at 3 months, but very little difference at 12 months. The baseline scores were about 33. Again there was 35% missing data, said to be missing-at-random. For constipation, there was no significant effect at any time (Figure 7). Baseline scores were around 38.

Figure 6
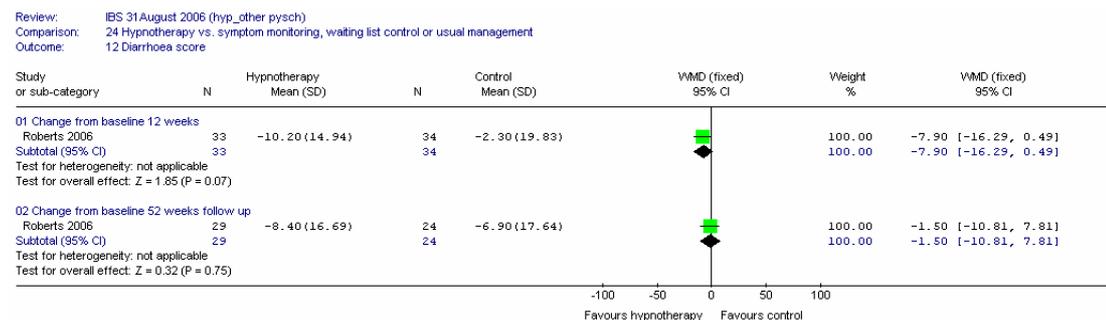
Figure 7
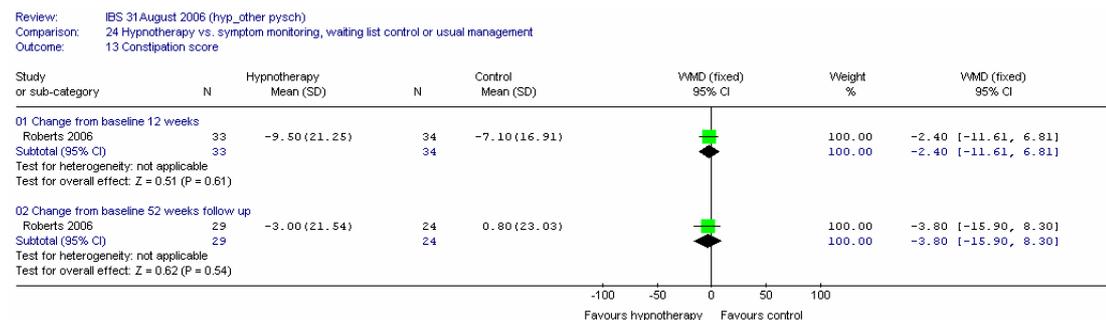
Abnormality of bowel habit was reported by Whorwell (1984). At baseline, this was more severely disordered in patients on hypnotherapy than in control patients (intervention group baseline score 17.2 versus psychotherapy 12.8 (i.e. baseline difference of 4.4); where abnormality of bowel habit was scored 0=none, 1=mild, 2=moderate or 3=severe, and scores totalled over 7 days, i.e. scale from 0-21, p=0.005 for baseline difference). The score fell from 17.2 to 1.6 (i.e. 15.6) in the hypnotherapy group compared with from 12.8 to 11.8 (i.e. 1.0) on psychotherapy (p<0.0001). The large baseline difference may have confounded this outcome measure.

Palsson (2002) reported stool frequency. There was no significant difference in the baseline values, or at 3 months.

Figure 8


```
Review:      IBS 31 August 2006 (hyp_other pysch)
Comparison:  24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome:     14 Stool frequency

Study                    Hypnotherapy          Control           WMD (fixed)      Weight      WMD (fixed)
or sub-category      N   Mean (SD)        N    Mean (SD)          95% CI           %           95% CI

01 daily frequency at 12 weeks
  Palsson 2002       15  2.30(1.94)       9    2.10(1.20)                          100.00      0.20 [-1.06, 1.46]
Subtotal (95% CI)    15                   9                                        100.00      0.20 [-1.06, 1.46]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.31 (P = 0.76)

                                          -4    -2      0      2      4
                                     Favours hypnotherapy  Favours control
```

3. Quality of life

IBS-specific quality of life (high = good) was reported by Roberts (2006) at 3 months (about 7 weeks after end of treatment) in 81 patients and at 6 and 12 month follow-up. There was no significant difference at any time, although the difference in QoL score did not appear to change over time. Again the scale was not given, but reference to the Wiklund (2003) study suggested that the scale was 26 items with a 7 point Likert scale, giving a possible maximum of 182. Baseline scores were about 50.

Figure 9


```
Review:      IBS 31 August 2006 (hyp_other pysch)
Comparison:  24 Hypnotherapy vs. symptom monitoring, waiting list control or usual management
Outcome:     09 Quality of life

Study                    Hypnotherapy          Control           WMD (fixed)      Weight      WMD (fixed)
or sub-category      N   Mean (SD)        N    Mean (SD)          95% CI           %           95% CI

01 3 months
  Roberts 2006       33  60.20(25.30)     34   51.50(22.70)                        100.00      8.70 [-2.82, 20.22]
Subtotal (95% CI)    33                   34                                       100.00      8.70 [-2.82, 20.22]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.48 (P = 0.14)

02 6-month follow up
  Roberts 2006       33  63.30(28.70)     34   53.80(26.20)                        100.00      9.50 [-3.67, 22.67]
Subtotal (95% CI)    33                   34                                       100.00      9.50 [-3.67, 22.67]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.41 (P = 0.16)

03 12-month follow up
  Roberts 2006       29  65.40(23.69)     24   55.80(25.47)                        100.00      9.60 [-3.75, 22.95]
Subtotal (95% CI)    29                   24                                       100.00      9.60 [-3.75, 22.95]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.41 (P = 0.16)

                                   -100   -50     0     50    100
                                     Favours control   Favours treatment
```

4. Use of IBS medication

One study (Roberts 2006) gave the self-reported use of prescription medication, either sometime, or continual, and over-the-counter medication (including antispasmodics, anti-motility agents, probiotics, herbal juices and teas, and incontinence pads) over the 12 months follow-up period. There were significantly more patients using prescription medication at some time during the 12 months; RR 0.61 (95%CI 0, 40, 0.94). This corresponded to an NNT of 4 (95%CI 2, 14), for a control group risk of 79%. There was no significant difference in the over-the-counter medication use.

Figure 10
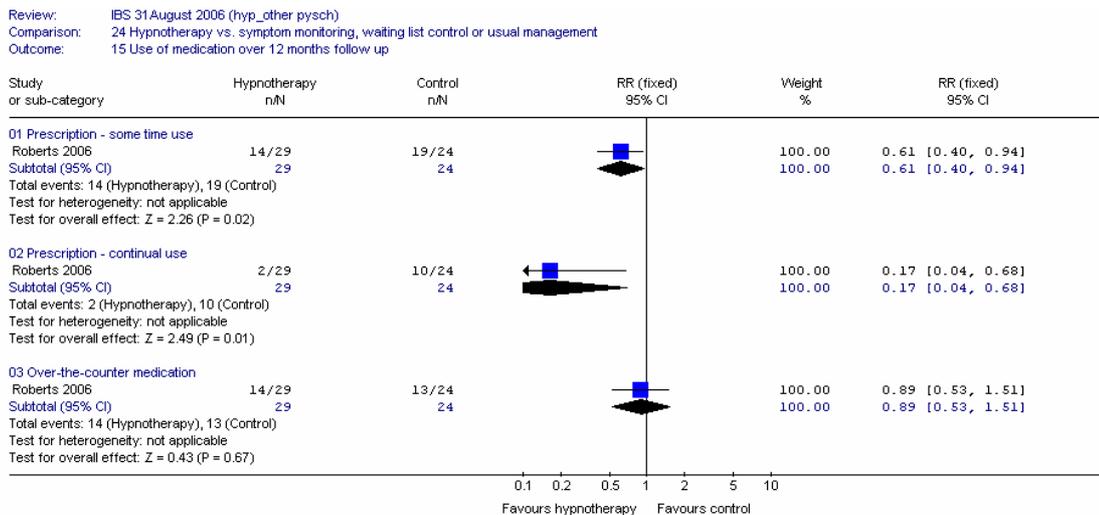
B. Hypnotherapy versus another intervention (relaxation)

Forbes (2000) compared hypnotherapy with relaxation in 52 patients. We noted that in both studies the two types of therapy were delivered by the same person, which could have introduced a therapist effect.

1. Global symptoms

Global improvement in symptoms (number of patients) was reported by Forbes (2000) at 12 weeks. There was no significant difference between interventions.

Figure 11
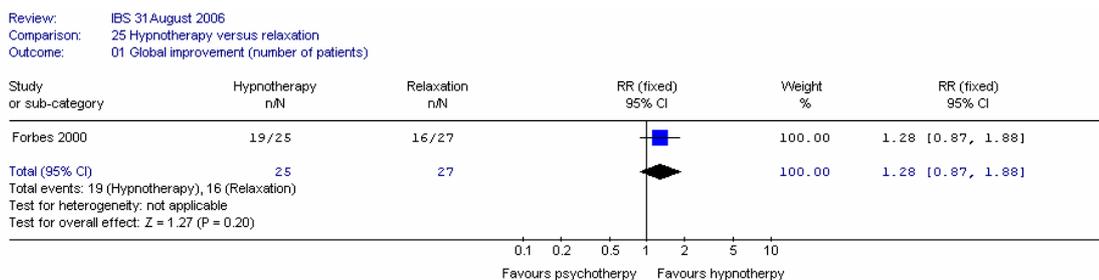
C. Group hypnotherapy versus individual hypnotherapy

One study in 33 patients (Harvey 1989) compared hypnotherapy on an individual basis versus hypnotherapy on a group basis (6 to 8 patients); the outcome reported was the global improvement in symptoms (number of patients). There was no significant difference between interventions.

1. Global outcomes

Global improvement in symptoms (number of patients)

Figure 12

Review:	IBS 31 August 2006
Comparison:	27 Group hypnotherapy vs. individual hypnotehrapy
Outcome:	01 Global improvement (number of patients)

Study or sub-category	Group hypno. n/N	Individual hypno. n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Harvey 1989	12/17	8/16		100.00	1.41 [0.79, 2.52]
Total (95% CI)	17	16		100.00	1.41 [0.79, 2.52]

Total events: 12 (Group hypno.), 8 (Individual hypno.)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.17 (P = 0.24)

0.1 0.2 0.5 1 2 5 10
Favours individual   Favours group

ECONOMIC LITERATURE FOR HYPNOTHERAPY

No relevant health economic analyses were identified on the cost-effectiveness of hypnotherapy in the management of IBS.

COST-EFFECTIVENESS ANALYSIS FOR HYPNOTHERAPY

This section describes the health economic analysis undertaken to inform recommendations on the use of hypnotherapy as a one-off intervention in the management of IBS. The general methods used in the economic analysis for all management interventions are described in detail in Chapter 5 and the model inputs and assumptions relevant to this particular intervention are described below.

▪ The effectiveness of hypnotherapy in addition to usual care compared to usual care alone in people with refractory IBS was based on the number of patients with an improvement in global symptoms (at the end of treatment) for hypnotherapy vs waiting list control, symptom monitoring, attention control or usual care. (OR 3.85, 95%CI 2.03- 7.29, from meta-analysis of Whorwell (1984) and Galvoski (1998), giving a RR of 2.23, 95% CI 1.16 – 2.80, for a 25% response rate in the control arm).

▪ We assumed that there is no further benefit after 12 months based on a non significant difference in mean global symptom score at 12 months (Roberts 2006). A linear fall-off was assumed between the end of treatment and 12 months. A sensitivity analysis assuming no further benefit after 6 months was carried out as the mean difference in global symptom score is similar at 6 months and 12 months but it is not possible to calculate statistical significance from the data presented for 6 months (Roberts 2006).

- The evidence included in the clinical effectiveness review did not allow a subtype specific estimate of clinical effectiveness to be estimated. Therefore it was assumed that hypnotherapy is equally effective in all IBS subtypes.
- Hypnotherapy was assumed to be given over 12 weeks as this was the duration of intervention in the Whorwell (1984) and Galovski (1998) studies.
- A 15 month time-frame was used so that the cost-effectiveness could be compared to against other psychological interventions for which there was 15 month efficacy data.

Modelled response rates

In the basecase scenario the response rate of 25% in the no treatment arm is taken from the mean placebo arm response rate from the behavioural therapy trials. This represents the group of patients whose symptoms improve without any specific intervention. The RR for an improvement in global symptoms for hypnotherapy vs no treatment at the end of treatment is 2.23, therefore the response rate in the intervention arm is 57% at the end of treatment (12 weeks). As shown in Figure 13, the response rate in the hypnotherapy arm has decreased to 46% by 6 months and 25% by one year, based on the assumptions regarding fall-off in effectiveness described above.

We have also considered a maintained benefits scenario in which the response to hypnotherapy is maintained for the one year after the end of treatment but there is no further benefit beyond this point.

There was no NHS reference cost available for hypnotherapy, even though it is funded in some regions of the NHS. A typical salary for a hypnotherapist falls within the Agenda for Change band 6 (based on personal communication from Peter Whorwell). This is the same salary used in estimating the reference cost for counsellors, on which the cost estimate for psychotherapy has been based. We have assumed that hypnotherapists have a similar working pattern to counsellors undertaking psychotherapy in terms of the proportion of their time that is spent on direct client contact and the proportion that is spent on research, administration, education and other activities. Therefore the cost per hour for hypnotherapy has been taken to be equivalent to the cost per hour for psychotherapy. The costs of hypnotherapy were based on the mean number and duration of sessions used in the Whorwell (1984) and Galoviski (1998) studies, weighted by their contribution to the meta-analysis. This gave a mean duration of 3.6 hours of hypnotherapy. As there were only two studies used to estimate the RR, the cost range was based on the range from the various studies included in the clinical effectiveness review (2.2 – 4.9 hours). This gave a total cost for hypnotherapy of £171, (range £105 - £237).

For hypnotherapy there was evidence from Roberts (2006) that hypnotherapy resulted in a significant reduction in the use of prescriptions in the 1 year following intervention: RR of 0.61 (0.40 – 0.91) for any prescription use and RR of 0.17 (0.04 to 0.68) for continual prescription use

for hypnotherapy compared to control. We have assumed no reduced resource use in the basecase analysis as reduced prescription rates are unlikely to have a significant cost impact. It was included in a sensitivity analysis by applying the reduction in resource use observed during the follow-up period of the Creed (2003) study for psychotherapy compared to usual care (£-4.08 per week, 95%CI-£8.11 to -£0.04) indirectly to hypnotherapy. This is plausible given that the odds ratio for an improvement in global symptom score at the end of treatment is larger for hypnotherapy compared to usual care (3.85, 95% CI 2.03 – 7.29) than the odds ratio observed for psychotherapy vs usual care in the Creed (2003) study (OR=2.44, 95%CI 1.28 – 4.67).

Figure 13: Response rate in the basecase analysis
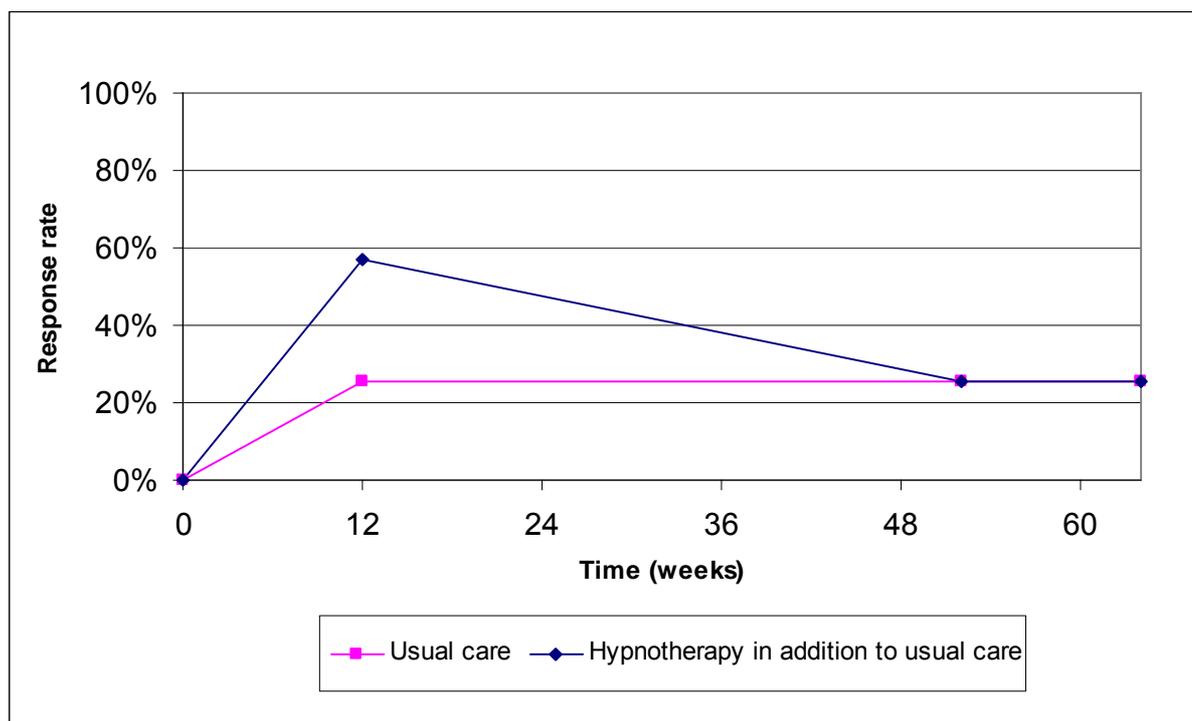
Table 1: Intervention specific parameters – Hypnotherapy

Description	Value	Evidence
RR of response for intervention vs placebo (at end of treatment)	2.23 (1.61 – 2.80)	Meta-analysis of RCT evidence for improvement in global symptoms
Fall-off in effect at 12 months compared to end of treatment	100%	Roberts (2006) global symptom score
Hypnotherapy cost: equiv to 3.6 hours per patient	£171, (range £105 - £237)	Weighted mean duration across studies and unit cost from Netten (2006)

Hypnotherapy in addition to usual care for 100 patients with refractory IBS is estimated to gain an additional 1.12 QALYs for an additional cost of £17,092 compared to usual care alone under the basecase assumptions. The incremental cost per QALY for is therefore £15,300. The probabilistic sensitivity analysis considers the uncertainty in this basecase estimate due to the uncertainty in the parameters used to estimate the cost-effectiveness. The CEAC in Figure 14, shows that given the parameter uncertainty, hypnotherapy in additional to usual care has a 59% probability of having a cost per QALY under £20,000 and a 81% probability of having a cost per QALY under £30,000, compared to usual care alone.

Figure 14: CEAC for hypnotherapy in addition to usual care compared to usual care alone in patients with refractory IBS
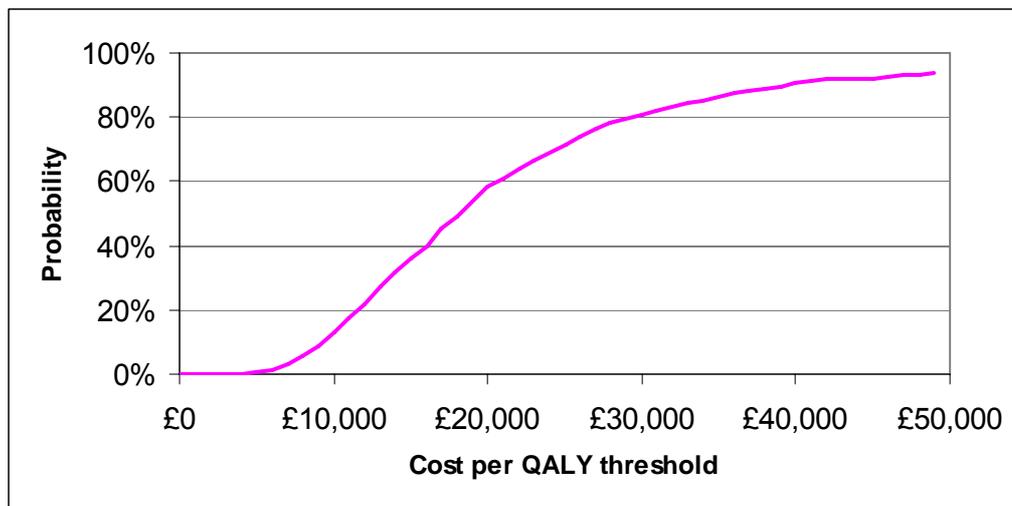
The incremental cost-effectiveness is dependent on the probability of an improvement for patients who receive usual care. When we applied a lower response rate of 9% in the usual care arm, the cost per QALY was increased to £25,809. It should be noted that the odds ratio rather than the RR was kept constant for this analysis as this was the efficacy estimate available from the clinical effectiveness review. As this sensitivity analysis significantly increased the cost per QALY estimate, the probabilistic sensitivity analysis was re-run using this lower response rate for the comparator arm. The mean cost per QALY from the 1000 samples was £25,770 and the cost per QALY had a 28% probability of being under £20,000 per QALY and a 51% probability of being under £30,000 per QALY.

The threshold analysis showed that a response to treatment would need to provide more than 0.054 QALYs per annum to give a cost per QALY of under £20,000 in the basecase analysis. When the utility gain associated with a response to treatment was increased to 0.135 (equivalent to the QALY gain expected for a complete remission of symptoms) the cost per QALY was significantly lower at £8,081

When we assumed there was no fall-off in response up to 52 weeks post-intervention, the cost per QALY was decreased to £6,859. This would be further reduced by any continued response beyond 52 weeks. When we assumed that there was no significant difference between hypnotherapy and usual care from 6 months, then the cost per QALY is increased to £30,601. Whilst these two scenarios represent extreme possibilities for the estimated fall-off in response, they demonstrate that the cost-effectiveness is sensitive to the rate of fall-off in response after the end of intervention.

When we assumed that the reduction in resource use observed in the one year after psychotherapy from the Creed (2003) study could also be expected in patients receiving hypnotherapy, the cost of providing hypnotherapy in additional to usual care was lower than the cost of providing hypnotherapy alone. Under these assumptions hypnotherapy in addition to usual care dominated usual care alone by providing significant health gains, whilst lowering cost.

Table 2: Sensitivity results for hypnotherapy in addition to usual care compared to usual care alone for 100 patients with refractory IBS (all subtypes)

Scenario	Usual care		Behavioural intervention and usual care		Incremental
	Cost	QALY	Cost	QALY	Cost per QALY
Basecase	£0	2.02	£17,092	3.14	£15,300
Lower response rate in comparator arm (9%)	£0	0.72	£17,092	1.38	£25,809
No fall-off in effect for 1 year	£0	2.02	£17,092	4.51	£6,859
Effect falls off over first 6 months	£0	2.02	£17,092	2.58	£30,601
Resource use reduction from Creed (2003) study	£0	2.02	-£9,026	3.14	-£8,080 Hypnotherapy dominates
High utility gain of 0.135	£0	3.83	£17,092	5.94	£8,081
Threshold analysis on lowest utility	A cost per QALY of £20,000 is reached when the QALY gain associated with responding to treatment lies between 0.054 and 0.055.				

Further analyses on the cost-effectiveness of hypnotherapy compared to other behavioural interventions are given in section 9.7.

EVIDENCE STATEMENTS

For this review, the evidence was assessed using the GRADE process and tables are shown in Appendix F. The following evidence statements are derived from the GRADE tables.

1. There is moderately good evidence to show a significant global improvement in symptoms after 12 weeks, for hypnotherapy compared with attention control or symptom monitoring or usual management, mainly in patients with refractory IBS, both in primary and secondary care.

2. There is moderately good evidence to show no significant improvement either in global symptoms or in pain after 52 weeks, for hypnotherapy compared with usual management, in patients with refractory IBS in primary care.

3. There is moderately good evidence to show a significant reduction in pain for hypnotherapy compared with attention control or usual management, in patients with refractory IBS, both in primary and secondary care.

4. There is limited evidence to show a significant reduction in bloating for hypnotherapy compared with attention control, in patients with refractory IBS, in secondary care.

5. There is moderately good evidence to show no significant improvement in diarrhoea or constipation or quality of life, after 12 weeks, for hypnotherapy compared with usual management, in patients with refractory IBS in primary care.

6. There is limited evidence to show a significant reduction over 12 months, in the number of prescriptions for other IBS medications, for hypnotherapy compared with usual management, in patients with refractory IBS in primary care.

7. There is limited evidence to show no significant difference between group and individual hypnotherapy, in patients with refractory IBS and psychological problems in secondary care.

HEALTH ECONOMIC STATEMENT

Evidence from a decision analytic model showed that the addition of hypnotherapy to usual care is cost-effective in individual with refractory IBS although the cost-effectiveness was sensitive to uncertainty around the proportion of patients experiencing an improvement in global symptom score with usual care alone. The ICER is sensitive to the proportion of patients experiencing an improvement in global symptom score with usual care alone and the efficacy in the 9 months after intervention. A threshold analysis showed that an improvement in global symptoms must result in a utility gain of at least 0.05 for the cost per QALY to remain below £20,000.

GDG DISCUSSION

The GDG's view was that hypnotherapy may be considered a developing intervention for IBS and the amount of evidence is limited. Despite this, the judgement and experience of GDG clinicians together with the limited RCT evidence from the review suggest that gut directed hypnotherapy strategies provide people with IBS with benefits in a cost-effective manner. Currently hypnotherapy is used as a second line therapy option, usually for people with unresolved IBS symptoms, who have failed to respond to a combination of management strategies. It features on the patient care pathway as one of the psychological interventions that primary care clinicians should consider if symptoms persist.

Although there is currently a lack of research in hypnotherapy, the GDG agreed there is potential for long-term benefits to the NHS from this behavioural therapy that need to be investigated further, including its use as a first line therapy. The GDG therefore decided to include hypnotherapy in one of its top five research recommendations, with the potential for this intervention to be considered as a first line therapy option.

EVIDENCE TO RECOMMENDATION

The evidence to recommendation statement for psychotherapy, CBT and hypnotherapy is detailed in section 9.8.

The combined guideline recommendation for psychotherapy, CBT and hypnotherapy is also stated in section 9.8.

9.7 Indirect comparison of psychological interventions

We have undertaken an indirect comparison to assess the relative cost-effectiveness of the psychological interventions (CBT, psychotherapy and hypnotherapy). It is indirect because it is based on the cost-effectiveness of each intervention compared to usual care as no trials were identified which compared behavioural interventions head-to-head. The results are presented for two scenarios. In the first scenario the basecase assumptions are maintained from the analysis of each intervention compared to usual care. In the second scenario the basecase assumptions are maintained except that the resource use reduction from the Creed (2003) study has been excluded from the cost-effectiveness estimate for psychotherapy. This has been done because the GDG felt that there was a lack of similar evidence for CBT and hypnotherapy but that this was due to a lack of trials reporting economic outcomes for these interventions rather than a true difference in the cost-effectiveness compared to psychotherapy.

Hypnotherapy provided the smallest QALY gain compared to usual care but is likely to be cost-effective compared to usual care as discussed in section 9.6. As CBT provided more QALY gain than hypnotherapy at additional cost, we have considered the incremental cost-effectiveness of CBT compared to hypnotherapy. The CEAC in Figure 1 shows that CBT has a 52% probability

of being cost-effective compared to hypnotherapy at a cost per QALY threshold of £20,000 and a 76% probability of at a threshold of £30,000. The mean cost per QALY for CBT compared to hypnotherapy under the basecase assumptions was £18,158 for the deterministic model. There was concern that this comparison had been biased by the use of different unit costs for therapy sessions for CBT and hypnotherapy so we carried out a sensitivity analysis using the unit costs for CBT for both psychological interventions. This gave a cost per QALY of £15,301.

Figure 1: CEAC for CBT vs hypnotherapy
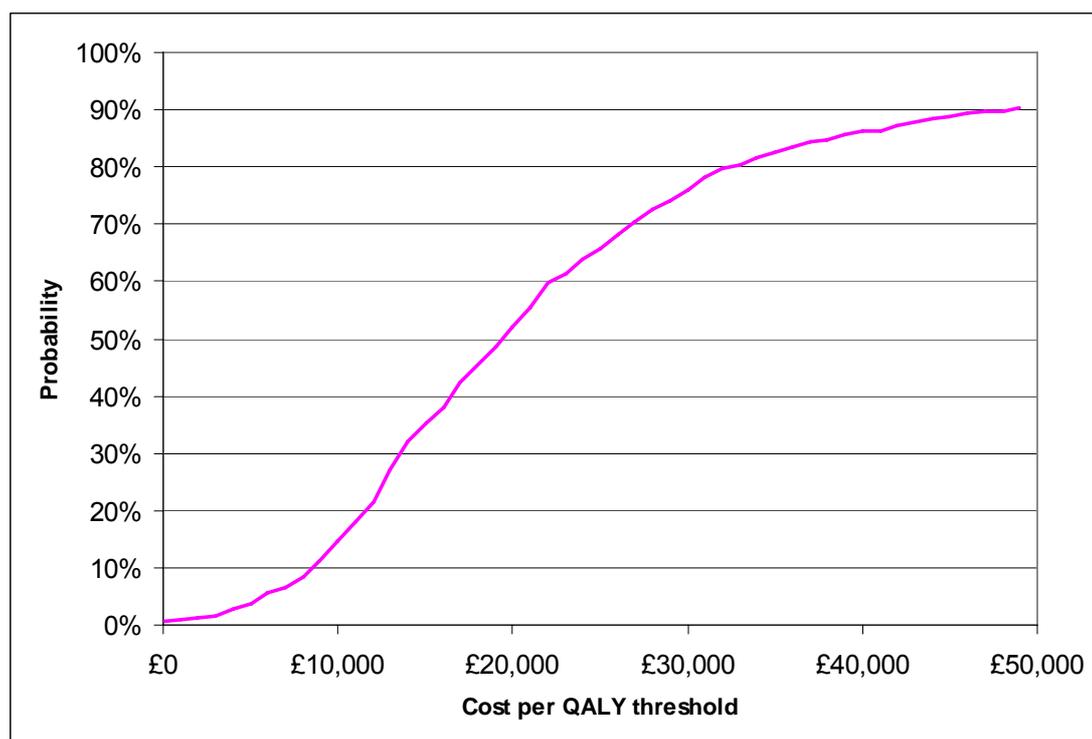
Under the basecase assumptions, psychotherapy provided additional QALY gain compared to CBT but the mean cost for psychotherapy was less than for CBT. The CEAC in Figure 2 shows the incremental cost-effectiveness of psychotherapy compared to CBT. Psychotherapy has a 68% probability of providing additional QALY gain at no additional cost compared to CBT and a 73% probability of providing additional QALY gain for less than £20,000 per QALY. There was concern that this comparison had been biased by the use of different unit costs for therapy sessions for CBT and psychotherapy so we carried out a sensitivity analysis using the unit costs for CBT for both psychological interventions. This raised the intervention cost for psychotherapy, but it still had a lower overall cost than CBT.

The lower cost of psychotherapy is driven by the assumption on lower resource use for psychotherapy compared to usual care. When this factor was excluded from the analysis psychotherapy had a mean cost per QALY of £11,314 compared to hypnotherapy with a 61% probability of a being under £20,000 and a 70% probability of being under £30,000 per QALY.

These results suggest that each of the interventions would result in the cost-effective use of NHS resources but it does not address which is the most cost-effective.

Figure 2: CEAC for psychotherapy compared to CBT

Figure 3 is a multi-way cost-effectiveness acceptability curve which shows the probability that each of the three psychological interventions is optimal compared to the other two at various cost per QALY thresholds. The optimal intervention is the one that provides the most QALY gain at a cost per QALY under the threshold. This is most easily described by considering the incremental net benefit of each intervention, which is the (monetary) value of a strategy compared with an alternative strategy for a given cost-effectiveness threshold. For example, if society is willing to pay £20,000 for an additional QALY then the incremental NB is:

Net benefit = (Additional QALY gain x £20,000) – additional cost

The strategy with the greatest incremental net benefit compared to usual care, at a given cost per QALY threshold, is the optimal strategy at that threshold. We used the results of the probabilistic sensitivity analysis to estimate the probability that each behavioural therapy is optimal at various cost per QALY thresholds. Figure 3 shows that under the basecase assumptions, psychotherapy has the highest probability of being the optimal strategy at willingness to pay thresholds of £10,000 to £50,000.

Figure 4 shows that when a similar exercise is carried out for the second scenario, in which we assumed that none of the three interventions result in reduced resource use, psychotherapy had a lower probability of being the optimal strategy, but it is still the most likely to be optimal for willingness to pay thresholds of £20,000 to £50,000 per QALY. Hypnotherapy has the highest probability of being the optimal strategy for cost per QALY thresholds under £20,000.

These results suggest that providing psychotherapy for people with refractory IBS is likely to result in more efficient use of NHS resources than providing CBT or hypnotherapy. However, the analysis did not take into account factors that may be important in deciding the optimal treatment for an individual. For example, if the effectiveness of these behavioural interventions is higher in patients who are committed to a particular intervention then choosing to provide the intervention preferred by the patient may result in treatment being more cost-effective. The results of the cost-effectiveness modelling suggest that all three behavioural interventions would result in the cost-effective use of NHS resources at willingness to pay thresholds of £20,000 to £30,000, given the evidence currently available.

HEALTH ECONOMIC STATEMENT

A decision analytic model was used to carry out an incremental analysis for the three psychological interventions. This was an indirect comparison based on the effectiveness of each behavioural therapy compared to usual care and therefore may be biased. Psychotherapy was the most cost-effective intervention when considering a cost per QALY threshold in the £20,000 to £30,000 range. This conclusion was not dependent on whether the reduction in resource use for psychotherapy compared to usual care, as observed in Creed (2003), was included in the analysis.

Figure 3: Multi-way CEAC for psychological interventions under the basecase assumptions
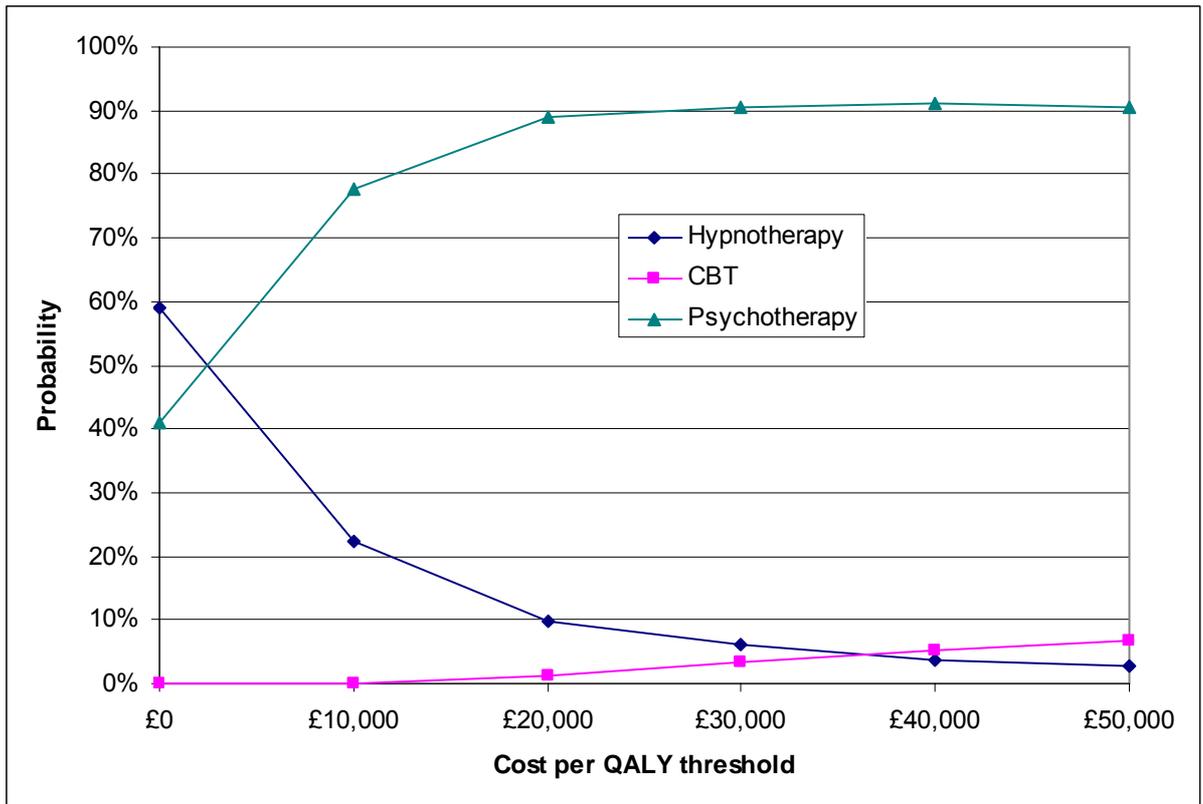
Figure 4: Multi-way CEAC for psychological interventions when we assumed that there is no resource use reduction associated with any of the therapies
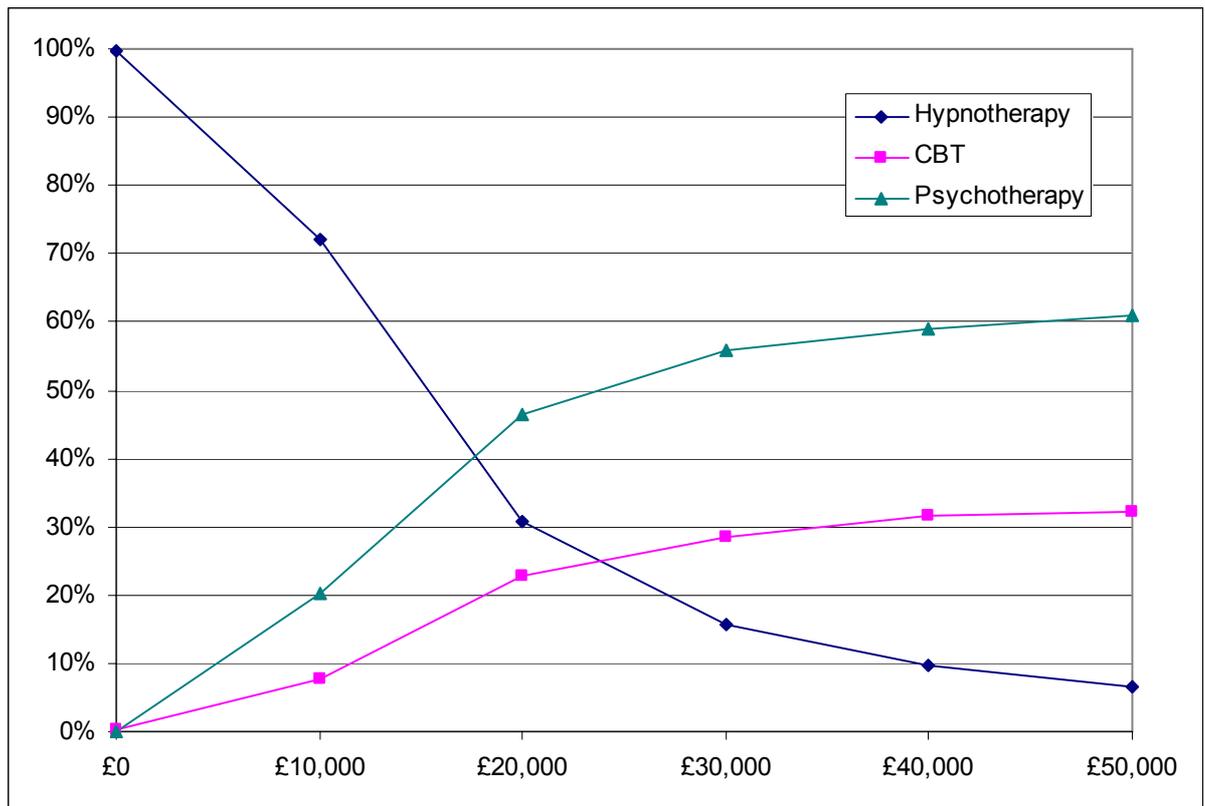
9.8 Evidence to recommendation: psychotherapy, CBT and hypnotherapy

The GDG considered CBT, hypnotherapy and psychotherapy, as a group of similar, but distinct therapies when making recommendations, and took into account several factors:

Firstly, they considered the clinical effectiveness reviews and cost effectiveness modelling that have been carried out mainly for the treatment of people with refractory IBS. The GDG interpreted the cost effectiveness analyses, including the indirect comparisons between the three therapies. The GDG noted that the trials were mainly in people with refractory IBS, and, for this group, the therapies were all cost effective.

Secondly, the GDG highlighted the current national variation relating to where these therapies are accessed, and noted that this is dependent on the commissioning patterns of individual strategic health authorities. Typically, they are more available in secondary care.

Thirdly, the GDG took into consideration the need to give people with IBS and their primary care clinician a choice in which behavioural therapy was most appropriate for them, and what might be available locally.

On balance, the GDG decided not to distinguish between the three therapies, and recommended that any one of them should be considered for people who have had IBS for at least 12 months, and who have not responded to first line therapies and whose symptoms continued. This patient profile has been defined for the purpose of this guideline as refractory IBS.

The GDG discussed whether there was an optimum time for treatment with any of these psychological therapies: leaving patients too long may have meant the person was no longer able to respond. In addition, the GDG was keen to determine whether these therapies could be used as first line treatments, as they had potential to enable people with IBS to cope with their symptoms by giving initial treatments which would have long term sustainability. This view was supported by evidence in children with IBS, which showed that hypnotherapy is clinically effective as a first line therapy. The GDG therefore proposed a recommendation for research to compare, head-to-head, the three therapies as first line therapies, with follow-up at various time points up to a year.

During GDG discussion relating to psychological therapies, it was recognised that it would be very useful for clinicians to be able to predict which people would have refractory IBS and which factors put them at risk. Therefore a second research recommendation was proposed to investigate what factors are important. These research recommendations are given in chapter 12.

RECOMMENDATION

Referral for psychological interventions (cognitive behavioural therapy [CBT], hypnotherapy and/or psychological therapy) should be considered for people with IBS who do not respond to pharmacological treatments after 12 months and who develop a continuing symptom profile (described as refractory IBS).

10 COMPLEMENTARY AND ALTERNATIVE THERAPIES

Clinical Questions

1. Is acupuncture an effective intervention in managing IBS symptoms?

2. Is reflexology an effective intervention in managing IBS symptoms?

3. Is herbal medicine an effective intervention in managing IBS symptoms?

BACKGROUND

Complementary and alternative medicine (CAM) may be defined 'as wide ranging therapies which may be used exclusively i.e. complete healing systems, or in combination with orthodox medical treatment' (House of Lords Select Committee on Science and Technology 2000 p.2). The terms 'Alternative' and 'Complementary' are used to define the use and setting of a therapy in relation to orthodox medicine. 'Alternative' usually refers to treatment modalities that are generally a substitute for orthodox treatment whereas 'complementary' refers to treatments that are used alongside orthodox medical treatments. CAM is usually considered to include the practice of therapies that are not considered integral to the dominant health care model of a country, society or culture.

The House of Lords Select Committee on Science and Technology Sixth Report addressed the future of CAM in relation to research, service delivery, education and training and regulation. The report stated that there is very little evidence about the efficacy of many complementary and alternative treatments but the use of CAM is widespread and is increasing across the developed world. There is a clear need for more effective guidance for the public and health professionals who advise patients as to what does and does not work and what is and is not safe.

In order to begin to establish the effectiveness of CAM it is important to identify specific therapies and particular conditions where the use of CAM may be appropriate. It is not uncommon for those suffering from chronic conditions, for whom conventional medicine has been less than successful in alleviating symptoms, to seek complementary and alternative medicine (CAM). Irritable Bowel Syndrome is an example of such a condition.

The guideline considered commonly used therapies, Acupuncture, Chinese Herbal Medicine, Homeopathy and Reflexology. Hypnotherapy was considered with the psychological interventions.

Homeopathy

Homeopathy is defined as a system of therapy based on the concept that disease can be treated with drugs (in minute doses) thought capable of producing the same symptoms in

healthy people as the disease itself. Homeopathy was invented by the German physician Samuel Hahnemann (1755-1843) in the late 18th and early 19th centuries. It was both refined and popularized by the American physician James Tyler Kent. Homeopathy is based on the theory that each naturally occurring element, plant, and mineral compound will, when ingested or applied, result in certain symptoms. Hahnemann believed that, by diluting these substances in a standardized manner, one could reach the true essence of that substance. Hahnemann described this process of dilution as "potentizing" (German: "potenziert") the substance. These dilute amounts could then be used to treat the very symptoms they were known to produce.

An initial search idenitified two trials using homeopathy for IBS, both conducted about 30 years ago and reported in German.No trials have been done since. Only randomised trials were to be considered for this review and the absence of further studies suggested no need to carry out a full review.

Acupuncture

Acupuncture is defined as a therapeutic and/or preventive medical procedure used in or adapted from Traditional Chinese Medicine (TCM) performed by the insertion of 1 or more specially manufactured solid metallic needle(s) into specific location(s) on the body. The intent is to stimulate acupuncture points, with or without subsequent manual manipulation. The acupuncture points are situated on fourteen major 'meridians'. The TCM theory is that acupuncture stimulates 'qi' (translated as life force) that circulates through the meridians. In optimum health the flow of 'qi' is unobstructed. Interruption or stagnation of the flow of 'qi' results in diverse symptoms. The theory is that insertion and manipulation of needles at particular points stimulates the energy flow, restoring the balance and thus normalising the function of the organ. An alternative theory is that acupuncture is a specialised sensory stimulation that is analysed through sensory neural pathways.

Reflexology

Reflexology is an ancient form of complementary medicine thought to originate in China, however research has shown that reflexology was also used by some early African tribes, Native American Indians and early Egyptians. Reflexology is a complementary therapy based on the theory that by the application of pressure to specific reflex points on the feet and hands, which correspond to the organs of the body, it is possible to 'normalise' function. In conventional medical terms reflexology could be said to facilitate homeostasis. Reflexology is a widely used therapy; it has been estimated that between 6 and 12% of the population use it and anecdotal evidence suggests that many people find it extremely effective for a range of chronic conditions including functional bowel conditions, although there is little rigorous research to support this view.

People with IBS may be drawn to acupuncture and reflexology's ancient roots and the desire for non-pharmacological treatment. Alterations in pain modulation, motility, and autonomic nervous

system function are likely mechanisms of IBS symptoms, which may have physiological responses to acupuncture and reflexology.

Herbal Medicine

Emerging evidence for herbal medicine (Chinese and non Chinese) suggesting a possible benefit informed one of the guideline's research recommendations.

People with IBS are interested in CAM and will continue to use these modalities as long as medical therapy fails to relieve their symptoms. To optimise the care of people with IBS there is a need for further evidence of the potential benefits and safety of these treatments. Integration of CAM into Western medical practice will require more than selection of a few isolated acupoints, or yoga positions. A wider understanding of the paradigm specific use of these techniques, mechanisms of action, and potential pitfalls is required.

10.1 Reflexology

SELECTION CRITERIA

The selection criteria described in the general methodology section were used. Interventions were any form of reflexology.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the AMED database was searched for this review. Search strategies are given in Appendix B.

The search strategy identified 560 studies. The titles and abstratcs of these studies were assessed. One was identified to be potentially relevant to this review and this paper was retrieved. The reference lists of these studies were inspected for further potential papers, but none were identified.

DESCRIPTION OF STUDIES

Types of Studies

Only one study was identified: a quasi randomised trial in a UK primary care setting (Tovey 2002).

Types of Participant

Thirty four patients were allocated treatments. The groups were comparable at baseline as regards age; gender; duration, and; severity of condition.

Inclusion criteria were that patients had to have a diagnosis of IBS in line with Rome II criteria, and they should be currently under the care of a primary care physician following referral to a

gastroenterologist to exclude organic GI disease. Patients were excluded if they had organic GI disease or had previously used reflexology.

Intervention

Six 30-minute sessions of reflexology delivered in a way that was as close as possible to normal practice conditions, over an eight week period. The control group received six 30-minute sessions of foot massage that excluded pressure on key points of the foot.

METHODOLOGICAL QUALITY

Sequence generation was by alternation, and the allocation concealment was inadequate. A power calculation was carried out and the sample size required was 18 patients per group for the outcome of abdominal pain. 4/19 (21%) in the reflexology group were lost to follow-up and 2/15 in the control group.

The study reported no significant differences in baseline characteristics (pain; diarrhoea; constipation; bloating).

RESULTS

Individual symptoms of IBS were recorded daily using a 5 point scale, but global symptoms were not reported.

a) Pain

Pain was the primary outcome measure. There was no significant difference between the reflexology and control groups for this outcome, either at assessment 2 weeks after completion of the intervention (p=0.32) or at 3 month follow-up.

Table 1.

Reflexology Baseline **Pain score (0-4 scale)**	Change from baseline: reflexology post treatment	Control Baseline	Change from baseline: control post treatment
Median:1.4 IQR: 0.6 to 2.1	End of treatment: Median: -0.10 (IQR: -0.80 to 0.10) 3 months follow up: Median 0.00	Median:0.7 IQR:0.5 to 1.3	Median:-0.40 (IQR: -0.90 to 0.00) 3 months follow up: Median -0.25

b) Bowel function

There was no significant difference, 2 weeks after completion of the intervention, in bowel function (change in constipation or diarrhoea) (p=0.47) between intervention and control groups.

Table 2.

Reflexology Baseline **Bowel Function** **(scale 0-4)**	Change from baseline: reflexology post treatment	Control Baseline	Change from baseline: control post treatment
Median: 1.9 IQR: 1.2 to 2.1	Median: 0.05 IQR: -0.53 to 0.43	Median: 1.2 IQR: 0.3 to 1.7	Median:-0.30 IQR: -0.80 to 0.20

c) Bloating

There was also no significant difference, 2 weeks after completion of the intervention, in bloating (p=0.0.17) between intervention and control groups.

Table 3.

Reflexology Baseline **Bloating** **(scale 0-4)**	Change from baseline: reflexology post treatment	Control Baseline	Change from baseline: control post treatment
Median: 2.5 IQR: 1.3 to 3.1	Median: -0.10 IQR: -0.60 to 0.20	Median:2.0 IQR: 1.0 to 2.2	Median: -0.40 IQR: -1.05 to -0.15

HEALTH ECONOMIC EVIDENCE

The cost effectiveness of reflexology was not taken into consideration for this review because reflexology is not prescribed with treatment being purchased independently by people with IBS.

EVIDENCE STATEMENT

There is limited evidence from a single study in people with IBS in Primary Care showing no significant effect on pain, bowel function and bloating compared with the foot massage placebo group.

GDG CONSENSUS

The GDG was concerned that the foot massage group may not have been reliable as a placebo group. The limited evidence from this small, quasi-randomised trial does not lend support to the use of reflexology in the management of IBS in adults. However, there may be a need for further research.

EVIDENCE TO RECOMMENDATION

The review reported limited evidence that showed reflexology is not effective in the management of IBS symptoms. The GDG's clinical view was that the current lack of effectiveness precludes a positive recommendation.

RECOMMENDATION

The use of reflexology should not be encouraged for the treatment of IBS.

10.2 Acupuncture

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some were specific to the acupuncture review and are reported below.

Types of studies

Crossover trials could be included, but those with a washout period of less than 2 weeks were to be excluded. All study designs were included for adverse effects. Specific searches for adverse effects were not carried out.

Types of intervention

Studies to be considered for inclusion included the following interventions:

- Single acupuncture needling point
- Combination acupuncture needling points.

Methods of acupuncture that do not involve needle insertion for example laser or acupressure were to be excluded. For the purposes of this review the minimum acceptable dose was to be two treatments of acupuncture. Studies that included a single acupuncture treatment were to be excluded.

Types of comparisons

The following comparisons were to be included:

- Single acupuncture versus sham acupuncture (placebo)
- Combination acupuncture versus sham acupuncture (placebo)
- Single acupuncture versus another type of treatment
- Combination acupuncture versus another type of treatment
- Acupuncture + treatment 2 versus treatment 2

Subgroup analyses

Subgroup analyses were to be carried out if there is heterogeneity as follows:

- Symptom severity
- Dose
- Type of acupuncture.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the AMED database was searched for this review. The search strategies are listed in Appendix B. The search strategy identified 764 studies. The titles and abstracts of these studies were assessed. Of these studies, 20 were identified on the basis of the title and abstract as being

potentially relevant to the review and these papers were retrieved in full. All reference lists of these studies were inspected for potential papers for inclusion in the review, but no further potential studies were found in addition to the titles already identified. Nineteen studies and one Cochrane review were identified (Manheimer 2006). Of these, eight were excluded and these are listed in Appendix E, along with reasons for exclusion.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

Twelve studies met the inclusion criteria for the review (Chung 2003; Burford-Mason 2003; White 2001; Ernst and White 1997; Ernst and White 2001; MacPherson 2001; Yamashita 2001; Schneider 2006; Lowe 2000; Forbes 2005; Conboy 2006; Fireman 2001). Five studies (Schneider 2006; Lowe 2000; Forbes 2005; Conboy 2006; Fireman 2001) investigated the effectiveness of acupuncture for the treatment of IBS. One was conducted in the UK (Forbes 2005) and one each in Germany, Canada and Israel. Seven studies investigated adverse effects (Chung 2003; Burford-Mason 2003; White 2001; Ernst and White 1997; Ernst and White 2001; MacPherson 2001; Yamashita 2001).

Study Design

All the studies in the review were parallel studies, with the exception of Fireman (2001) which was a crossover study. The latter had a three week washout period, but first period results were also reported, which were used in preference because of the uncertainty about carry-over effects. One study (Lowe 2000) was only reported as a conference abstract.

All the studies took place in secondary care. The studies investigating adverse effects included medical doctors in primary and secondary care and non-medical acupuncturists.

The Cochrane Review 'Acupuncture for treatment of irritable bowel syndrome' (Manheimer 2006) included six trials. Two were excluded from the guideline review: one used electrical ear acupuncture (Liu 1995) and the other was a study in Chinese (Liao 2000). The Cochrane review authors stated that there was a possibility that Liao (2000) was not an RCT.

Population

All studies included people with a diagnosis of IBS, although the definition varied. Three used the Rome I criteria (Fireman 2001; Forbes 2005; Lowe 2000) and two used the Rome II criteria (Schneider 2006; Conboy 2006). All studies included a combination of IBS types and none of the studies stated that any participants had IBS as result of gastrointestinal infection. All studies included some participants with bloating. One study (Schneider 2006) identified all patients as having bloating, and in another (Fireman 2001) 80% had bloating.

All of the studies described symptom severity as mixed. The age range of participants was 17 to 79 years with the average mean age being approximately 46 years. No study particularly identified elderly participants. All studies had more women than men.

The Forbes (2005) study only included people who were refractory to other treatments; Fireman (2001) had participants who had had clinical symptoms for at least a year.

The numbers of participants ranged from 25 to over 100 (Conboy 2006; Liu 1997).

Interventions

The included studies all used different acupuncture protocols, but all used Chinese style acupuncture. One study used a single acupuncture point (Fireman 2001) and the remainder used a combination of points (Lowe 2000; Forbes 2005; Conboy 2006; Schneider 2006). The number of sessions of acupuncture varied from two (Fireman 2001) to ten (Forbes 2005; Schneider 2006).

Comparisons

The majority of studies compared true acupuncture with sham acupuncture. One study compared acupuncture plus psychotherapy with acupuncture alone and psychotherapy alone (Liu 1997).

The sham acupuncture varied between studies:

- Multiple needling versus sham needling at non acupuncture points (Schneider 2006)
- Single needling versus sham needling at an inappropriate acupuncture point (Fireman 2001)
- Multiple needling versus sham needling at inappropriate acupuncture points (Forbes 2005)
- Multiple needling versus non-needling at the same acupuncture points (Lowe 2000; Conboy 2005).

Two studies used a validated sham needling device (Conboy 2005; Schneider 2006).

Outcomes

The studies measured a range of outcomes using different scales.

1. Global score

a) Number of people with global improvement of symptoms

The Forbes (2005) study reported the number of people who recorded a reduction in symptom score of four points, which constituted an improvement. The Lowe (2001) study recorded a patient-determined success rate, which was based on individual patient expectations stated at baseline.

b) Global improvement of symptoms score

Fireman (2001) used a visual analogue scale (1 to 5), on which 5 equated to significant improvement in global symptoms. Liu (1997) used a three point scale, 1 = cured, 2 = improved, 3 = no effect.

c) Global symptom score

Forbes (2005) used a global symptom score with a scale of 0 to 30 based on symptom diaries plus the Bristol Stool Form Scale. A reduction of 4 points was considered clinically significant.

2. Individual symptoms

a) Pain

Three studies reported a pain score (Fireman 2001, Lowe 2000, Scheider 2006). In all cases the highest rating meant worst symptoms, although the scales used were not the same. The Lowe (2000) study only gave p-values for the pre-post comparison for each group and the Scheider (2006) study recorded scores on the pain subscale of SF36.

b) Bloating

One study reported bloating as an individual symptom (Fireman 2001).

c) Bowel habits

No studies reported bowel habit as an individual symptom in all patients, although Fireman (2001) reported diarrhoea scores for 11 patients with diarrhoea and defaecation difficulty scores in 13 people with constipation. We decided that these small subgroups broke the randomisation and were likely to give unreliable results.

3. Mental health

One study (Forbes 2005) assessed participants using the Hospital Anxiety and Depression (HAD) scale and recorded the change score.

4. Quality of Life

One study (Forbes 2005) assessed participants using using the EuroQol quality of life questionnaire.

Schneider (2006) used the FDDQL (scale 0 to 100), which assesses the disease related impact of bowel symptoms on quality of life; and the SF36 health-related quality of life measure. The primary outcome of the study was improvement in the global score of the FDDQL i.e. a reduction in score after 10 sessions. Lowe (2001) used the validated quality of life tool, IBS-36, but only reported p-values for changes from baseline.

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in Appendix D.

The method of randomisation was reported in three studies, all of which were classified as adequate (computer generated: Forbes 2005; Schneider 2006; Conboy 2006). The other studies did not state the method of randomisation (Fireman 2001; Lowe 2000).

Allocation concealment was reported in two studies (Forbes 2005; Schneider 2006). The Schneider (2006) study had adequate concealment (sequence retained by a central telephone centre) and the Forbes (2005) had partial concealment (sealed envelopes).

Three studies reported that the outcome assessors and the patients were blinded to the interventions (Fireman 2001; Forbes 2005; Schneider 2006). It was unclear whether the patients were blinded in Lowe (2000).

Most studies described the details of the placebo and active intervention giving the location of acupuncture points used. Lowe (2000) was the exception.
Four studies (Lowe 2000; Forbes 2005; Schneider 2006; Conboy 2006) described an *a-priori* power calculation. Two studies used an intention to treat analysis (Schneider 2006; Forbes 2005). Most studies included in the review demonstrated some level of baseline comparability of the groups, but one provided no data regarding baseline characteristics (Lowe 2000). The number of people who withdrew from the studies or were lost to follow-up was minimal. None of the studies were considered to be at high risk of bias.

RESULTS

1. Global symptoms

a) Number of people with global improvement of symptoms

Two studies recorded the number of people with an improvement in global symptoms (Lowe 2000, Forbes 2005). These two studies were combined in a meta-analysis of 109 participants, even though the studies used different types of sham acupuncture. There was no statistically significant difference between acupuncture and sham acupuncture.

Figure 1: Number of people with global improvement of symptoms

Review: IBS 31 August 2006
Comparison: 27 Acupuncture vs sham acupuncture
Outcome: 01 Global improvement of IBS symptoms

Study or sub-category	Acupuncture n/N	Sham acupuncture n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Multiple point acupuncture vs non needling at same points					
Lowe 2000	16/28	10/22		55.03	1.26 [0.72, 2.20]
Subtotal (95% CI)	28	22		55.03	1.26 [0.72, 2.20]
Total events: 16 (Acupuncture), 10 (Sham acupuncture)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.80 (P = 0.42)					
02 Multiple point acupuncture vs sham needling at inappropriate points					
Forbes 2005	11/27	10/32		44.97	1.30 [0.66, 2.59]
Subtotal (95% CI)	27	32		44.97	1.30 [0.66, 2.59]
Total events: 11 (Acupuncture), 10 (Sham acupuncture)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.76 (P = 0.45)					
Total (95% CI)	55	54		100.00	1.28 [0.83, 1.98]
Total events: 27 (Acupuncture), 20 (Sham acupuncture)					
Test for heterogeneity: Chi² = 0.01, df = 1 (P = 0.94), I² = 0%					
Test for overall effect: Z = 1.10 (P = 0.27)					

0.1 0.2 0.5 1 2 5 10

Favours sham acupi   Favours acupuncture

b) Global improvement of symptoms score

Fireman (2001) recorded the global improvement in symptoms score (based on symptoms of pain; defaecation difficulties; diarrhoea; alternating diarrhoea and constipation; bloating; abdominal discomfort relieved by defaecation, and; mucus in stools), in 25 patients, using a visual analogue scale (1 to 5), on which 5 equated to significant improvement in global symptoms. As this study was a crossover design, data were used from the first period only. There was no significant difference between interventions.

Figure 2: Global improvement of symptoms score

Review: Acupuncture
Comparison: 02 Single point Acupuncture vs sham at inappropriate AC points
Outcome: 01 Global Improvement in Symptom Score - 1st treatment period

Study or sub-category	Acupuncture N	Mean (SD)	Sham Acupuncture N	Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Fireman et al 2001	13	4.10 (0.86)	12	3.62 (0.84)		100.00	0.48 [-0.19, 1.15]
Total (95% CI)	13		12			100.00	0.48 [-0.19, 1.15]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.41 (P = 0.16)							

-4 -2 0 2 4

Favours Sham Acupunct   Favours Acupuncture

c) Global symptom score

One study (Forbes 2005) recorded the global symptom score on a scale of 0 to 30. There was no significant difference between acupuncture and sham acupuncture.

Figure 3: Global symptom score


```
Review:        IBS 31 August 2006
Comparison:    27 Acupuncture vs sham acupuncture
Outcome:       02 Global symptom score

Study                          Acupuncture          Sham acupuncture        WMD (fixed)      Weight      WMD (fixed)
or sub-category        N       Mean (SD)       N     Mean (SD)               95% CI           %           95% CI

01 Multiple point acupuncture vs sham needling at inappropriate points
  Forbes 2005          27      11.60(5.13)     32    11.20(4.17)                 ■            100.00      0.40 [-2.01, 2.81]
Subtotal (95% CI)      27                      32                               ◆            100.00      0.40 [-2.01, 2.81]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.32 (P = 0.75)

Total (95% CI)         27                      32                               ◆            100.00      0.40 [-2.01, 2.81]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.32 (P = 0.75)

                                                              -10    -5     0     5    10
                                                         Favours acupuncture  Favours sham acupunc
```

2. Individual symptoms

a) Pain

Two studies recorded pain scores, Fireman (2001) and Scheider (2006). The latter used the discomfort subscale of SF36. The studies differed in the type of acupuncture used (single versus multiple point, respectively), and therefore were not combined in a meta-analysis. There was no significant difference between interventions in either study, although the sham acupuncture is favoured in the Fireman (2001) study.

Figure 4: Pain score


```
Review:        Acupuncture
Comparison:    02 Single point Acupuncture vs sham at inappropriate AC points
Outcome:       02 Pain Score - 1st treatment period

Study                          Acupuncture          Sham Acupuncture        WMD (fixed)      Weight      WMD (fixed)
or sub-category        N       Mean (SD)       N     Mean (SD)               95% CI           %           95% CI

  Fireman et al 2001   12      4.20(0.90)      11    3.50(0.96)                  ■           100.00      0.70 [-0.06, 1.46]

Total (95% CI)         12                      11                              ◆            100.00      0.70 [-0.06, 1.46]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.80 (P = 0.07)

                                                              -4    -2     0     2    4
                                                         Favours Acupuncture  Favours Sham
```
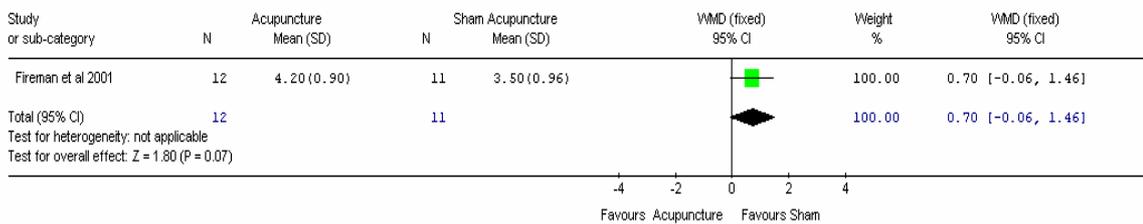
Figure 5: Oain component of SF 36


```
Review:        IBS 31 August 2006
Comparison:    27 Acupuncture vs sham acupuncture
Outcome:       04 Pain score

Study                          Acupuncture          Sham acupuncture        WMD (fixed)      Weight      WMD (fixed)
or sub-category        N       Mean (SD)       N     Mean (SD)               95% CI           %           95% CI

02 Multiple needling versus sham needling at non acupuncture points
  Schneider 2006       22      44.86(22.70)    21    41.62(17.03)               ■            100.00      3.24 [-8.72, 15.20]
Subtotal (95% CI)      22                      21                              ◆            100.00      3.24 [-8.72, 15.20]
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.53 (P = 0.60)

                                                          -100   -50    0    50   100
                                                         Favours acupuncture  Favours sham acupunc
```

3. Bloating

Fireman (2001) reported a bloating score on a VAS of 1 to 5 in 20 participants. There was no significant difference between acupuncture and sham acupuncture.

Figure 6: Bloating score


```
Review:      Acupuncture
Comparison:  02 Single point Acupuncture vs sham at inappropriate AC points
Outcome:     03 Bloating score- 1st treatment period
```

Study or sub-category	N	Acupuncture Mean (SD)	N	Sham Acupuncture Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Fireman et al 2001	10	3.70 (0.93)	10	3.50 (0.87)		100.00	0.20 [-0.59, 0.99]
Total (95% CI)	10		10			100.00	0.20 [-0.59, 0.99]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.50 (P = 0.62)

Favours Sham Acpunct   Favours Acupuncture

4. Quality of life

Three studies reported quality of life measurements (Forbes 2005, Lowe 2000 and Schneider 2006). Forbes 2005 reported a small improvement in the EuroQol scores over baseline in both the acupuncture (59.4 to 64.6%) and sham acupuncture (64.6 to 65.6%) groups, neither difference was significant. Lowe (2000) reported a marked improvement in the IBS-36 quality of life score in both true and sham groups. There was no significant difference between the two groups. Schneider reported a mean difference of 1.98 (95%CI -3.59, 7.39) in 43 people, at the end of treatment, on a scale of 0 to 100, i.e. no significant difference. After three months follow-up, there was still no significant difference (MD 3.41 (95%CI -3.02, 9.83)

5. Adverse effects

The benefit of acupuncture cannot be evaluated without considering the risks associated with treatment. The incidence of adverse effects is largely unknown. However, for the purposes of this review, we included seven studies investigating adverse effects (Chung 2003; Burford-Mason 2003; White 2001; Ernst and White 1997; Ernst and White 2001; MacPherson 2001; Yamashita 2001). Three of these were systematic reviews (Ernst and White 1997; Ernst and White 2001; Yamashita 2001), two were surveys of acupuncture practice and one a commentary. The systematic reviews identified ten reports which included surveys from Europe and eighty-nine reports from the Far East. The most common adverse events identified in Europe were:

- Pain at the site of needling
- Pain due to aggravation of the presenting condition
- Bleeding – petechia, ecchymosis, haematoma
- Nausea and/or vomiting
- Fainting
- Tiredness.

Potentially serious adverse effects are rare: two cases of pneumothorax and two cases of needle fracture requiring surgical removal of the fragment, and one case of burn injury following moxibustion. There were no reports of infection complications or transmission of disease through needling.

The review from the Far East (Yamashita 2001) synthesised 89 papers that reported 124 cases of adverse events. These were classified into five categories:

- Injuries or foreign bodies (42 cases)
- Infections (32 cases, including 11 cases of Hepatitis B)
- Neurological problems (29 cases, including 18 cases of spinal cord injury, 10 of which were caused by needle breakage)
- Dermatological problems (17 cases)
- Other (4 cases).

The reviewers had previously demonstrated that severe adverse effects seem to be uncommon in standard practice for adequately trained acupuncturists.

The two Ernst and White, European reviews also found that there was no standard definition of adverse effects and estimated that there may be under-reporting of adverse events. It is also possible that there is over reporting of adverse effects so that the true incidence of serious complications may be very low. They emphasise the need to ensure appropriate training standards and appropriate regulatory and surveillance systems to enable more accurate assessment.

HEALTH ECONOMIC EVIDENCE

The cost effectiveness of acupuncture was not taken into consideration for this review because acupuncture is not prescribed, with the majority of acupuncture treatment being purchased independently by people with IBS.

EVIDENCE STATEMENTS

1. There is fair evidence to show no significant effect of acupuncture on IBS global symptoms, pain, and quality of life compared with placebo.

2. There is limited evidence of potentially serious adverse effects associated with acupuncture treatments.

GDG DISCUSSION

The GDG was concerned about the reported adverse effects (some of which were severe), non-registration and the safety of acupuncture. They noted an additional adverse effect that occurs with moxibustion, which can lead to burns. Members of the GDG were not surprised that acupuncture has been shown to have no significant effect in IBS: this might be expected because acupuncture is thought to work by producing endorphins which give pain relief, but they have no effect on visceral pain.It was noted that the patient community widely supports Traditional Chinese Medicine acupuncture.

The GDG's clinical view was that, although people with IBS widely support the use of acupuncture, the current lack of effectiveness and potential harm precludes a positive recommendation.

EVIDENCE TO RECOMMENDATION

The GDG took into consideration the lack of effectiveness of acupuncture, the limited evidence showing harm, and registration and regulation difficulties, and decided they would not recommend the use of acupuncture for IBS.

RECOMMENDATION

The use of acupuncture should not be encouraged for the treatment of IBS.

10.3 Herbal Medicine

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, but some specific to the herbal medicine review are reported below.

Types of studies

The GDG decided that crossover studies should not be included in this review because it was unclear whether herbal medicines effected longer term changes or how long they were retained in the gut.

Types of intervention

Studies were to include the following classes of interventions:

- Single Chinese herbal medicines
- Combination Chinese herbal medicines
- Single non-Chinese herbal medicines
- Combination non-Chinese herbal medicines.

The following comparisons were to be included:

- Single Chinese herbal medicines versus placebo
- Combination Chinese herbal medicines versus placebo
- Single non-Chinese herbal medicines versus placebo
- Combination non-Chinese herbal medicines versus placebo
- Single Chinese herbal medicines versus Combination Chinese herbal medicines
- Single non-Chinese herbal medicines versus Combination non-Chinese herbal medicines
- Herbal medicine type 1 versus type 2.

The review was concerned only with longer-term maintenance treatment.

Subgroup analyses

We planned to carry out subgroup analyses as follows:

- Sub-types of IBS (diarrhoea-predominant, constipation-predominant and alternating)
- Dose of herbal medicine.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL and *The Cochrane Library* (1966 to current day with guidance from the GDG). The MEDLINE search strategy is given in Appendix B.

Six studies met the inclusion criteria for the review. The reference lists of these were inspected for further potential papers, but none were identified.

DESCRIPTION OF STUDIES INCLUDED IN THE REVIEW

Six studies met the inclusion criteria for the review (Bensoussan 1998; Brinkhaus 2005; Leung 2006; Madisch 2004; Wang 2006; Yadav 1989). Two studies were conducted in China (Leung 2006; Wang 2006), two in Germany (Brinkhaus 2005; Madisch 2004) and one each in Australia (Bensoussan 1998) and India (Yadav 1989).

The total sample sizes ranged from 60 to 208, with all but one study (Wang 2006) having more than 100 participants.

Study Design

Setting: The setting was not stated in majority of studies; one was in primary care (Madisch 2004) and one took place in secondary care (Wang 2006).

All the studies included in the review had a parallel design.

Three studies had two arms (Leung 2006; Wang 2006; Yadav 1989). Two studies (Bensoussan 1998; Brinkhaus 2005) had three arms: one study (Bensoussan 1998) compared standard Chinese herbal medicine, individualised Chinese herbal medicine and placebo and the other study (Brinkhaus 2005) compared the single non-Chinese herb curcuma with another single non-Chinese herb fumitory and with placebo. This gave a total of 12 comparisons in the review.

Population

The definition of IBS varied between studies: one used the Rome I criteria (Bensoussan 1998); two used the Rome II criteria (Leung 2006; Wang 2006); two used authors' definitions (Brinkhaus 2005; Madisch 2004) and one used the Sandler (1984) criteria (Yadav 1989).

All studies but two included people who had a range of IBS types; the other two studies specified diarrhoea predominant IBS symptoms (Leung 2006; Wang 2006). No studies stated that the participants had IBS as result of gastrointestinal infection.

The majority of studies did not state the number of participants with bloating; only two stated that some people had bloating (Madisch 2004; Yadav 1989).

One study described symptom severity as severe (Brinkhaus 2005); the rest did not state symptom severity.

The age range of participants across studies was 13 to 65 years, with the mean age (where given) ranging from 28.4 to 48 years. No study particularly identified elderly participants. There were approximately equal numbers of men and women in total.

Interventions

The studies varied in the type of herbal medicines used:

- Single Chinese herbal medicines: none
- Combination Chinese herbal medicines
  - One used individualised combination herbal preparations from a formulary of 81 Chinese herbs; 15 capsules (dose not stated) (Bensoussan 1998)
  - One used a standard formula Chinese powdered herb capsule containing a combination of 20 different herbs; 15 capsules (Bensoussan 1998)
  - One used a standard formula Chinese herbal preparation containing 11 herbs; water-extracted herbs as granules, which were dissolved in water (Leung 2006)
  - One used the combination Chinese herbal formula Tong-xie-ning granule (TXNG), aqueous extract containing 4 herbs; three times per day, 15 g/day (Wang 2006)
- Single non-Chinese herbal medicines
  - One study used curcuma tablets; three times per day, 60 mg/day of aqueous spray dried extract (Brinkhaus 2005)
  - One study used fumitory tablets; three times per day, 1500 mg/day of aqueous spray dried extract (Brinkhaus 2005)
  - One used bitter candytuft; three times per day, 20 drops (Madisch 2004)
- Combination non-Chinese herbal medicines
  - One used the commercially available preparation STW 5 containing 9 herbs; three times per day, 20 drops (Madisch 2004)
  - One used STW 5-II, a research preparation containing 6 herbs; three times per day, 20 drops (Madisch 2004)
  - One used an Ayurvedic combination containing two herbs plus excipient (polyvinylpyrrolidone, glucose, citric acid); granules taken three times per day, 15 g/day (Yadav 1989).

The doses used, where given, varied across studies, ranging from 60mg to 15g/day, with some studies being unclear on the dose used (Bensoussan 1998; Leung 2006; Madisch 2004).

The duration of the intervention ranged from 3 to 18 weeks. One study (Wang 2006) was only 3 weeks duration and the GDG decided not to consider this further. One study had a duration of 4 weeks (Madisch 2004) one had 6 weeks (Yadav 1989), one 8 weeks (Leung 2006), one 16 weeks (Bensoussan 1998) and one 18 weeks (Brinkhaus 2005).

Comparisons

The included studies covered the following comparisons:

- Single Chinese herbal medicines versus placebo: none
- Combination Chinese herbal medicines versus placebo:

- o Individualised Chinese combination herbal preparation versus placebo (Bensoussan 1998)
  - o Standard formula Chinese powdered herb capsule containing a combination of 20 different herbs versus placebo (Bensoussan 1998)
  - o Standard formula Chinese herbal preparation containing 11 herbs versus placebo (Leung 2006)
  - o Tong-xie-ning granule (TXNG) containing 4 herbs versus placebo (Wang 2006)
- Single non-Chinese herbal medicines versus placebo:
  - o Curcuma versus placebo (Brinkhaus 2005)
  - o Fumitory versus placebo (Brinkhaus 2005)
  - o Bitter candytuft versus placebo (Madisch 2004)
- Combination non-Chinese herbal medicines versus placebo:
  - o STW 5 containing 9 herbs versus placebo (Madisch 2004)
  - o STW 5-II, a research preparation containing 6 herbs versus placebo (Madisch 2004)
  - o Ayurvedic combination containing 2 herbs versus placebo (Yadav 1989).
- Combination Chinese herbal medicines versus single Chinese herbal medicines: none
- Combination non-Chinese herbal medicines versus single non-Chinese herbal medicines
  - o STW 5 containing 9 herbs versus bitter candytuft (Madisch 2004)
  - o STW 5-II, a research preparation containing 6 herbs versus bitter candytuft (Madisch 2004)
- Single Chinese herbal medicines type 1 versus type 2: none
- Combination Chinese herbal medicines type 1 versus type 2:
  - o Individualised Chinese combination versus standard formula Chinese powdered herb capsule (Bensoussan 1998)
- Single non-Chinese herbal medicines type 1 versus type 2:
  - o Curcuma versus fumitory (Brinkhaus 2005)
- Combination non-Chinese herbal medicines type 1 versus type 2:
  - o Commercially available STW 5 containing 9 herbs versus STW 5-II, a research preparation containing 6 herbs (Madisch 2004).

OUTCOMES

The studies measured a range of outcomes.

1. Global symptoms

a) Number of people with an improvement in global symptoms

Four studies recorded the participants' assessment of improvement at the end of the study (ranged from 6 to 18 weeks) (Bensoussan 1998; Brinkhaus 2005; Leung 2006; Yadav 1989).

b) Global symptom score (mean)

The global symptom score was recorded by two studies: Madisch 2004 at 4 weeks on a scale of 0 to 12 (severe) and Bensoussan 1998 at 16 weeks on a 4 x 100mm VAS.

c) Global assessment of IBS discomfort

This outcome was recorded by one study at 4 weeks (Madisch 2004) on a VAS of 0 to 100mm (most intense discomfort).

d) Global assessment of treatment on symptoms

This outcome was recorded by one study (Madisch 2004). Madisch (2004) recorded the physician's assessment of the efficacy of the drug at 4 weeks.

2. Individual symptoms

a) Pain

Pain was reported in several ways, either giving the number of people with pain at the end of the study, the number of people whose pain improved or worsened compared with the baseline, and pain scores. The latter recorded a range of features, including severity, frequency and duration, or a combination of these. In addition, studies recorded the final scores, mean daily scores or the change from baseline. The studies reporting these outcomes are as follows:

- Number of people with an improvement in pain symptoms: one study (Yadav 1989).
- Pain score: two studies (Leung 2006 [scale 0 to 4 (high = bad)]; Madisch 2004 (scale 0 to 21 [high = bad]).
- Change in pain score: one study (Brinkhaus 2005); scale 0 to 50mm VAS (high = worse).

b) Bloating

- Number of people with no bloating: one study (Madisch 2004)
- Change in bloating score: one study (Brinkhaus 2005); scale 0 to 50mm VAS (high = worse).

c) Bowel habits

- Stool score: one study using Bristol stool score (Leung 2006).
- Stool frequency: one study (Leung 2006).
- Number of people with improvement in bowel habit (alternating diarrhoea and constipation): one study (Yadav 1989)
- Number of people with absence of bowel habit problems (diarrhoea or constipation or both): one study (Madisch 2004)
- Number of people with an improvement in bowel symptoms (Yadav 1989).

3. Quality of Life

- One study reported quality of life as an outcome (Leung 2006).

- Two studies reported the number of patients with improvement in psychological distress (Brinkhaus 2005 [psychosocial stress], Yadav 1989 [Hamilton scale]).

4. Adverse effects

- Two studies reported adverse effects (Brinkhaus 2005; Madisch 2004).

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in the Appendix. The method of randomisation was adequate (computer generated) in two studies (Leung 2006; Wang 2006), partially adequate in two (Bensoussan 1998; Madisch 2004) and unclear in two (Brinkhaus 2005; Yadav 1989).

Allocation concealment was adequate in two studies (Brinkhaus 2005; Leung 2006), and partially adequate in one (Wang 2006).

All the studies reported that the participants were blinded to the interventions; outcome assessors were blinded in four studies (Bensoussan 1998; Leung 2006; Madisch 2004; Wang 2006) but this was unclear in two studies (Brinkhaus 2005; Yadav 1989). All described in detail the appearance and taste of the placebo and active intervention.

Four studies (Bensoussan 1998; Leung 2006; Madisch 2004; Wang 2006) described an *a-priori* power calculation. Three studies used an intention to treat analysis (Brinkhaus 2005; Madisch 2004; Wang 2006). All studies included in the review demonstrated some level of baseline comparability of the groups.

Two studies had no loss to follow-up (Brinkhaus 2005; Madisch 2004). One study (Yadav 1989) reported that more than 20% of participants in at least one arm (or overall) were not analysed or were lost to follow-up (21% overall; 11/68 on herbal therapy [16%] and 18/70 on placebo [26%]), while in the other three studies, fewer than 20% of participants were lost to follow-up.

One study had a duration of intervention of only three weeks (Wang 2006) and was not considered further.

The risk of bias was assessed for each included study and only Yadav (1989) was considered to be at higher risk of bias due to the attrition rate of 26% in the placebo arm. This was considered, where possible, in sensitivity analyses.

RESULTS

A. Herbal medicines versus placebo

A1. Combination Chinese herbal medicines versus placebo

All comparisons were of combinations of herbs versus placebo, rather than single herbs:

- Individualised Chinese combination herbal preparation versus placebo (Bensoussan 1998)
- Standard formula Chinese powdered herb capsule containing a combination of 20 different herbs versus placebo (Bensoussan 1998)
- Standard formula Chinese herbal preparation containing 11 herbs versus placebo (Leung 2006)
- Combination Chinese herbal formula Tong-xie-ning granule (TXNG) containing 4 herbs versus placebo (Wang 2006).

I. Global symptoms

a) Number of people with an improvement in global symptoms

Two studies (three comparisons) in 218 people recorded the participants' assessment of improvement at the end of treatment (Bensoussan 1998; Leung 2006).

Figure 1: Global improvement of symptoms (number of people)
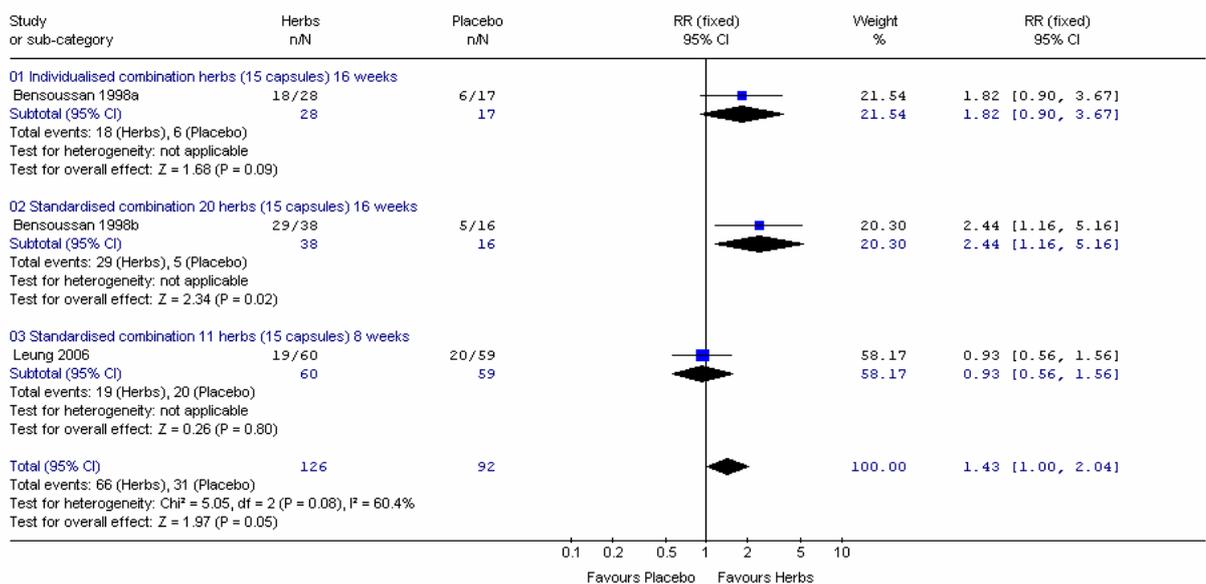
There was significant heterogeneity across studies (I^2=60, p=0.08). It is unclear if the cause of the heterogeneity is:

- Different types of herbs
- Different durations of intervention (Bensoussan 1998 had 16 weeks; Leung 2006 had 8 weeks)
- Differences in dose (although the doses were not stated)

- Differences in the solubility of the herbs (the Leung 2006 intervention was aqueous granules; Bensoussan 1998 had encapsulated herbs)
- Differences in the type of IBS (Leung 2006 patients had diarrhoea predominant; Bensoussan 1998 had mixed types of IBS).

Bensoussan (1998) also measured this outcome at a follow up 14 weeks after the end of the trial, however only 51% of the participants originally randomised to placebo were available at this point, and we decided not to report the results here.

b) Global symptom score (mean)

The global symptom score was recorded by one study (Bensoussan 1998) in 61 people at 16 weeks on a scale of maximum 400.

Figure 2: Global symptom score
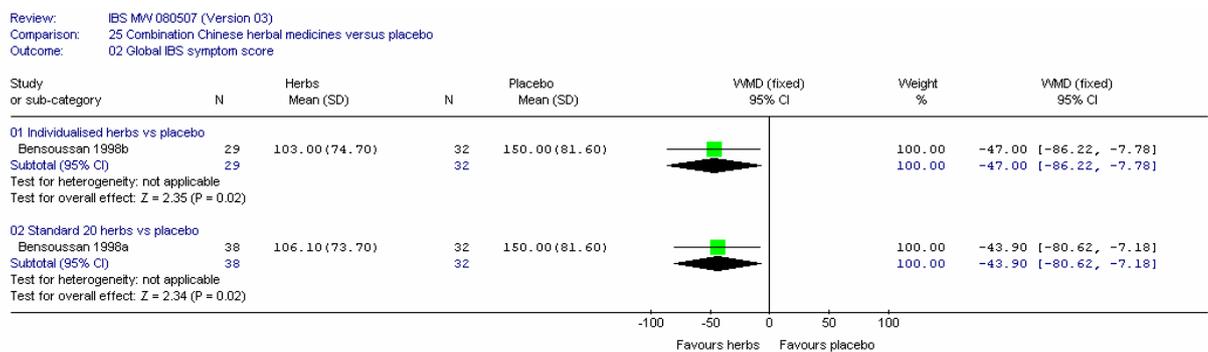
There was a significantly lower score for the herbs groups compared with placebo.

II. Individual symptoms

a) Pain

i. Pain score

One study in 119 people recorded a pain score (Leung 2006). This was rated from 0 (none) to 4 (very severe). The median score for people taking herbs was 1 (range 0 to 4) and for those on placebo was 1.5 (range 0 to 4); this was not significantly different.

b) Bowel habits

i. Stool score

Stool score was assessed in one study (Leung 2006) using the Bristol Stool Scale. The participants in the herb group scored a median of 5 (range 3 to 6) and those on placebo scored 5 (1 to 6); this was not significantly different.

ii. Stool frequency

Stool frequency (daily) was assessed in one study (Leung 2006). The people in the herb group scored a median of 2 (range 1 to 7) and those on placebo scored 2 (1 to 6); this was not significantly different.

III. Quality of Life

One study in 119 people reported quality of life as an outcome using the SF-36 physical and mental scales, each of which has a maximum score of 100 (Leung 2006). There was no significant difference between herbs and placebo.

Figure 3: Quality of life

Review: IBS (Version herbal jacoby)
Comparison: 41 Combination Chinese Herbal Medicine vs Placebo
Outcome: 04 Quality of life (SF36)

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 Physical component (0-100)							
Leung 2006	46	38.00(10.20)	49	34.60(12.00)		100.00	3.40 [-1.07, 7.87]
Subtotal (95% CI)	46		49			100.00	3.40 [-1.07, 7.87]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.49 (P = 0.14)							
02 Mental component (0-100)							
Leung 2006	46	41.50(10.20)	49	39.60(12.10)		100.00	1.90 [-2.59, 6.39]
Subtotal (95% CI)	46		49			100.00	1.90 [-2.59, 6.39]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.83 (P = 0.41)							

-100   -50   0   50   100
Favours placebo   Favours herbs

A2. Single non-Chinese herbal medicines versus placebo:

- Single non-Chinese herb curcuma versus placebo (Brinkhaus 2005)
- Single non-Chinese herb fumitory versus placebo (Brinkhaus 2005)
- Single non-Chinese herbal preparation bitter candytuft versus placebo (Madisch 2004).

I. Global symptoms

a) Global improvement in symptoms (number of patients)

One study recorded the participants' assessment of improvement in 67 people (Brinkhaus 2005). There was no significant difference between interventions.

Figure 4: Global improvement of symptoms

Review: IBS (Version herbal jacoby)
Comparison: 42 Single non-Chinese Herbal Medicine vs Placebo
Outcome: 01 Global improvement of IBS symptoms

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Curcuma tablets 60mg/day (aqueous extract) 18 weeks					
Brinkhaus 2005a	12/21	8/12		51.81	0.86 [0.50, 1.48]
Subtotal (95% CI)	21	12		51.81	0.86 [0.50, 1.48]
Total events: 12 (Herbs), 8 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.55 (P = 0.58)					
02 Fumitory tablets 1500mg/day (aqueous extract) 18 weeks					
Brinkhaus 2005a	14/23	7/11		48.19	0.96 [0.55, 1.66]
Subtotal (95% CI)	23	11		48.19	0.96 [0.55, 1.66]
Total events: 14 (Herbs), 7 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.16 (P = 0.88)					
Total (95% CI)	44	23		100.00	0.91 [0.61, 1.33]
Total events: 26 (Herbs), 15 (Placebo)					
Test for heterogeneity: Chi² = 0.08, df = 1 (P = 0.78), I² = 0%					
Test for overall effect: Z = 0.50 (P = 0.61)					

0.1   0.2   0.5   1   2   5   10
Favours Placebo   Favours Herbs

b) Global symptom score (mean)

The global symptom score (scale 0 to 12) was recorded in one study (Madisch 2004) in 104 people at 4 weeks. There was no significant difference between interventions.

Figure 5: Global symptom score
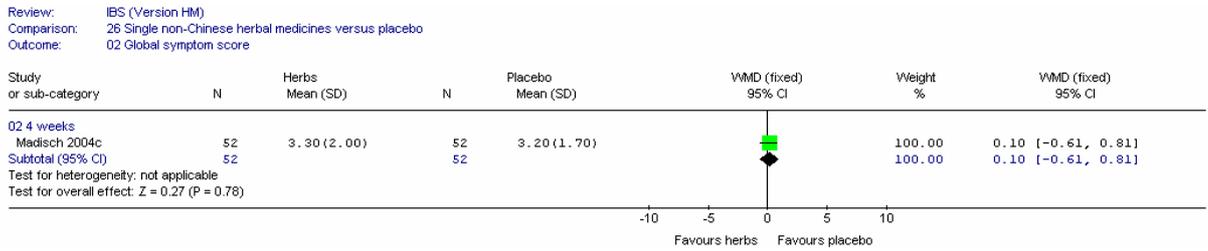
Review:	IBS (Version HM)
Comparison:	26 Single non-Chinese herbal medicines versus placebo
Outcome:	02 Global symptom score

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004c	52	3.30(2.00)	52	3.20(1.70)		100.00	0.10 [-0.61, 0.81]
Subtotal (95% CI)	52		52			100.00	0.10 [-0.61, 0.81]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.27 (P = 0.78)							

-10 -5 0 5 10
Favours herbs   Favours placebo

c) Global assessment of IBS discomfort

This outcome was recorded by one study at 4 weeks (Madisch 2004) on a 100mm VAS. There was significantly more global discomfort for the group given placebo.

Figure 6: Global assessment of IBS discomfort
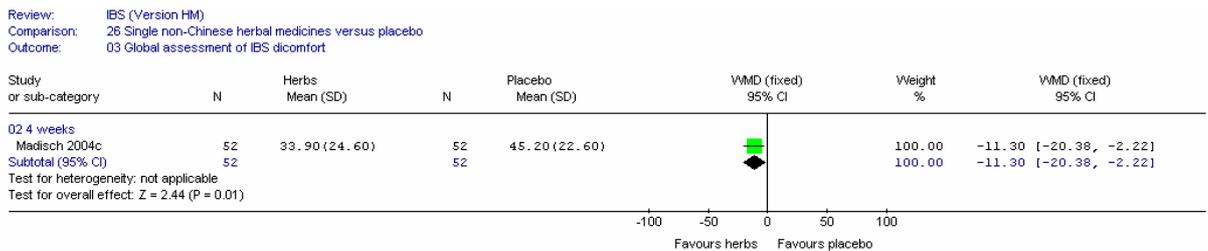
Review:	IBS (Version HM)
Comparison:	26 Single non-Chinese herbal medicines versus placebo
Outcome:	03 Global assessment of IBS dicomfort

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004c	52	33.90(24.60)	52	45.20(22.60)		100.00	-11.30 [-20.38, -2.22]
Subtotal (95% CI)	52		52			100.00	-11.30 [-20.38, -2.22]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.44 (P = 0.01)							

-100 -50 0 50 100
Favours herbs   Favours placebo

d) Global assessment of treatment on symptoms

This outcome was recorded by one study (Madisch 2004), which recorded the physician's assessment of the efficacy of the intervention at 4 weeks (number of people recorded as 'good' or 'very good'). There was no significant difference between interventions.

Figure 7: Global assessment of efficacy (by clinician)

Review:	IBS (Version HM)
Comparison:	26 Single non-Chinese herbal medicines versus placebo
Outcome:	04 Global assessment of efficacy ('good' or 'very good')

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Madisch 2004c	25/53	20/52		100.00	1.23 [0.78, 1.92]
Total (95% CI)	53	52		100.00	1.23 [0.78, 1.92]
Total events: 25 (Herbs), 20 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.90 (P = 0.37)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours herbs

II. Individual symptoms

a) Pain

i. Pain score

Two studies recorded a pain score. Brinkhaus (2005) (in 106 participants) recorded the change in pain on a VAS 0 to 50mm, with high being good; Madisch (2004) (in 104 participants) reported a pain score on a scale of 0 to 21, where a high score was more severe pain. These studies were analysed separately, but neither showed a significant effect of herbs on pain.

Figure 8: Pain score and change in pain score
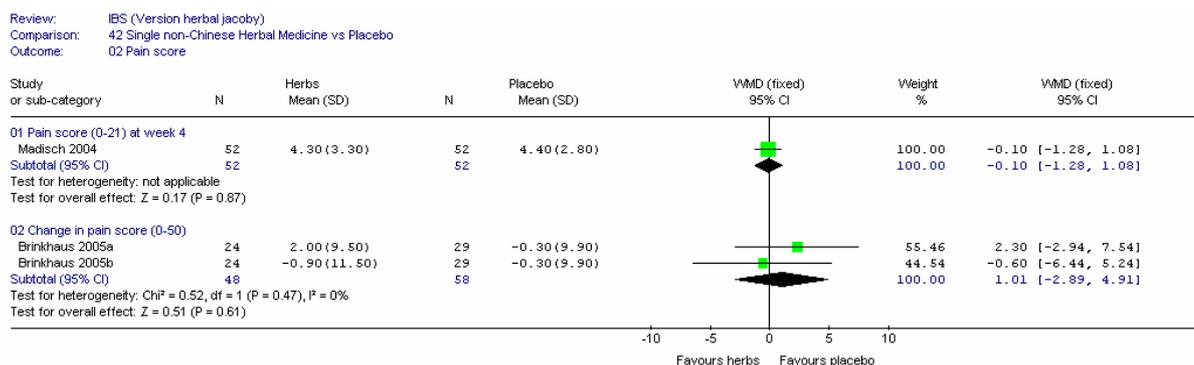
b) Bloating

i. Bloating score

One study recorded the change in bloating score on a scale of 0 to 50mm (Brinkhaus 2005). There was no significant difference between interventions.

Figure 9: Change in bloating score
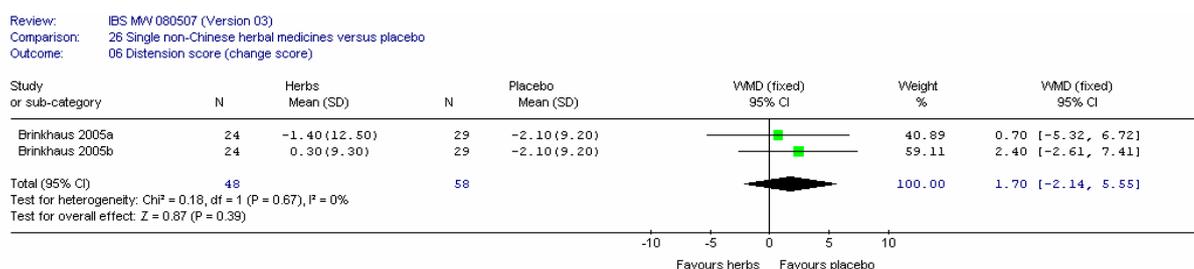
ii. Number of people with bloating

One study (Madisch 2004), in 104 participants, recorded the number of people with no bloating. There were significantly more people with an absence of bloating for the herbs group compared with placebo; RR 2.00 (95%CI 1.15, 3.46).

Figure 10: Number of people with no bloating

Review: IBS (Version herbal jacoby)
Comparison: 42 Single non-Chinese Herbal Medicine vs Placebo
Outcome: 03 Bloating (No of people with no bloating)

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
02 Number with absence of sensation of fullness					
Madisch 2004	25/51	13/53		100.00	2.00 [1.15, 3.46]
Subtotal (95% CI)	51	53		100.00	2.00 [1.15, 3.46]
Total events: 25 (Herbs), 13 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.47 (P = 0.01)					

0.1   0.2   0.5   1   2   5   10
Favours placebo   Favours Herbs

c) Bowel habits

One study (Madisch 2004) recorded the number of people with complete relief or absence of changed bowel habit (constipation, diarrhoea or alternating bowel habit). There was a significant effect in favour of the herbal medicine, but the confidence interval is very wide, so this conclusion is uncertain.

Figure 11: Number of people with absence of changed bowel habit

Review: IBS (Version HM)
Comparison: 26 Single non-Chinese herbal medicines versus placebo
Outcome: 07 Number of patients with relief of changes in bowel habit (constipation, diarrhoea or alternating)

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
02 4 weeks					
Madisch 2004c	8/50	1/49		100.00	7.84 [1.02, 60.37]
Subtotal (95% CI)	50	49		100.00	7.84 [1.02, 60.37]
Total events: 8 (Herbs), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.98 (P = 0.05)					

0.01   0.1   1   10   100
Favours placebo   Favours herbs

III. Psychological distress

One study in 90 participants reported the number of people whose psychological distress caused by IBS was improved with treatment (Brinkhaus 2005). There was no significant difference between interventions.

Figure 12: Number of people with reduction in psychological stress

Review: IBS (Version HM)
Comparison: 26 Single non-Chinese herbal medicines versus placebo
Outcome: 09 Number of patients whose psychological stress caused by IBS was improved with treatment

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Brinkhaus 2005a	11/21	13/23		50.84	0.93 [0.54, 1.59]
Brinkhaus 2005b	12/23	12/23		49.16	1.00 [0.58, 1.74]
Total (95% CI)	44	46		100.00	0.96 [0.65, 1.42]
Total events: 23 (Herbs), 25 (Placebo)					
Test for heterogeneity: Chi² = 0.04, df = 1 (P = 0.85), I² = 0%					
Test for overall effect: Z = 0.19 (P = 0.85)					

0.1   0.2   0.5   1   2   5   10
Favours placebo   Favours herbs

IV. Adverse effects

a) Adverse effects related to treatment

Two studies reported adverse effects related to the study treatment (Brinkhaus 2005; Madisch 2004). There were significantly more adverse effects in the herbal medicine group, but the confidence interval was wide, and the type of herbal medicine seemed to be important.

Figure 13: Number of people with adverse effects

Review: IBS (Version herbal jacoby)
Comparison: 42 Single non-Chinese Herbal Medicine vs Placebo
Outcome: 04 No of people with adverse effects

Study or sub-category	Herbs n/N	Placebo n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
01 Curcuma 18 weeks					
Brinkhaus 2005a	1/24	1/29		44.23	1.22 [0.07, 20.55]
Subtotal (95% CI)	24	29		44.23	1.22 [0.07, 20.55]
Total events: 1 (Herbs), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.14 (P = 0.89)					
02 Fumitory tablets 18 weeks					
Brinkhaus 2005a	8/24	1/29		30.77	14.00 [1.60, 122.33]
Subtotal (95% CI)	24	29		30.77	14.00 [1.60, 122.33]
Total events: 8 (Herbs), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.39 (P = 0.02)					
03 Bitter candytuft 4 weeks					
Madisch 2004	1/53	0/52		25.00	3.00 [0.12, 75.34]
Subtotal (95% CI)	53	52		25.00	3.00 [0.12, 75.34]
Total events: 1 (Herbs), 0 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 0.67 (P = 0.50)					
Total (95% CI)	101	110		100.00	5.60 [1.38, 22.64]
Total events: 10 (Herbs), 2 (Placebo)					
Test for heterogeneity: Chi² = 1.95, df = 2 (P = 0.38), I² = 0%					
Test for overall effect: Z = 2.42 (P = 0.02)					

0.01 0.1 1 10 100
Favours herbs   Favours placebo

NB: Scale 0.01 to 100

b) Number of people reporting tolerability

Two studies gave the number of people reporting tolerability to treatment. There was no significant difference between interventions and no heterogeneity across studies (I^2=9%, p=0.33)

Figure 14: Number of people reporting tolerability

Review: IBS (Version HM)
Comparison: 26 Single non-Chinese herbal medicines versus placebo
Outcome: 10 Tolerability 'good' or 'very good'

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Brinkhaus 2005a	19/21	21/23		22.78	0.99 [0.82, 1.20]
Brinkhaus 2005b	23/23	21/23		24.44	1.09 [0.94, 1.27]
Madisch 2004c	44/53	46/52		52.78	0.94 [0.80, 1.10]
Total (95% CI)	97	98		100.00	0.99 [0.90, 1.09]
Total events: 86 (Herbs), 88 (Placebo)					
Test for heterogeneity: Chi² = 2.19, df = 2 (P = 0.33), I² = 8.8%					
Test for overall effect: Z = 0.24 (P = 0.81)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours herbs

A3. Combination non-Chinese herbal medicines versus placebo:

- Non-Chinese herbal preparation STW 5 containing 9 herbs versus placebo (Madisch 2004)

- Non-Chinese herbal preparation STW 5-II, a research preparation containing 6 herbs versus placebo (Madisch 2004)
- Non-Chinese Ayurvedic combination containing two herbs (Yadav 1989).

I. Global symptoms

a) Number of people with an improvement in global symptoms

One study (Yadav 1989) in 109 participants recorded their assessment of improvement ('good' or 'satisfactory' response) after 6 weeks. There were significantly more people with improvement in symptoms in the herbal medicine group; RR 1.99 (95%CI 1.29, 3.07). However, this study was considered to be at higher risk of bias due to the attrition rate of 26% in the placebo arm.

Figure 15: Global improvement of symptoms

Review: IBS (Version herbal jacoby)
Comparison: 43 Combination non-Chinese Herbal Medicine vs Placebo
Outcome: 01 Global improvement of IBS symptoms

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Ayurvedic combination of two herbs					
Yadav 1989	37/57	17/52		100.00	1.99 [1.29, 3.07]
Subtotal (95% CI)	57	52		100.00	1.99 [1.29, 3.07]
Total events: 37 (Herbs), 17 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.10 (P = 0.002)					
Total (95% CI)	57	52		100.00	1.99 [1.29, 3.07]
Total events: 37 (Herbs), 17 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.10 (P = 0.002)					

0.1 0.2 0.5 1 2 5 10
Favours Placebo   Favours Herbs

b) Global symptom score (mean)

The global symptom score was recorded by one study (Madisch 2004) in 151 people at 4 weeks. The scale was 0 to 12. There was a significant difference between interventions, favouring the combination herbal medicine group.

Figure 16: Global symptom score

Review: IBS (Version HM)
Comparison: 27 Combination non-Chinese herbal medicines versus placebo
Outcome: 02 Global IBS score

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004a	48	2.10 (1.40)	26	3.20 (1.70)		53.47	-1.10 [-1.86, -0.34]
Madisch 2004b	51	2.10 (1.80)	26	3.20 (1.70)		46.53	-1.10 [-1.92, -0.28]
Subtotal (95% CI)	99		52			100.00	-1.10 [-1.66, -0.54]
Test for heterogeneity: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%							
Test for overall effect: Z = 3.86 (P = 0.0001)							

-10 -5 0 5 10
Favours herbs   Favours placebo

c) Global assessment of IBS discomfort

This outcome was recorded by one study at 4 weeks on a 100mm VAS (Madisch 2004). There was significantly more global discomfort for the group given placebo.

Figure 17: Global assessment of IBS discomfort

```
Review:        IBS (Version HM)
Comparison:    27 Combination non-Chinese herbal medicines versus placebo
Outcome:       03 Global assessment of IBS discomfort

Study                          Herbs                  Placebo             WMD (fixed)         Weight      WMD (fixed)
or sub-category        N       Mean (SD)       N      Mean (SD)             95% CI              %           95% CI

02 4 weeks
  Madisch 2004a       48      27.30(19.30)    26     45.20(22.60)                             48.48     -17.90 [-28.16, -7.64]
  Madisch 2004b       51      26.10(17.70)    26     45.20(22.60)                             51.52     -19.10 [-29.05, -9.15]
Subtotal (95% CI)     99                      52                                             100.00    -18.52 [-25.66, -11.37]
Test for heterogeneity: Chi² = 0.03, df = 1 (P = 0.87), I² = 0%
Test for overall effect: Z = 5.08 (P < 0.00001)

                                                          -100    -50     0    50    100
                                                            Favours herbs   Favours placebo
```

d) Global assessment of treatment on symptoms

This outcome was recorded by one study (Madisch 2004) in 154 participants, which recorded the physician's assessment of the efficacy of the intervention at 4 weeks. The herbal medicines were judged to be significantly more efficacious than placebo; RR 1.78 (95%CI 1.23, 2.58).

Figure 18: Global assessment of treatment on symptoms

```
Review:        IBS (Version herbal jacoby)
Comparison:    43 Combination non-Chinese Herbal Medicine vs Placebo
Outcome:       05 Global assessment of efficacy (by clinician)

Study                        Herbs          Placebo            RR (fixed)         Weight       RR (fixed)
or sub-category              n/N             n/N                 95% CI             %            95% CI

01 STW 5 containing 9 herbs (4 weeks)
  Madisch 2004              33/51           10/26                                 50.00      1.68 [0.99, 2.85]
Subtotal (95% CI)            51              26                                   50.00      1.68 [0.99, 2.85]
Total events: 33 (Herbs), 10 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.94 (P = 0.05)

02 STW 5-II, a research preparation containing 6 herbs  (4 weeks)
  Madisch 2004              37/51           10/26                                 50.00      1.89 [1.13, 3.16]
Subtotal (95% CI)            51              26                                   50.00      1.89 [1.13, 3.16]
Total events: 37 (Herbs), 10 (Placebo)
Test for heterogeneity: not applicable
Test for overall effect: Z = 2.42 (P = 0.02)

Total (95% CI)              102              52                                  100.00      1.78 [1.23, 2.58]
Total events: 70 (Herbs), 20 (Placebo)
Test for heterogeneity: Chi² = 0.09, df = 1 (P = 0.76), I² = 0%
Test for overall effect: Z = 3.08 (P = 0.002)

                                          0.1  0.2   0.5   1    2    5   10
                                            Favours Placebo   Favours Herbs
```

II. Individual symptoms

a) Pain

i. Number of people with improved pain response

One study (Yadav 1989) reported the number of people with an improvement in pain symptoms (excellent', 'good' or 'satisfactory' improvement). There was no significant difference between interventions. However, this study was considered to be at higher risk of bias due to the attrition rate of 37% in the placebo group and 29% in the herbal medicines group for this outcome.

Figure 19: Improvement in pain symptoms

Review: IBS (Version herbal jacoby)
Comparison: 43 Combination non-Chinese Herbal Medicine vs Placebo
Outcome: 02 No of patients with pain response of excellent, good or satisfactory

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Ayurvedic combination of two herbs					
Yadav 1989	21/48	13/44		100.00	1.48 [0.85, 2.59]
Subtotal (95% CI)	48	44		100.00	1.48 [0.85, 2.59]
Total events: 21 (Herbs), 13 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.38 (P = 0.17)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours Herbs

ii. Pain score

One study recorded a pain score in 151 participants on a scale of 0 to 21 at 4 weeks (Madisch 2004). There was a statistically significant reduction in pain for the herbal medicines group: WMD -1.65 (95%CI -2.58, -0.72).

Figure 20: Pain score

Review: IBS (Version HM)
Comparison: 27 Combination non-Chinese herbal medicines versus placebo
Outcome: 06 Pain score

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004a	48	2.90(2.60)	26	4.40(2.80)		50.72	-1.50 [-2.80, -0.20]
Madisch 2004b	51	2.60(2.80)	26	4.40(2.80)		49.28	-1.80 [-3.12, -0.48]
Subtotal (95% CI)	99		52			100.00	-1.65 [-2.58, -0.72]
Test for heterogeneity: Chi² = 0.10, df = 1 (P = 0.75), I² = 0%							
Test for overall effect: Z = 3.48 (P = 0.0005)							

-10 -5 0 5 10
Favours herbs   Favours placebo

b) Bowel habits

i. Improvement in symptoms of bowel habit (constipation, diarrhoea or alternating)

One study (Yadav 1989) reported the number of people with an improvement in bowel symptoms. Overall there was a significant improvement for the people given herbal medicines. However, this study was considered to be at higher risk of bias due to the attrition rate of 26% in the placebo group.

The study also reported sub-group analyses by type of IBS, but the participants were not stratified before randomisation and so we have not included the results in this review.

One other study (Madisch 2004) reported the number of participants with an absence of bowel problems and found significantly more people had no bowel problems in the herbal medicine groups (shown combined in Figure 21). The confidence interval was very wide, however.

Figure 21: Improvement of bowel symptoms

Review: IBS (Version herbal jacoby)
Comparison: 43 Combination non-Chinese Herbal Medicine vs Placebo
Outcome: 03 No of patients with improvement in bowel symptoms

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Ayurvedic combination herbs: number with improved bowel symptoms					
Yadav 1989	42/57	22/52		94.58	1.74 [1.22, 2.48]
Subtotal (95% CI)	57	52		94.58	1.74 [1.22, 2.48]
Total events: 42 (Herbs), 22 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 3.08 (P = 0.002)					
02 STW5 combination: number with no bowel symptoms					
Madisch 2004	17/95	1/49		5.42	8.77 [1.20, 63.96]
Subtotal (95% CI)	95	49		5.42	8.77 [1.20, 63.96]
Total events: 17 (Herbs), 1 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.14 (P = 0.03)					
Total (95% CI)	152	101		100.00	2.12 [1.46, 3.08]
Total events: 59 (Herbs), 23 (Placebo)					
Test for heterogeneity: Chi² = 3.16, df = 1 (P = 0.08), I² = 68.4%					
Test for overall effect: Z = 3.95 (P < 0.0001)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours Herbs

III. Psychological improvement

One study reported the number of patients with a psychological improvement (Yadav 1989). There was a significantly greater improvement for the herbal medicine group; RR 2.10 (95% CI 1.23, 3.50). However, this study was considered to be at higher risk of bias due to the attrition rate of over 20%.

Figure 22: Psychological improvement

Review: IBS (Version herbal jacoby)
Comparison: 43 Combination non-Chinese Herbal Medicine vs Placebo
Outcome: 04 No of patients with psychological improvement

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Ayurvedic combination herbs					
Yadav 1989	27/45	12/42		100.00	2.10 [1.23, 3.58]
Subtotal (95% CI)	45	42		100.00	2.10 [1.23, 3.58]
Total events: 27 (Herbs), 12 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.72 (P = 0.007)					
Total (95% CI)	45	42		100.00	2.10 [1.23, 3.58]
Total events: 27 (Herbs), 12 (Placebo)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 2.72 (P = 0.007)					

0.1 0.2 0.5 1 2 5 10
Favours placebo   Favours Herbs

IV. Adverse effects

The Madisch (2004) study reported that only one person in the STW 5 herbal medicine group had an adverse event and none in the placebo group. The Yadav (1989) study reported that two participants had an adverse effect (drowsiness) and there was none in the placebo group. One study (Madisch 2004) reported tolerability (Figure 23). There was no significant difference between interventions.

Figure 23: Tolerability

Review:	IBS (Version HM)				
Comparison:	27 Combination non-Chinese herbal medicines versus placebo				
Outcome:	14 Tolerability 'good' or 'very good'				

Study or sub-category	Herbs n/N	Placebo n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Madisch 2004a	50/51	23/26		50.00	1.11 [0.96, 1.28]
Madisch 2004b	46/51	23/26		50.00	1.02 [0.86, 1.20]
Total (95% CI)	102	52		100.00	1.06 [0.95, 1.19]
Total events: 96 (Herbs), 46 (Placebo)					
Test for heterogeneity: Chi² = 0.56, df = 1 (P = 0.45), I² = 0%					
Test for overall effect: Z = 1.11 (P = 0.27)					

0.1 0.2 0.5 1 2 5 10

Favours placebo   Favours herbs
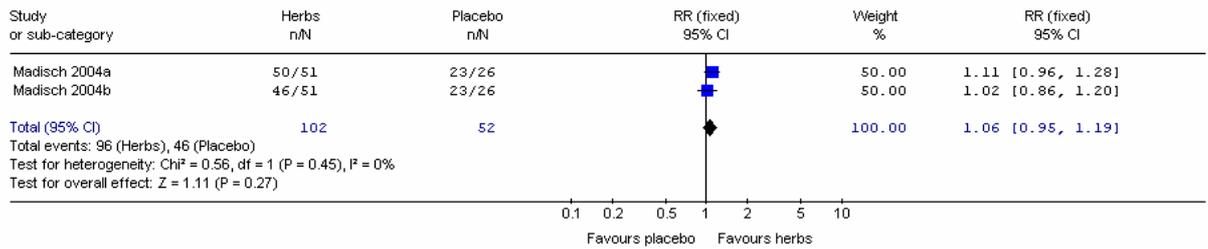
B. Herbal medicine type 1 versus herbal medicine type 2

B1. Combination non-Chinese herbal medicines versus single non-Chinese herbal medicines

- Commercially available non-Chinese herbal preparation STW 5 containing 9 herbs versus single non-Chinese herbal preparation bitter candytuft (Madisch 2004)
- Non-Chinese herbal preparation STW 5-II, a research preparation containing 6 herbs versus single non-Chinese herbal preparation bitter candytuft (Madisch 2004)

I. Global symptoms

a) Global symptom score (mean)

A global symptom score was recorded by one study (Madisch 2004) in 102 participants at 4 weeks. The scale was 0 to 12. There were significantly lower symptom scores for the herbal medicine group: WMD -1.20 (95%CI -1.83, -0.57).

Figure 24: Global symptom score

Review:	IBS (Version herbal jacoby)				
Comparison:	44 Combination non-Chinese Herbal Medicine vs single non-Chinese herbal medicine				
Outcome:	01 Global symptom score				

Study or sub-category	N	Herbs Mean (SD)	N	Placebo Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 STW 5 versus bitter candytuft							
Madisch 2004	48	2.10(1.40)	26	3.30(2.00)		52.75	-1.20 [-2.06, -0.34]
Subtotal (95% CI)	48		26			52.75	-1.20 [-2.06, -0.34]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.72 (P = 0.007)							
02 STW 5-II (research) versus bitter candytuft							
Madisch 2004	51	2.10(1.80)	26	3.30(2.00)		47.25	-1.20 [-2.11, -0.29]
Subtotal (95% CI)	51		26			47.25	-1.20 [-2.11, -0.29]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 2.57 (P = 0.01)							
Total (95% CI)	99		52			100.00	-1.20 [-1.83, -0.57]
Test for heterogeneity: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%							
Test for overall effect: Z = 3.74 (P = 0.0002)							

-10 -5 0 5 10

Favours herbs   Favours placebo

b) Global assessment of IBS discomfort

This outcome was recorded by one study at 4 weeks (Madisch 2004) on a 100mm VAS. There was no significant difference between interventions.

Figure 25: Global assessment of IBS discomfort
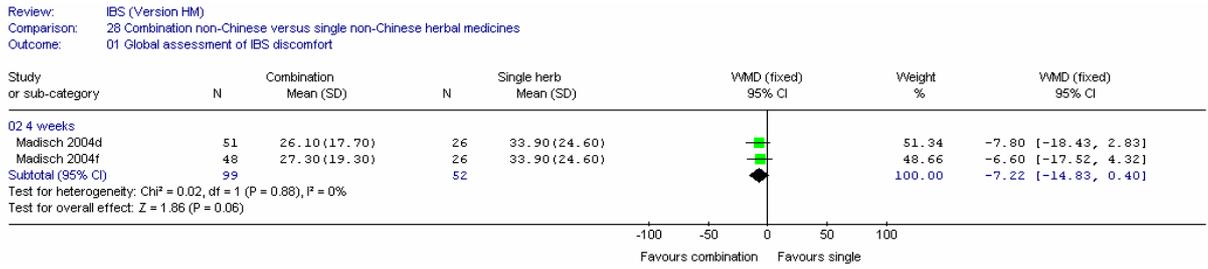
c) Global assessment of treatment on symptoms

This outcome was recorded by one study (Madisch 2004), which recorded the clinician's assessment of the efficacy of the intervention at 4 weeks (number of people recorded as 'good' or 'very good'). There were significantly more people assessed to be efficaciously treated in the combined group compared to the single herbal medicine: RR 1.46 (95% CI 1.06, 1.99). This corresponds to an NNT of 5 (95% CI 3, 20).

Figure 26: Global assessment of efficacy (by clinician)
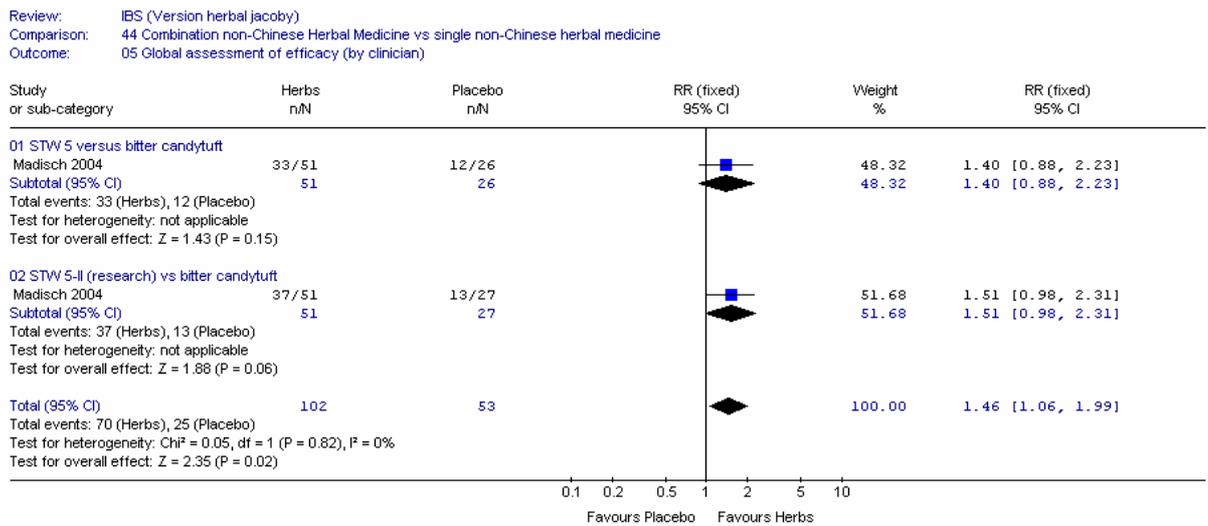
II. Individual symptoms

a) Pain

One study (Madisch 2004) reported a pain score on a scale of 0 to 21. There was significantly less pain for the group treated with combination herbal medicines.

Figure 27: Pain score
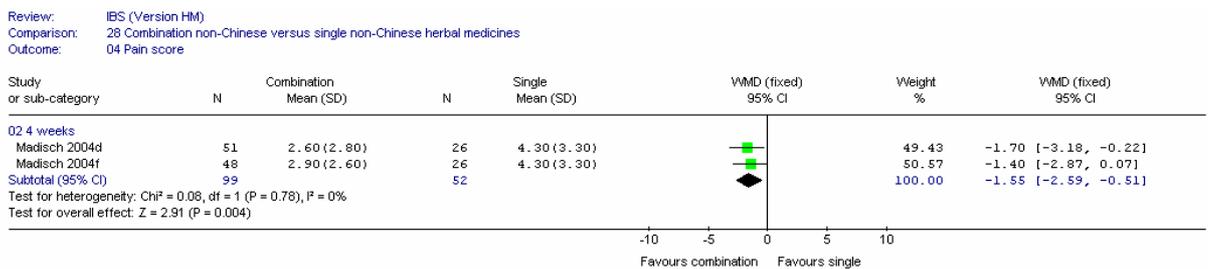
b) Bowel habits

One study (Madisch 2004) recorded the number of people with complete relief or absence of changes in bowel habit (constipation, diarrhoea or alternating bowel habit). There was no significant difference between interventions.

Figure 28: Number of people with no change in bowel habit
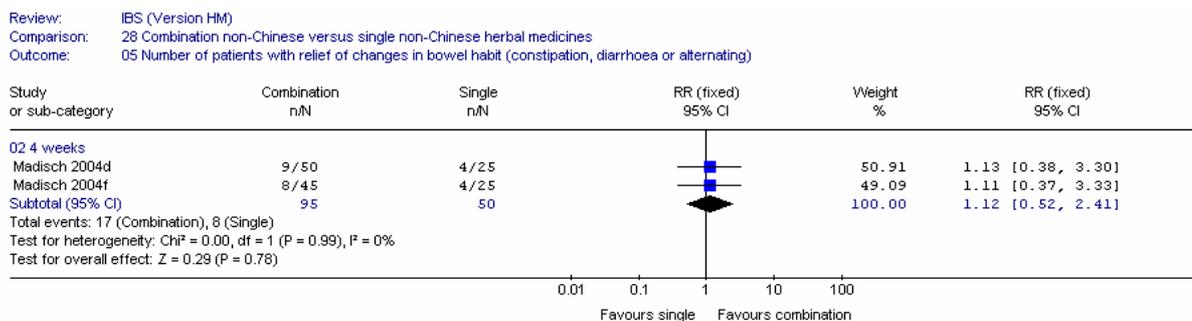
III. Adverse effects

One study reported tolerability (Madisch 2004). The combination herbal medicines were significantly better tolerated than the single bitter candytuft herb.

Figure 29: Tolerability
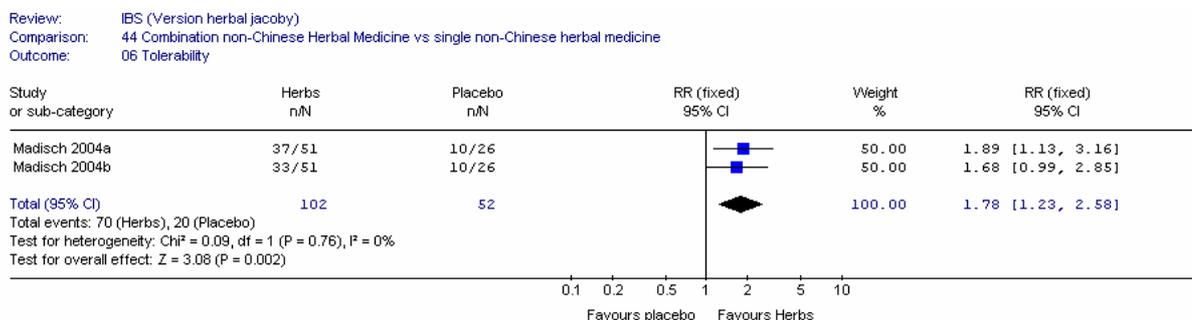
B2. Combination Chinese herbal medicines type 1 versus type 2

- Individualised Chinese combination versus standard formula Chinese powdered herb capsule (Bensoussan 1998c).

I. Global symptoms

a) Number of people with an improvement in global symptoms

One study recorded the participants' assessment of improvement in 66 people at the end of the 16-week treatment period and at follow up 14 weeks later (Bensoussan 1998). There was no significant difference between interventions.

Figure 30: Global improvement of symptoms

Review: IBS (Version HM)
Comparison: 29 Combination Chinese herbal medicine type 1 versus type 2
Outcome: 01 Patients rating of response: 'improved'

Study or sub-category	Standard herbs n/N	Individualised herbs n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 End of treatment					
Bensoussan 1998c	29/38	18/28		100.00	1.19 [0.86, 1.65]
Subtotal (95% CI)	38	28		100.00	1.19 [0.86, 1.65]
Total events: 29 (Standard herbs), 18 (Individualised herbs)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.03 (P = 0.31)					
02 Follow up					
Bensoussan 1998c	22/35	18/24		100.00	0.84 [0.59, 1.18]
Subtotal (95% CI)	35	24		100.00	0.84 [0.59, 1.18]
Total events: 22 (Standard herbs), 18 (Individualised herbs)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.01 (P = 0.31)					

0.1 0.2 0.5 1 2 5 10
Favours individual   Favours standard

b) Global symptom score (mean)

The global symptom score was recorded by one study: Bensoussan (1998) at 16 weeks (end of treatment) and at follow up 14 weeks later, on a scale of 4 x 100mm. There was no significant difference between interventions.

Figure 31: Global symptom score

Review: IBS MW 080507 (Version 03)
Comparison: 29 Combination Chinese herbal medicine type 1 versus type 2
Outcome: 02 Global symptom score

Study or sub-category	N	Standard herb Mean (SD)	N	Individualised herb Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
01 End of treatment (16 weeks)							
Bensoussan 1998c	38	106.10(73.70)	29	103.00(74.70)		100.00	3.10 [-32.79, 38.99]
Subtotal (95% CI)	38		29			100.00	3.10 [-32.79, 38.99]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.17 (P = 0.87)							
02 Follow up (14 weeks later)							
Bensoussan 1998c	35	132.60(90.20)	24	99.40(74.80)		100.00	33.20 [-9.09, 75.49]
Subtotal (95% CI)	35		24			100.00	33.20 [-9.09, 75.49]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 1.54 (P = 0.12)							

-100 -50 0 50 100
Favours standard   Favours individual

B3. Single non-Chinese herbal medicines type 1 versus type 2:

- Single non-Chinese herb curcuma versus single non-Chinese herb fumitory (Brinkhaus 2005).

I. Global symptoms

a) Number of people with an improvement in global symptoms

One study recorded the participants' assessment of improvement in 44 people (Brinkhaus 2005). There was no significant difference between interventions.

Figure 32: Global improvement of symptoms

Review: IBS (Version HM)
Comparison: 30 Single non-Chinese type 1 versus type 2
Outcome: 01 Global improvement in symptoms

Study or sub-category	Curcuma n/N	Fumitory n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Brinkhaus 2005c	12/21	14/23		100.00	0.94 [0.57, 1.54]
Total (95% CI)	21	23		100.00	0.94 [0.57, 1.54]

Total events: 12 (Curcuma), 14 (Fumitory)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.25 (P = 0.80)

0.1 0.2 0.5 1 2 5 10
Favours fumitory   Favours curcuma

II. Individual symptoms

a) Pain

One study recorded a pain score on a VAS of 0 to 50mm (Brinkhaus 2005). There was no significant difference between interventions.

Figure 33: Pain score

Review: IBS (Version HM)
Comparison: 30 Single non-Chinese type 1 versus type 2
Outcome: 02 Change in pain score

Study or sub-category	N	Curcuma Mean (SD)	N	Fumitory Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Brinkhaus 2005c	24	2.00(9.50)	24	-0.90(11.50)		100.00	2.90 [-3.07, 8.87]
Total (95% CI)	24		24			100.00	2.90 [-3.07, 8.87]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.95 (P = 0.34)

-100 -50 0 50 100
Favours curcuma   Favours fumitory

c) Bloating

One study (Brinkhaus 2005) recorded a bloating score on a VAS of 0 to 50mm. There was no significant difference between interventions

Figure 34: Bloating score

Review: IBS (Version HM)
Comparison: 30 Single non-Chinese type 1 versus type 2
Outcome: 03 Change in distension score

Study or sub-category	N	Curcuma Mean (SD)	N	Fumitory Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Brinkhaus 2005c	24	-1.40(12.50)	24	0.30(9.30)		100.00	-1.70 [-7.93, 4.53]
Total (95% CI)	24		24			100.00	-1.70 [-7.93, 4.53]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.53 (P = 0.59)

-100 -50 0 50 100
Favours curcuma   Favours fumitory

III. Psychological distress

One study reported the number of people whose psychological distress caused by IBS was improved with treatment (Brinkhaus 2005). There was no significant difference between interventions.

Figure 35: Global improvement of symptoms

Review: IBS (Version HM)
Comparison: 30 Single non-Chinese type 1 versus type 2
Outcome: 04 Number of people whose psychological distress caused by IBS was improved by treatment

Study or sub-category	Curcuma n/N	Fumitory n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Brinkhaus 2005c	11/21	12/23		100.00	1.00 [0.57, 1.77]
Total (95% CI)	21	23		100.00	1.00 [0.57, 1.77]

Total events: 11 (Curcuma), 12 (Fumitory)
Test for heterogeneity: not applicable
Test for overall effect: Z = 0.01 (P = 0.99)

0.1 0.2 0.5 1 2 5 10
Favours fumitory   Favours curcuma

IV. Adverse effects

One study reported tolerability (Brinkhaus 2005). There was no significant difference between interventions.

Figure 36: Tolerability

Review: IBS (Version HM)
Comparison: 30 Single non-Chinese type 1 versus type 2
Outcome: 05 Tolerability = 'good'

Study or sub-category	Curcuma n/N	Fumitory n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Brinkhaus 2005c	19/21	23/23		100.00	0.91 [0.77, 1.06]
Total (95% CI)	21	23		100.00	0.91 [0.77, 1.06]

Total events: 19 (Curcuma), 23 (Fumitory)
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.22 (P = 0.22)

0.1 0.2 0.5 1 2 5 10
Favours fumitory   Favours curcuma

B4. Combination non-Chinese herbal medicines type 1 versus type 2:

- Commercially available non-Chinese herbal preparation STW 5 containing 9 herbs versus non-Chinese herbal preparation STW 5-II, a research preparation containing 6 herbs (Madisch 2004).

I. Global symptoms

a) Global symptom score (mean)

The global symptom score was recorded by one study in 99 people (Madisch 2004) at 4 weeks on a scale of 0 to 12. There was no significant difference between interventions.

Figure 37: Global symptom score
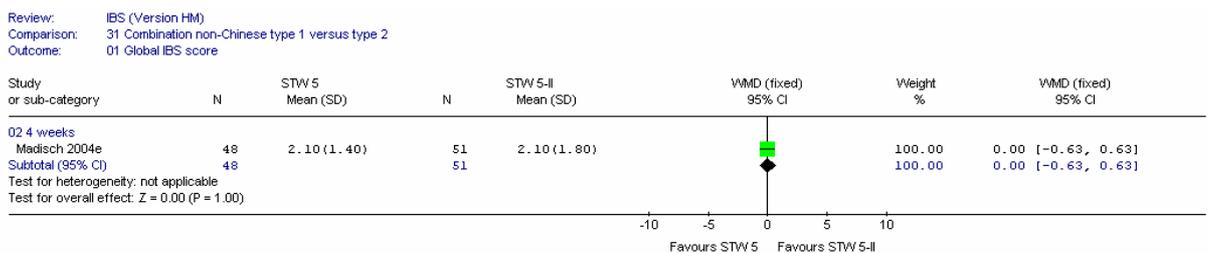
Review: IBS (Version HM)
Comparison: 31 Combination non-Chinese type 1 versus type 2
Outcome: 01 Global IBS score

Study or sub-category	N	STW 5 Mean (SD)	N	STW 5-II Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004e	48	2.10 (1.40)	51	2.10 (1.80)		100.00	0.00 [-0.63, 0.63]
Subtotal (95% CI)	48		51			100.00	0.00 [-0.63, 0.63]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.00 (P = 1.00)

-10 -5 0 5 10
Favours STW 5   Favours STW 5-II

b) Global assessment of IBS discomfort

This outcome was recorded by one study at 4 weeks (Madisch 2004) on a scale of 0 to 100mm.
There was no significant difference between interventions.

Figure 38: Global assessment of IBS discomfort
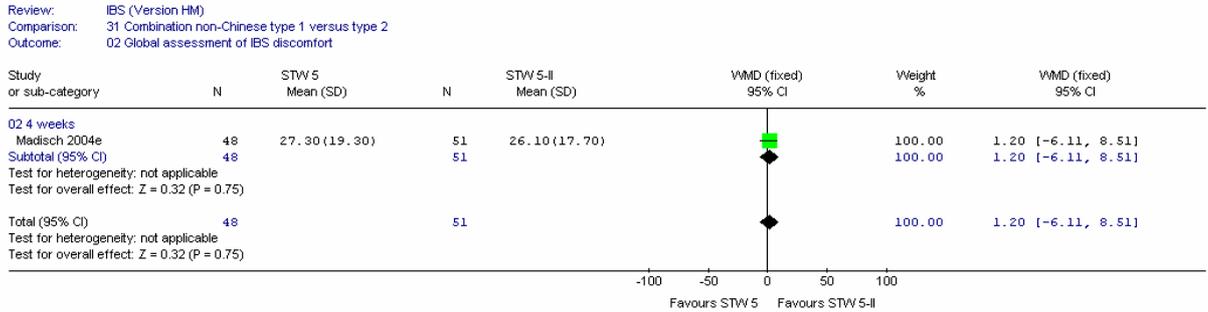
c) Global assessment of treatment on symptoms

This outcome was recorded by one study (Madisch 2004) in 102 people, which recorded the
clinician's assessment of the efficacy of the drug at 4 weeks (number of people recorded as
'good' or 'very good'). There was no significant difference between interventions.

Figure 39: Global assessment of treatment on symptoms
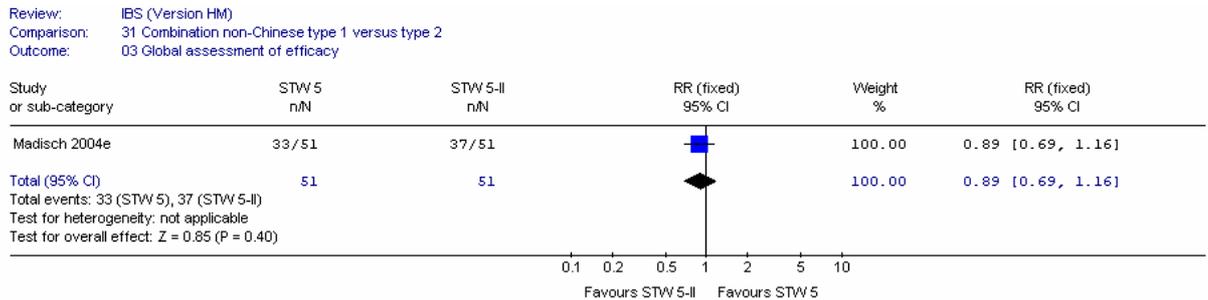
II. Individual symptoms

a) Pain

One study (Madisch 2004) reported a pain score on a scale of 0 to 21. There was no significant
difference between interventions.

Figure 40: Pain score
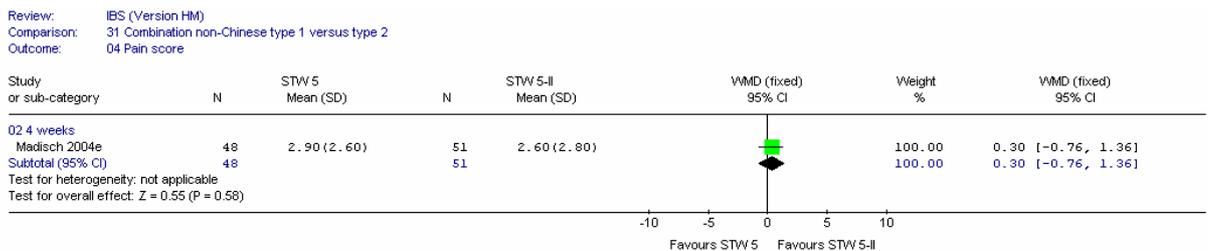
b) Bowel habits

One study (Madisch 2004) recorded the number of people with complete relief or absence of changes in bowel habit (constipation, diarrhoea or alternating bowel habit). There was no significant difference between interventions.

Figure 41: Number of people with no changes in bowel habit

Review: IBS (Version HM)
Comparison: 31 Combination non-Chinese type 1 versus type 2
Outcome: 04 Pain score

Study or sub-category	N	STW 5 Mean (SD)	N	STW 5-II Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
02 4 weeks							
Madisch 2004e	48	2.90(2.60)	51	2.60(2.80)		100.00	0.30 [-0.76, 1.36]
Subtotal (95% CI)	48		51			100.00	0.30 [-0.76, 1.36]
Test for heterogeneity: not applicable							
Test for overall effect: Z = 0.55 (P = 0.58)							

-10 -5 0 5 10
Favours STW 5   Favours STW 5-II
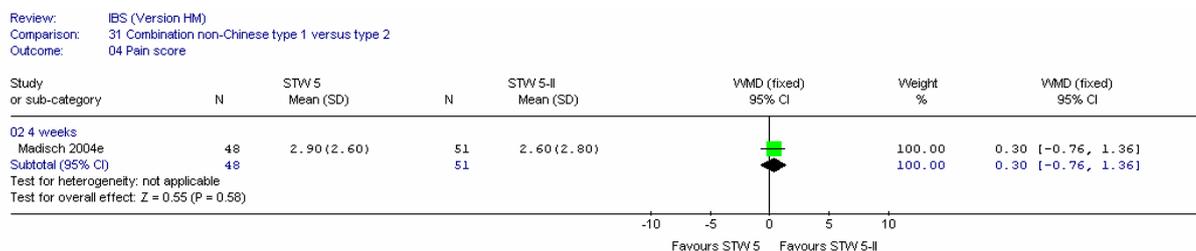
III. Adverse effects

One study reported tolerability (Madisch 2004). There was no significant difference between interventions.

Figure 42: Tolerability

Review: IBS (Version HM)
Comparison: 31 Combination non-Chinese type 1 versus type 2
Outcome: 07 Tolerability 'good' or 'very good'

Study or sub-category	STW 5 n/N	STW 5-II n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
Madisch 2004e	50/51	46/51		100.00	1.09 [0.99, 1.20]
Total (95% CI)	51	51		100.00	1.09 [0.99, 1.20]
Total events: 50 (STW 5), 46 (STW 5-II)					
Test for heterogeneity: not applicable					
Test for overall effect: Z = 1.66 (P = 0.10)					

0.1 0.2 0.5 1 2 5 10
Favours STW 5-II   Favours STW 5

EVIDENCE STATEMENTS

1. There is inconsistent evidence to show if Chinese combination herbs give an improvement in global symptoms compared with placebo.

2. There is moderate evidence to show there is no significant effect of combination Chinese herbal medicines on pain, bowel habits or quality of life.

3. There is moderate evidence to show no significant difference between single non-Chinese herbal medicines and placebo in global symptoms or in pain score, but inconsistency or uncertainty in the bloating and bowel habit outcomes respectively.

4. There is moderate evidence to suggest that combination non-Chinese herbal medicines give a significantly greater improvement in global symptoms, compared with placebo, however, there was inconsistency or uncertainty in the pain and bowel habit outcomes respectively.

5. There is moderate evidence to suggest that combination non-Chinese herbal medicines give a significantly greater improvement in global symptoms, pain and tolerability, but no significant difference in bowel habits, compared with single non-Chinese medicines.

6. There is moderate fair evidence to suggest that there is no significant difference between individualised combination Chinese medicines and standard formula Chinese medicines, in global symptoms.

7. There is limited evidence to show that there is no significant difference between two single non-Chinese herbs, curcuma and fumitory, in global symptoms, pain, bloating, tolerability or psychological distress.

8. There is moderate evidence to show that there is no significant difference between two combination non-Chinese herbs, STW5 and STW-II, in global symptoms, pain, bowel habits, tolerability or psychological distress.

EVIDENCE TO RECOMMENDATION

The review evidence suggests that some herbal preparations may be clinically effective in people with IBS and are well tolerated. However, the GDG believed there were too many uncertainties regarding type and dose of herbal medicines to make a recommendation for practice, and proposed that these interventions should be investigated further in a research recommendation.

11 Psychosocial interventions: patient information and support groups

<div style="border:1px solid">

Clinical Questions

1. Do psychosocial interventions have a role in managing IBS symptoms?

2. Do self help/support groups have a role in managing IBS symptoms?

3. Information determines patient experience by facilitating informed choices.

4. What role does patient information play in IBS?

</div>

BACKGROUND

Psychosocial factors contribute to the predisposition to IBS; some authors believe it the most important factor in terms of who manifests IBS, how severe it becomes and how people with IBS cope with managing the condition. The multifaceted nature of IBS requires an appreciation and understanding of psychosocial principles that relate to the disorder and the way these may be addressed in treatment strategies. It is important to explore the possible indicators of psychological distress which may affect the ways in which a person with IBS presents their condition and the associated coping behaviours. A history of physical and sexual abuse is twice more common in people first presenting with IBS than in people with organic gastrointestinal disease prior to a definitive diagnosis. Anxiety and other major life stress and/or trauma have been shown to correlate with the development of IBS and the severity of its symptoms (Gunn 2003; Camilleri 2001; Jones 2000). The presence of psychosocial factors is also an indicator for the likelihood that people will seek medical attention for IBS as well as other medical conditions.

Addressing psychosocial factors with an ongoing collaborative multi-disciplinary approach leads to improvement in the clinical outcomes and while psychosocial factors do not cause IBS symptoms, they do influence the patients' response both to the condition and treatment (Gaynes 1999).

Support groups and patient information

People with IBS often experience a sense of frustration, isolation, and a need to identify a niche in the health/sick role continuum. Frustration may arise from their perceived inability to control symptoms, prevent episodes, identify episode triggers, and obtain medical validation of the condition. Many people with IBS consider their condition to be severe and greatly affecting their lives. They feel that some health care professionals do not give credence to IBS as a chronic debilitating condition and that information which may help them understand more about how to live with the condition is often not forthcoming. The constant anticipation of the next IBS episode, the nature of the bowel symptoms, the

requirement for quick and easy access to toilet facilities, often results in embarrassment and withdrawal from social activities with resultant isolation (Bertram 2001). Providing people with IBS appropriate information about their disorder may promote a strong physician-patient interaction and may reduce healthcare use. Most people with IBS feel insufficiently informed, particularly in relation to risk of serious disease and the role of diet (Dancey 1993; O'Sullivan 2000).

The isolation people with IBS experience in many aspects of their lives may also be addressed by the use of support groups. Support groups have been suggested as a way for people to help one another by having the opportunity to discuss coping strategies with others who are experiencing similar issues. However provision and access to IBS support groups may vary throughout the UK.

"*There is no self help group near to me at the moment, but maybe that will happen in due course. I am sure that to talk with fellow sufferers must be a great help. So many people are striving to get the better of this awful affliction without much luck, but, ever the optimist, I shall continue to look for an answer.*"

This anonymous quote is not atypical of the IBS patient experience, and to address these concerns through effective diagnosis and management interventions is an essential aspect of this clinical guideline. Support group discussion may provide people with an opportunity to share others' difficulties with IBS, which may affirm their own IBS experiences. They may be relieved to finally be able to discuss their problems and symptoms with others who understand the challenges.

People appear to cope better with this chronic illness if they have sufficient information about IBS and appropriate support networks. Within the context of the whole IBS patient journey, evidence suggests that an important feature of effective coping and improved quality of life is for people to take responsibility through shared management with their primary care clinician (Kennedy 2003; Lacy 2007; Rogers 2007). Therefore, due consideration of the information needs of people with IBS is fundamental to the provision of effective management strategies.

Review of the literature for psychosocial interventions, support groups and patient information led to two reviews, one on support groups and self help, and the other on patient information. There was no evidence for other psychosocial interventions. The support groups and patient information reviews are closely linked: a common theme is the investigation of the effectiveness of a guidebook giving patient information. The two reviews are presented in sections 11.1 and 11.2; section 11.3 describes the process of evidence to recommendation for both reviews, leading to a single recommendation.

11.1 Support Groups and Self Help

SELECTION CRITERIA

The selection criteria described in the general methodology section were used, except that crossover studies were excluded as inappropriate due to the carry-over effect of the interventions.

The following comparisons were to be included:

- Support group versus waiting list control
- Support group plus other intervention versus other intervention only
- Support group versus other intervention.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library* (1966 to current day with guidance from the GDG). Additionally, the PSYCINFO database was searched for this review. The search strategies are given in Appendix B.

The titles and abstracts of these studies were assessed. Three studies were identified as being potentially relevant to the review and these papers were retrieved in full. The reference lists for each of the retrieved studies were inspected for further potential papers, but none were identified. The one excluded study is listed in Appendix E, along with reasons for exclusion.

CHARACTERISTICS OF INCLUDED STUDIES

Study Design

Two randomised trials were found (Payne 1995; Robinson 2006). One was carried out in the USA (the setting was not stated) (Payne 1995); the other was carried out in primary care in the UK (Robinson 2006).

Population

The 34 patients (5 men and 29 women) in the Payne (1995) study had IBS satisfying the Rome I criteria; their mean age was around 40 years (range 22 to 70 years); 29/34 had an Axis I disorder.

The 420 patients (50 men and 370 women) in the Robinson (2006) study had IBS, of whom 38% satisfied Rome II criteria (the rest diagnosed by GP or specialist if they had previously been referred); patients were excluded if they were unable to read or understand English; their mean age was 40 years (SD 14.4 years); psychiatric co-morbidities were not stated. They had had bowel symptoms for a mean of 6 years (SD 7.2 years).

Neither of the studies reported the number of participants with bloating or whether the symptoms were post-infective.

Interventions and comparisons

Payne (1995) compared three groups: 1) support group; 2) cognitive behavioural therapy; 3) waiting list control for 8 weeks. The self-help group intervention involved guided discussion on aspects of IBS, for example, stress and diet, for 1 hour 15 minutes per week for 8 weeks.

Robinson (2006) compared three groups: 1) self-help support group plus educational guidebook; 2) guidebook only; 3) usual care. The self-help guidebook included information on: lifestyle; diet; drugs, and; alternative therapies. The self-help meeting was a one-off 2-hour meeting of 8 to 12 people at a time; only 59 of 139 attended. The study carried out some additional qualitative research and noted that some people were unwilling to discuss bowel related symptoms with strangers, which may have been the cause of the poor attendance rate. The control group had usual care at the discretion of the primary care physician. Data were collected at one year.

People in Payne (1995) continued to take their medication unchanged. Participants in Robinson (2006) were informed that they were free to continue to visit their primary care physician without restriction.

Comparisons were:

- Support group versus waiting list control for 8 weeks
- Self-help support group plus educational guidebook versus guidebook only, followed at one year
- Guidebook only versus usual care, followed at one year
- Support group versus cognitive behavioural therapy (this is reported in the CBT review), at 8 weeks.

Outcomes

The outcomes examined were:

1. Global symptoms:
   a) Global improvement in symptoms (number of patients) (Payne 1995)
   b) Global symptom score on a 7-point scale from unbearable to no symptoms (i.e. higher score is better) (Robinson 2006)
   c) Global improvement of IBS symptoms (mean Composite Primary Symptom Reduction [CPSR] score; CPSR represents the proportional reduction in score from baseline); i.e. high = bad (Payne 1995)
   d) Global assessment of treatment on symptoms (Robinson 2006).

2. Mental health outcomes (overall mental health; depression; anxiety)

Overall anxiety and psychological distress (Anxiety, State-Trait Anxiety Inventory [STAI]); Scale range 20 to 80; high = bad (Payne 1995)

Beck depression inventory (scale maximum 63; high=bad) (Payne 1995).

METHODOLOGICAL QUALITY

The quality assessment for included trials is shown in Appendix D.

The method of randomisation was adequate in Robinson (2006), which stated that it used a randomisation system based on minimisation. Allocation concealment was adequate in one study using a central telephone randomisation system (Robinson 2006) and not stated in the other (Payne 1995). The patients were not blinded (because of the type of intervention). Neither study reported an *a priori* power calculation. Payne (1995) demonstrated baseline comparability between the groups; this was not stated in Robinson (2006). All the participants were followed up in Payne (1995), while in Robinson (2006) data were missing for 56 patients overall (13%). Overall, neither study was considered to be at risk of bias.

RESULTS

A. Support group versus waiting list control

1) Global symptoms

Global improvement of IBS symptoms was reported by Payne (1995) at 8 weeks, in 22 people, using the mean Composite Primary Symptom Reduction [CPSR] score; CPSR represents the proportional reduction in score from baseline. The study gave individual patient data, allowing calculation of standard deviations. There was no significant difference between interventions, but the confidence interval was fairly wide.

Figure 1:


```
Review:       IBS 31 August 2006 (Version workotherpysch)
Comparison:   28 Support group versus waiting list control
Outcome:      01 Global improvement in symptoms (CPSR)

Study                        Support group         Control          WMD (fixed)      Weight      WMD (fixed)
or sub-category        N      Mean (SD)       N    Mean (SD)           95% CI           %           95% CI

Payne 1995            12      0.31(0.42)      10   0.10(0.23)                          100.00    0.21 [-0.07, 0.49]

Total (95% CI)        12                      10                                       100.00    0.21 [-0.07, 0.49]
Test for heterogeneity: not applicable
Test for overall effect: Z = 1.49 (P = 0.14)
                                                              -1   -0.5   0   0.5   1
                                                          Favours control   Favours support gp
```
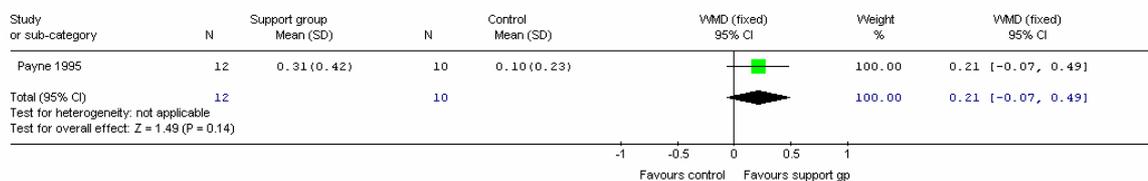
2) Mental health outcomes

a) Beck Depression Inventory

Payne (1995) reported Beck Depression Inventory scores (scale maximum 63; high=bad) at the end of treatment (8 weeks). There was no significant difference between interventions.
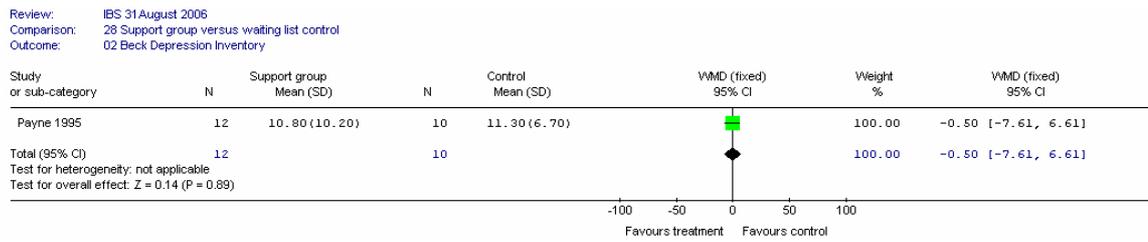
Figure 2:

b) Overall anxiety and psychological distress: State-Trait Anxiety Inventory (STAI)

Payne (1995) reported STAI scores (20 to 80; high=bad) at the end of treatment (8 weeks).
Again there was no significant difference between interventions.

Figure 3:
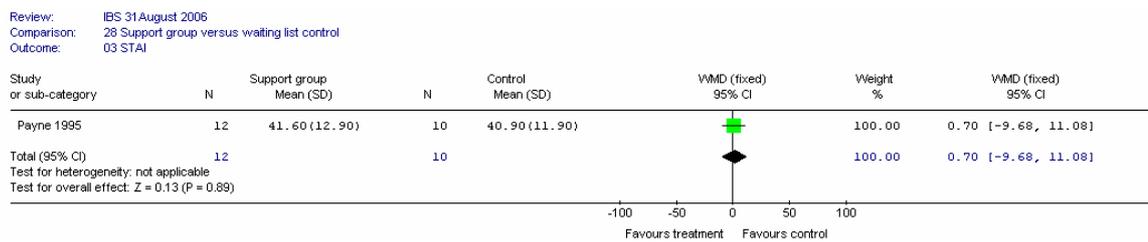
B. Self-help support group plus educational guidebook versus guidebook only

1. Global symptoms

Robinson (2006) reported a global IBS symptom score on a 7-point scale from unbearable to
no symptoms (i.e. higher score is better) at the 52-week follow-up in 247 patients. There was
no significant difference between interventions. We note, however, that only 59 of the 139
participants attended the support group meeting (although the global symptoms scores were
still reported).

Figure 4:
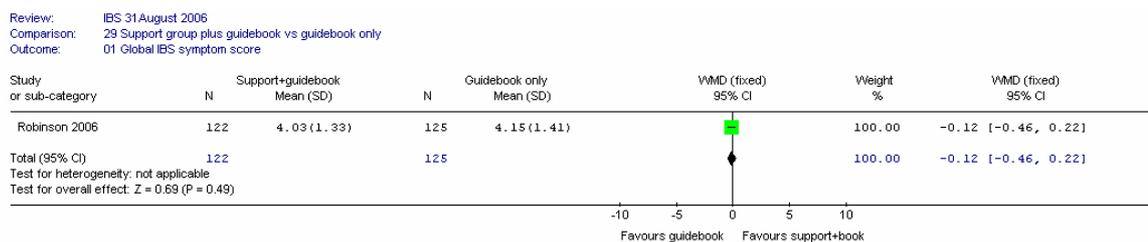
People also reported their assessment of treatment on symptoms (global improvement of
symptoms score) on a 7-point scale from very much worse to very much improved (i.e. higher
score is better). There was no significant difference between interventions.

Figure 5:

Review: IBS 31 August 2006
Comparison: 29 Support group plus guidebook vs guidebook only
Outcome: 02 Global improvement symptom score

Study or sub-category	N	Support+guidebook Mean (SD)	N	Guidebook only Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Robinson 2006	122	5.06(1.27)	125	5.14(1.20)		100.00	-0.08 [-0.39, 0.23]
Total (95% CI)	122		125			100.00	-0.08 [-0.39, 0.23]

Test for heterogeneity: not applicable
Test for overall effect: Z = 0.51 (P = 0.61)

-10   -5   0   5   10
Favours guidebook   Favours support+book

C. Guidebook only versus usual care

1) Global symptoms

Robinson (2006) reported a global IBS symptom score on a 7-point scale from unbearable to no symptoms (i.e. higher score is better) at the 52-week follow up in 242 patients. There was no significant difference between interventions.

Figure 6:

Review: IBS 31 August 2006
Comparison: 31 Guidebook versus usual care
Outcome: 01 Global IBS symptom score

Study or sub-category	N	Guidebook Mean (SD)	N	Usual care Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Robinson 2006	125	4.15(1.41)	117	3.93(1.31)		100.00	0.22 [-0.12, 0.56]
Total (95% CI)	125		117			100.00	0.22 [-0.12, 0.56]

Test for heterogeneity: not applicable
Test for overall effect: Z = 1.26 (P = 0.21)

-4   -2   0   2   4
Favours usual care   Favours guidebook

Patients also reported their assessment of treatment on symptoms (global improvement of symptoms) on a 7-point scale from very much worse to very much improved (i.e. higher score is better). There was a statistically significant difference, in favour of the guidebook.

Figure 7:

Review: IBS 31 August 2006
Comparison: 31 Guidebook versus usual care
Outcome: 02 Assessment of treatment on symptoms

Study or sub-category	N	Guidebook Mean (SD)	N	Usual care Mean (SD)	WMD (fixed) 95% CI	Weight %	WMD (fixed) 95% CI
Robinson 2006	125	5.14(1.20)	117	4.63(1.41)		100.00	0.51 [0.18, 0.84]
Total (95% CI)	125		117			100.00	0.51 [0.18, 0.84]

Test for heterogeneity: not applicable
Test for overall effect: Z = 3.02 (P = 0.003)

-4   -2   0   2   4
Favours usual care   Favours guidebook

The results from Robinson (2006) suggest that the guidebook may have helped patients, with little additional benefit from the support group (a single 2-hour meeting, which only 59 of the 139 patients in this randomisation arm actually attended).

ECONOMIC LITERATURE FOR PSYCHOSOCIAL INTERVENTIONS / SUPPORT GROUPS

One relevant health economic analysis was identified on the cost-effectiveness of psychosocial interventions or support groups in the management of IBS. Robinson (2006) was a trial based economic analysis looking at the impact of two self help interventions (a guidebook and a self-

help group session) on clinical and economic outcomes in primary care patients with IBS. Only the economic outcomes are described here as the clinical effectiveness results have been described in the clinical effectiveness review. The economic outcomes reported were GP visits, hospital consultation rates, prescription costs and overall costs.

This study provided evidence that the provision of a self-help guidebook reduced GP visits (-1.56 visits per annum, P<0.001), hospital visits (-0.22 visits per annum, p=0.038) and prescription costs (£24, p=0.031) but the addition of a self-help group session did not further reduce resource use. Overall costs for GP visits, hospital visits and prescribed drugs were reduced for those who received the guidebook (-£73, 95%CI -£43 to -£103, p<0.001) but there was no significant effect on overall costs from the addition of the self-help session. The guidebook was also associated with a significant increase in the use of self-care activities such as dietary interventions and relaxation therapy.

This study was a partial economic evaluation as it did not assess the incremental cost of any benefit achieved by the provision of a guidebook or self-help group session in the form of a cost-effectiveness ratio. A particular limitation of this study was the failure to include intervention costs for the guidebook or self-help group. This limitation would only affect the conclusions drawn from this study if the cost of providing the guidebook exceeded the cost-savings resulting from reduced resource use in patients provided with a guidebook. The evidence provided by this study was considered relevant to the guideline as it considered both costs and effects for the intervention in an appropriate population and setting. No potential areas of significant bias were identified except for the exclusion of intervention costs. Whilst this study did not provide a full economic analysis of the provision of guidebooks, with or without a self-help group session, this evidence was considered alongside the clinical effectiveness evidence to inform recommendations on the use of self-help groups and self-help information in the management of IBS.

EVIDENCE STATEMENTS
1. There was limited evidence to show no significant difference in global improvement of symptoms or in depression on the Beck inventory for a self help group intervention, involving guided discussion on aspects of IBS, for example, stress and diet, for 1hour 15 minutes per week for eight weeks, compared with waiting list control.

2. There was good evidence to show no significant additional effect on global symptoms of a single two-hour self help meeting of 8 to 12 people at a time, in people already receiving a guidebook. It is noted that less than half the people attended the self help meeting.

3. There was good evidence to show no significant additional effect on the number of primary care consultations and hospital visits of a single two-hour self help meeting of 8 to 12 people

at a time, in people already receiving a guidebook. It is noted that less than half the people attended the self help meeting.

4. There was good evidence to show no significant additional effect of a single two-hour self help meeting of 8 to 12 patients at a time, in patients already receiving a guidebook, on the overall cost of GP visits, hospital visits and prescription drugs. It is noted that less than half the people attended the self help meeting.

5. There was good evidence to show a significant improvement in global symptoms for people receiving a self-help guidebook, which included information on lifestyle, diet, drugs and alternative therapies, in comparison with usual care.

6. There was good evidence to show a significant decrease in the number of primary care consultations and hospital visits for people receiving a self-help guidebook, which included information on lifestyle, diet, drugs and alternative therapies, in comparison with usual care.

7. There was good evidence to show a significant reduction in the overall cost of GP visits, hospital visits and prescription drugs for people receiving a self-help guidebook, which included information on lifestyle, diet, drugs and alternative therapies, in comparison with usual care.

GDG DISCUSSION

The GDG commented that people may not attend support groups because of travel difficulties (due to lack of control of symptoms) and a general reluctance to discuss bowel problems with others. The superior effect of the guidebook compared to usual care was not surprising, and the GDG noted that a simple guide to IBS has proved popular in the past. It was also noted that many people with IBS do make a great effort to attend support groups and those that attend find these beneficial.

11.2 Patient Information

OBJECTIVE

To review the evidence on the information needs of people who have been diagnosed with IBS, assessing the impact that information has in their self management of the syndrome and their ability to maximise quality of life when living with the syndrome.

CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

Types of studies

Quantitative (RCTs, prospective studies, survey) and qualitative (eg. focus group) study designs were considered for this review.

SEARCH STRATEGY FOR IDENTIFICATION OF STUDIES

Searches were performed on the following core databases: MEDLINE, EMBASE, CINAHL, and *The Cochrane Library.* Searches were performed from the beginning of each database and updated to June 2007. The search strategies are given in Appendix B:

Following sifting, five studies were included in the review.

CHARACTERISTICS OF CLINICAL STUDIES INCLUDED IN THE REVIEW

Five studies were included in this review: three were prospective studies (Kennedy 2003, Lacy 2007, Bogalo 2006) and two were separate papers from a randomised trial (Robinson 2006, Rogers 2007). One study was excluded and is given in Appendix E.

Prospective non-randomised studies

The Lacy (2007) study used a questionnaire that addressed two main domains:

- Participant knowledge of IBS (epidemiology and natural history; aetiology; symptoms; diagnosis and treatment)
- Participant attitudes towards IBS (relationships of IBS to functional status; concerns and fears about IBS; ability of the medical system to address patients' needs).

People with IBS were identified from a search of the medical records in Lebanon, New Hampshire, USA and the records were examined to ensure that the participants met the Rome II criteria. 261 of 664 contacted (39%) returned the questionnaires.

Bogalo (2006) was a prospective study of 31 people assigned to the treatment group in an RCT. These participants received a self-help treatment manual over six weeks, with one chapter per week. Each chapter was task oriented. Topics covered in each of the six weeks were: IBS explained; assessing symptoms and self monitoring; managing IBS symptoms; cognitive restructuring, personal expectations and activity patterns; relaxation and stress management; and maintenance. The study investigated the hypothesis that treatment group participants who had a higher level of engagement in the homeworking tasks would experience greater relief from their IBS symptoms.

The Kennedy (2003) study used focus groups of 12 people to explore participants' knowledge and experience of IBS. Participants were recruited from an article in a regional paper asking for volunteers. Focus group meetings were held over a two week period and lasted 1¼ to 1½ hours. Each session was taped and transcribed. Transcripts were read and analysed using a framework developed for the study. Four main areas were outlined (perceptions and expectations; experience of IBS; information needs and sources; managing IBS) and these were divided into 16 subheadings. Each comment in the focus group was allocated to one of the sub-headings. The patients' views and experiences were used in the development of a self-help guidebook.

Randomised studies

The Robinson (2006) study measured the clinical and cost effectiveness of three interventions: the Kennedy self-help guidebook; the guidebook together with a two-hour one-off support group meeting; and usual care. It is noted that less than half of the people in the support group arm attended the support group meeting. This study has been reported and discussed in the Support groups review (section 11.1)

Rogers (2007) is a report of a qualitative study of a purposefully selected group of 12 of the Robinson (2006) trial patients: four of these participants had received the guidebook only; four had received the guidebook and had attended the support group meeting; one who had received the guidebook but did not attend the meeting; and three control group participants. Interviews were carried out with the participants, lasting between 40 and 90 minutes. These were transcribed and transcripts were analysed thematically, against one another by constant comparison. Key themes included: the lived experience of IBS (impact on everyday life; experience of symptoms; and reaction of others); ways of managing; lay epidemiology; experience of medical management and diagnosis; alternative help-seeking views about medication; the guidebook as projected identification with others; use of the guidebook, together with perceived changes; and continuity from being part of the trial.

Quantitative and qualitative narrative review was carried out to assimilate the evidence on the reported benefits of information in enabling people with IBS to better understand their condition and make lifestyle adjustments. A thematic analysis was carried out across included studies.

RESULTS

From the qualitative study data, several distinct themes emerge. These are:

- There is a lack of clear information to support people with IBS, which creates misunderstanding and misconception
- People with IBS are often misinformed at the point of diagnosis, not fully understanding the diagnosis and its potential impact on quality of life
- Medical management is often one-dimensional, with no attention given to lifestyle and other therapeutic interventions
- Developing coping mechanisms are augmented by structured information. Patient experience was improved through exposure to a guidebook which focused on self-management in partnership with the primary care clinician
- Structured information provides an instant source of help, sensitive to the episodic nature of IBS, facilitating the sharing of experience and ongoing IBS management
- Patient information is essential for shared decision-making and partnership between clinicians and people with IBS
- Information should be patient-centered with involvement of the person with IBS

- Primary care clinicians should take responsibility to ensure that their knowledge of IBS enables expressed concerns to be answered, offering support through clear explanation
- People with IBS need information relating to cause, cure and long-term prognosis for IBS
- Appropriate sources of information for people with IBS include magazine articles, leaflets in shops, books, support groups, internet, medical journals and books.

Information needs

The importance of good information is highlighted in this review of published papers relating to the information needs, experiences, and quality of life issues for people with IBS. Given the chronic nature of IBS, information quality will contribute to the development of an effective shared decision making model between primary care clinician and the person with IBS. The primary care clinician is a key resource for the person who presents with IBS, and following positive diagnosis the importance of information sharing and shared decision making relating to their symptom profile and treatment response is a key aspect.

Respondents in the Kennedy study (2003) highlighted that primary care clinicians had little knowledge about IBS, and subsequently were unable to provide much in the way of clear information that encouraged self-help strategies for people with IBS. Magazine articles, leaflets, books, support groups, internet sources and published journal articles were highlighted as useful. The inclusion of patients in producing patient information emerged as a contemporary theme within this study, this is consistent with NICE's approach to developing patient information, and is produced as part of this clinical guideline suite of information. Lacy (2007) supports points raised for discussion from the Kennedy study, highlighting that specific guidance on diagnosis, treatment, misconceptions about natural history of the disease and subsequent confusion should all feature in prepared information.

Information clearly has an educative role, in correcting inappropriate concerns relating to cancer, and in developing the shared care model that provides clarity in symptom based presentation and subsequent treatment interventions that are appropriate for that particular person with IBS.

The Robinson (2006) study found that use of the self-help guidebook compared with usual care resulted in fewer primary care consultations, and a greater improvement in global symptoms. Well prepared information appears to be at least cost neutral, with a cost per patient reduction reported as £73. This aspect is discussed further in the support groups review (section 11.1)

Understanding the importance of patients as agents of change is a key aspect to effective implementation of evidence, where patients become the drivers for change in healthcare behaviours. The success of the patient self-help booklet reported in the Kennedy study (2003) is a clear example of the importance of patient information, in meeting the needs of people learning to live with IBS or adapting lifestyle for those people who have lived with IBS for a significant period of time.

Quality of life

Chronic illness remains a significant challenge to the individual in terms of effective coping, and to the NHS in identifying the appropriate level of support to that individual. In this review, quality of life issues are raised consistently in studies that ask people with IBS questions relating to the level of impact that IBS has on their daily living activities. People often express the desire for cure and information relating to long term prognosis, reflecting the over-medicalised language relating to effective coping. Living with IBS is the challenge, and symptom based management relating to the quality of life experience is key in the shared care model. In the Lacy (2007) study, nearly all participants (n=261) reported that IBS affected their lives in some way. Clearly this relates to severity and quality of life when considered as a continuum, and one can see the person moving from coping to not coping, reflecting the episodic nature of the syndrome.

The role of information and its added value in addressing misconceptions, diagnosis, providing reliable answers to questions, treatment interventions and indicating when access to a primary care clinician should be considered due to continued worsening of symptoms are all aspects highlighted within this clinical guideline. If addressed, collectively they should provide a foundation for the patient to develop effective coping strategies.

GDG DISCUSSION

Information should be clear, concise and relate to the symptom-based management of the syndrome. It should deal with areas of misconception, embarrassment and quality of life issues. This should be provided at the earliest opportunity by the primary care clinician following positive diagnosis of IBS.

The role of well prepared information provides the basis for the development of the shared care model between the primary care clinician and the person with IBS. Evidence supports the use of information booklets that encourage self help activity.

EVIDENCE STATEMENTS

1. There is fair evidence indicating that if people with IBS receive directed information and encouragement to be actively involved in the management of their condition that this contributes to:
   - A positive impact on treatment outcomes
   - An improvement in quality of life perception and reduction in symptom severity
   - Reduction in primary care consultations.

2. There is weak evidence indicating that people with IBS lack appropriate information relating to their condition. This can lead to misunderstanding and reduced quality of life experiences.

11.3 Evidence to Recommendation

The GDG took into consideration the evidence in both reviews. This included the clinical and cost effectiveness results from the quantitative studies, and the qualitative analysis in the patient information review, which was supported by the experiences of all members of the GDG. The evidence from the support groups review showed improved clinical outcomes and reduced health care costs for people provided with a self-help guidebook, and indicated that providing self-help information in the management of IBS is likely to be clinically and cost-effective. There did not appear to be an additional effect from a support group meeting, but this may have been confounded by the poor attendance. Some members of the GDG had extensive experience of self-help groups and reported that they are a current voluntary sector patient provision, providing social support, and that people with IBS comment positively on their involvement in self help organisations. The GDG decided not to make a recommendation on the usefulness of support groups.

Based on the qualitative data in the patient information review and their own experience, the GDG recognised the need for clear unambiguous information, and indicated its role and importance in helping people with IBS develop coping strategies in partnership with their primary care clinician. The GDG noted that patient information should be provided early in the patient pathway following a positive diagnosis of IBS, and that self-help information should be wide ranging, covering areas such as general lifestyle, physical activity, diet and symptom-targeted medication.

RECOMMENDATION

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.

---

12   RECOMMENDATIONS FOR RESEARCH

1. Tricyclic antidepressants, SSRIs and SNRIs

Are low-dose tricyclic antidepressants (TCAs), SSRIs and SNRIs effective in the treatment of IBS as a first line therapy, and which is the more effective and the safer option?

Why is this important?

Reviews have shown that TCAs and SSRIs have each been compared with placebo, but not at low dose. In practice, TCAs are used at higher doses and concordance with treatment is poor because of side effects. GDG clinicians believe that at low doses (e.g. 5 to 10 mg equivalent of amytriptyline), TCAs could be the treatment of choice, but there is a lack of evidence. Newer antidepressants, SNRIs, maybe useful in the treatment of IBS pain. A large randomised trial is proposed, comparing an SSRI, a TCA, an SNRI and placebo. Participants should be adults with a positive diagnosis of IBS, stratified by type of IBS and randomised to treatments. The primary outcome should be global improvement in IBS symptoms. Health related quality of life should also be measured. Adverse effects should be recorded. Study outcomes should be assessed at 12, 26 and 52 weeks after the start of therapy.

2. Psychological interventions

Are psychological interventions (psychological therapy, hypnotherapy and CBT) equally effective in the management of IBS symptoms, either as first line therapies in primary care, or in the treatment of people with IBS that is refractory to other treatments?

Why is this important?

Reviews show some evidence of effect when comparing psychological interventions with control, mainly in people with refractory IBS. Many trials are small in size. The psychological interventions of psychological therapy, hypnotherapy and CBT are thought to be useful in helping people with IBS cope with their symptoms, but it is unclear at what stage these should be given, including their use as first line therapies in primary care. A large randomised trial is proposed, comparing CBT, hypnotherapy and psychological therapy (psychodynamic interpersonal therapy). Participants should be adults with a positive diagnosis of IBS, and they should be stratified into those with and without refractory IBS and then randomised to treatments. The primary outcome should be global improvement in IBS symptoms. Health related quality of life should also be measured. Adverse effects should be recorded. Study outcomes should be assessed at 12, 26 and 52 weeks after the start of therapy.

3. What factors contribute to refractory symptoms in IBS?
Why is this important?

Most individuals with IBS experience symptoms that are relatively short lived or only trouble them on an intermittent basis. Some people, however, develop chronic and severe symptoms

that are difficult to treat. There are relatively few prospective studies that have investigated this problem.

A large, prospective, population based cohort study is proposed, which would evaluate people in the community with IBS symptoms, according to measures of bowel symptomatology, physical symptom profile, psychological symptoms, childhood adversity, past history of psychiatric disorder, social supports, quality of life and other relevant potential predictors. Individuals would be re-evaluated 12 and 24 months later using similar measures. Baseline variables would be used to predict chronicity of symptoms, quality of life and healthcare utilisation at 12 months and at 24 months.

4. Relaxation and biofeedback

What is the effect of relaxation and biofeedback therapies on IBS symptoms and patient-related outcomes?

Why is this important?

Reviews of biofeedback and relaxation therapies suggest a positive effect on the control of IBS symptoms, but evidence is limited and not sufficient to make recommendations. Patient representation within the group supports this view, from a personal and anecdotal perspective.

Recent developments in computer-aided biofeedback methods merit investigation. A large randomised trial is proposed to compare relaxation therapy, computer-aided biofeedback therapy and attention control in primary care. Participants should be adults with a positive diagnosis of IBS, and they should be stratified into those with and without refractory IBS, and then randomised to treatments. The primary outcome should be global improvement in IBS symptoms. Health related quality of life should also be measured. Adverse effects should be recorded. Study outcomes should be assessed at 12, 26 and 52 weeks after the start of therapy. Qualitative data should be generated relating to how people with IBS perceive their IBS condition.

5. Herbal medicines

Are Chinese and non Chinese herbal medicines safe and effective as first-line therapy in the treatment of IBS, and which is the more effective and safer option?

Why this is important

Reviews of herbal medicines suggest a positive effect on the control of IBS symptoms, but evidence is limited and not sufficient to make recommendations (8 comparisons from the six trials provides heterogeneous data which are very difficult to interpret). A large randomised placebo controlled trial is proposed, comparing Chinese and non Chinese herbal medicine, single and multiple compounds, which should be available within the UK as standard preparations. Participants should be adults with a positive IBS diagnosis and should be

stratified by IBS type and then randomised to treatments. The primary outcome should be global improvement in IBS symptoms, with symptom scores being recorded using a validated scale. Health-related quality of life should also be measured, and adverse events recorded. Study outcomes should be assessed at 12, 26 and 52 weeks post-intervention.

13   IMPLEMENTATION OF THE GUIDELINE

The Healthcare Commission assesses the performance of NHS organisations in meeting core and developmental standards set by the Department of Health in 'Standards for better health', issued in July 2004. Implementation of clinical guidelines forms part of the developmental standard D2. Core standard C5 says that national agreed guidance should be taken into account when NHS organisations are planning and delivering care.

NICE has developed tools to help organisations implement this guidance (listed below). These are available on the NICE website (www.nice.org.uk/CG61).

- Slides highlighting key messages for local discussion.
- Costing tools:
    - Costing report to estimate the national savings and costs associated with implementation
    - Costing template to estimate the local costs and savings involved.
- Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
- Audit criteria to monitor local practice.

14 RELATED NICE GUIDANCE

Published

Depression: Depression, management of depression in primary and secondary care. NICE clinical guideline 23 (2004). Available from www.nice.org.uk.

Referral for Suspected Cancer: Referral Guidelines for Suspected Cancer in Adults and Children. NICE clinical guideline 27 (2005). Available from www.nice.org.uk.

Physical Activity. NICE public health intervention guidance PH1002 (2006). Available from www.nice.org.uk.

15 UPDATE OF THE GUIDELINE

NICE clinical guidelines are updated as needed so that recommendations take into account important new information. We check for new evidence 2 and 4 years after publication, to decide whether all or part of the guideline should be updated. If important new evidence is published at other times, we may decide to do a more rapid update of some recommendations.

16 REFERENCES

[No authors listed] (177) Irritable bowel syndrome: practitioner awareness, patient education can limit utilization, *Disease Management Advisor,* 8(12):185-7.

[No authors listed] (2000) Irritable bowel syndrome and women's health, *National Women's Health Report,* 22(3):1-2.

Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005] (2007) in: Higgins J and Green S, (editors) *The Cochrane Library, Issue 3, 2005.* Chichester, UK: John Wiley & Sons, Ltd.,

Abrahamsson H (1982) Irritable bowel syndrome. Diagnosis, *Scandinavian Journal of Gastroenterology - Supplement,* 79:20-3.

Adeniji OA, Barnett CB, and Di Palma JA (2004) Durability of the diagnosis of irritable bowel syndrome based on clinical criteria, *Digestive Diseases and Sciences,* 49(4):572-4.

Agrawal A and Whorwell PJ (2006) Irritable bowel syndrome: Diagnosis and management, *British Medical Journal,* 332(7536):280-3.

Akehurst RL, Brazier JE, Mathers N, O'Keefe C, Kaltenthaler E, Morgan A, Platts M, and Walters SJ (2002) Health-related quality of life and cost impact of irritable bowel syndrome in a UK primary care setting, *Pharmacoeconomics,* 20(7):455-62.

Aller R, de Luis DA, Izaola O, la Calle F, del Olmo L, Fernandez L, Arranz T, and Gonzalez Hernandez JM (2004) Effects of a high-fiber diet on symptoms of irritable bowel syndrome: a randomized clinical trial, *Nutrition,* 20(9):735-7.

Allison MC, Sercombe J, and Pounder RE (1988) A double-blind crossover comparison of lidamidine, loperamide and placebo for the control of chronic diarrhoea, *Alimentary Pharmacology and Therapeutics,* 2(4):347-51.

Amery W, Duyck F, Polak J, and van den BG (1975) A multicentre double-blind study in acute diarrhoea comparing loperamide (R 18553) with two common antidiarrhoeal agents and a placebo, *Current Therapeutic Research, Clinical and Experimental,* 17(3):263-70.

Arthurs Y and Fielding JF (1983) Double blind trial of ispaghula/poloxamer in the Irritable Bowel Syndrome, *Irish Medical Journal,* 76(5):253.

Asp NG (1992) Resistant starch. Proceedings from the second plenary meeting of EURESTA: European FLAIR concerted action no. 11 on physiological implications of the consumption of resistant starch in man. Preface, *European Journal of Clinical Nutrition,* 46(SUPPL. 2):S1.

Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW, Jr., Zaza S, and GRADE Working Group (2004) Grading quality of evidence and strength of recommendations, *British Medical Journal,* 328(7454):1490.

Atkinson W, Sheldon TA, Shaath N, and Whorwell PJ (2004) Food elimination based on IgG antibodies in irritable bowel syndrome: a randomised controlled trial., *Gut,* 53(10):1459-64.

Attar A, Lemann M, Ferguson A, Halphen M, Boutron MC, Flourie B, Alix E, Salmeron M, Guillemot F, Chaussade S, Menard AM, Moreau J, Naudin G, and Barthet M (1999) Comparison of a low dose polyethylene glycol electrolyte solution with lactulose for treatment of chronic constipation, *Gut,* 44(2):226-30.

Babic Z, Simek M, Altabas K, Bilic A, and Maras A (1997) Irritable bowel syndrome, *Croatian Journal of Gastroenterology and Hepatology,* 6(3-4):55-9.

Badia X, Mearin F, Balboa A, Baro E, Caldwell E, Cucala M, Diaz-Rubio M, Fueyo A, Ponce J, Roset M, and Talley NJ (2002) Burden of illness in irritable bowel syndrome comparing Rome I and Rome II criteria, *Pharmacoeconomics,* 20(11):749-58.

Barbara G, Stanghellini V, Brandi G, Cremon C, Nardo GD, De Giorgio R, and Corinaldesi R (2005) Interactions between commensal bacteria and gut sensorimotor function in health and disease, *American Journal of Gastroenterology,* 100(11):2560-8.

Barbezat G, Poulton R, Milne B, Howell S, Fawcett JP, and Talley N (2002) Prevalence and correlates of irritable bowel symptoms in a New Zealand birth cohort, *New Zealand Medical Journal,* 115(1164):U220.

Barbezat GO, Clain JE, and Halter F (1979) A double-blind trial of loperamide in the treatment of chronic diarrhoea, *South African Medical Journal,* Suid-Afrikaanse Tydskrif Vir Geneeskunde. 55(13):502-3.

Barnes J (1996) When to investigate and treat IBS symptoms, *Practitioner,* 240(1560):184-7.

Bellini M, Tosetti C, Costa F, Biagi S, Stasi C, Del Punta A, Monicelli P, Mumolo MG, Ricchiuti A, Bruzzi P, and Marchi S (2005) The general practitioner's approach to irritable bowel syndrome: from intention to practice, *Digestive and Liver Disease,* 37(12):934-9.

Bellini M, Tosetti C, Stasi C, Biagi S, Costa F, Bruzzi P, and Marchi S (2006) The general practitioner's management of patients with a new diagnosis of irritable bowel syndrome, *Journal of Clinical Gastroenterology,* 40(1):87.

Bennett P and Wilkinson S (1985) A comparison of psychological and medical treatment of the irritable bowel syndrome, *British Journal of Clinical Psychology,* 24(Pt 3):215-6.

Bensoussan A (1998) Treatment of irritable bowel syndrome with Chinese herbal medicine: a randomized controlled trial, *Journal of the American Medical Association,* 280(18):1585-9.

Bergeron CM (1984) A comparison of cognitive stress management, progressive muscle relaxation, and biofeedback in the treatment of irritable bowel syndrome, *Dissertation Abstracts International,* 44(10-B):3186.

Berthelot J and Centonze V (1981) Etude controlee en double aveugle Duspatalin (Mebeverine) contre placebo, dans le traitement du colon irritable., *Gazette Medicale De France,* 88(16):2341-3.

Bertram S, Kurland M, Lydick E, Locke GR, III, and Yawn BP (2001) The patient's perspective of irritable bowel syndrome, *Journal of Family Practice,* 50(6):521-5.

Bi L and Triadafilopoulos G (2003) Exercise and gastrointestinal function and disease: an evidence-based review of risks and benefits, *Clinical Gastroenterology and Hepatology,* 1(5):345-55.

Bijkerk CJ, de Wit NJ, Stalman WAB, Knottnerus JA, Hoes AW, and Muris JWM (2003) Irritable bowel syndrome in primary care: The patient's and doctors' views on symptoms, etiology and management, *Canadian Journal of Gastroenterology,* 17(6):363-8.

Bittner AC, Croffut RM, and Stranahan MC (2005) Prescript-assist probiotic-prebiotic treatment for irritable bowel syndrome: A methodologically oriented, 2-week, randomized, placebo-controlled, double-blind clinical study, *Clinical Therapeutics,* 27(6):755-61.

Blanchard EB and Schwarz SP (1987) Adaptation of a multicomponent treatment for irritable bowel syndrome to a small-group format, *Biofeedback & Self Regulation.,* 12(1):63-9.

Blanchard EB and Radnitz C. (1987) Psychological changes associated with self-regulatory treatments of irritable bowel syndrome, *Biofeedback & Self Regulation.,* 12(1):31-7.

Blanchard EB, Schwarz SP, Suls JM, Gerardi MA, Scharff L, Greene B, Taylor AE, Berreman C, and Malamood HS (1992) Two controlled evaluations of multicomponent psychological treatment of irritable bowel syndrome, *Behaviour Research and Therapy,* 30(2):175-89.

Blanchard EB, Greene B, Scharff L, and Schwarz-McMorris SP (1993) Relaxation training as a treatment for irritable bowel syndrome, *Biofeedback & Self Regulation.,* 18(3):125-32.

Block J, Myren J, Larssen SE, and Larsen S (1983) [Trimipramin (Surmontil) and placebo in irritable colon syndrome. A double-blind study], *Tidsskrift for Den Norske Laegeforening,* 103(11):903-5.

Boerner D, Eberhardt R, Metz K, and Schick E (1988) Wirksamkeit ind vertraeglichkeit eines antidepressivums beim colon irritablie, *Therapiewoche,* 38:201-8.

Bogalo L and Moss-Morris R (2006) The Effectiveness of Homework Tasks in an Irritable Bowel Syndrome Cognitive Behavioural Self-Management Programme., *New Zealand Journal of Psychology,* 35(3):120-5.

Bommelaer G, Dorval E, Denis P, Czernichow P, Frexinos J, Pelc A, Slama A, and El Hasnaoui A (2002) Prevalence of irritable bowel syndrome in the French population according to the Rome I criteria, *Gastroenterologie Clinique Et Biologique,* 26(12):1118-23.

Bouhnik Y, Neut C, Raskine L, Michel C, Riottot M, Andrieux C, Guillemot F, Dyard F, and Flourie B (2004) Prospective, randomized, parallel-group trial to evaluate the effects of lactulose and polyethylene glycol-4000 on colonic flora in chronic idiopathic constipation, *Alimentary Pharmacology and Therapeutics,* 19(8):889-99.

Bowling A (2002) *Research Methods in Health*, 2nd Edition. Maidenhead: Open University Press.

Bozzani A (2002) Managing patients with irritable bowel syndrome in primary care: Review of the specific literature, *Journal of Functional Syndromes,* 2(1):40-2.

Brinkhaus B, Hentschel C, Von Keudell C, Schindler G, Lindner M, Stutzer H, Kohnen R, Willich SN, Lehmacher W, and Hahn EG (2005) Herbal medicine with curcuma and fumitory in the treatment of irritable bowel syndrome: a randomized, placebo-controlled, double-blind clinical trial, *Scandinavian Journal of Gastroenterology,* 40(8):936-43.

Burden S (2001) Dietary treatment of irritable bowel syndrome: Current evidence and guidelines for future practice, *Journal of Human Nutrition & Dietetics,* 14(3):231-41.

Cann PA, Read NW, Holdsworth CD, and Barends D (1984) Role of loperamide and placebo in management of irritable bowel syndrome (IBS), *Digestive Diseases and Sciences,* 29(3):239-47.

Carling L, Svedberg L-E, and Hulten S (1989) Short term treatment of the irritable bowel syndrome: A placebo-controlled trial of peppermint oil against hyoscyamine, *Opuscula Medica,* 34(3):55-7.

Casellas F, Lopez VJ, and Malagelada JR (2005) Perceived health status in celiac disease, *Revista Espanola De Enfermedades Digestivas,* 97(11):794-804.

Cash BD, Schoenfeld P, and Chey WD (2002) The utility of diagnostic tests in irritable bowel syndrome patients: a systematic review, *American Journal of Gastroenterology,* 97(11):2812-9.

Cash BD and Chey WD (2002) Irritable bowel syndrome: an evidence-based management approach: case study and commentary, *Journal of Clinical Outcomes Management,* 9(7):409-18.

Cash BD and Chey WD (2004) Irritable bowel syndrome: A systematic review, *Clinics in Family Practice,* 6(3):647-69.

Cash BD and Chey WD (2005) Diagnosis of irritable bowel syndrome, *Gastroenterology Clinics of North America,* 34(2):205-20.

Catassi C, Kryszak D, Louis-Jacques O, Duerksen DR, Hill I, Crowe SE, Brown AR, Procaccini NJ, Wonderly BA, Hartley P, Moreci J, Bennett N, Horvath K, Burk M, and Fasano A (2007) Detection of Celiac disease in primary care: a multicenter case-finding study in North America, *American Journal of Gastroenterology,* 102(7):1454-60.

Chapman ND, Grillage MG, Mazumder R, and Atkinson SN (1990) A comparison of mebeverine with high-fibre dietary advice and mebeverine plus ispaghula in the treatment of irritable bowel syndrome: an open, prospectively randomised, parallel group study, *British Journal of Clinical Practice,* 44(11):461-6.

Charapata C and Mertz H (2006) Physician knowledge of Rome symptom criteria for irritable bowel syndrome is poor among non-gastroenterologists, *Neurogastroenterology and Motility,* 18(3):211-6.

Chaussade S and Minic M (2003) Comparison of efficacy and safety of two doses of two different polyethylene glycol-based laxatives in the treatment of constipation. *Alimentary Pharmacology and Therapeutics,* 17(1):165-72.

Chey WD, Olden K, Carter E, Boyle J, Drossman D, and Chang L (2002) Utility of the Rome I and Rome II criteria for irritable bowel syndrome in U.S. women, *American Journal of Gastroenterology,* 97(11):2803-11.

Chinn S (2000) A simple method for converting an odds ratio to effect size for use in meta-analysis, *Statistics in Medicine,* 19(22):3127-31.

Christie AH, Culbert P, Guest JF, Christie AH, Culbert P, and Guest JF (2002) Economic impact of low dose polyethylene glycol 3350 plus electrolytes compared with lactulose in the management of idiopathic constipation in the UK, *Pharmacoeconomics,* 20(1):49-60.

Colwell LJ, Prather CM, Phillips SF, and Zinsmeister AR (1998) Effects of an irritable bowel syndrome educational class on health- promoting behaviors and symptoms, *American Journal of Gastroenterology,* 93(6):901-5.

Conboy LA, Wasserman RH, Jacobson EE, Davis RB, Legedza AT, Park M, Rivers AL, Morey EB, Nam BH, Lasagna L, Kirsch I, Lembo AJ, Kaptchuk TJ, and Kerr CE (2006) Investigating placebo effects in irritable bowel syndrome: a novel research design, *Contemporary Clinical Trials,* 27(2):123-34.

Cook IJ, Irvine EJ, Campbell D, Shannon S, Reddy SN, and Collins SM (1990) Effect of dietary fiber on symptoms and rectosigmoid motility in patients with irritable bowel syndrome. A controlled, crossover study, *Gastroenterology,* 98(1):66-72.

Corazziari E, Badiali D, Habib FI, Reboa G, Pitto G, Mazzacca G, Sabbatini F, Galeazzi R, Cilluffo T, Vantini I, Bardelli E, and Baldi F (1996) Small volume isosmotic polyethylene glycol electrolyte balanced solution (PMF-100) in treatment of chronic nonorganic constipation, *Digestive Diseases and Sciences,* 41(8):1636-42.

Corazziari E, Badiali D, Bazzocchi G, Bassotti G, Roselli P, Mastropaolo G, Luca MG, Galeazzi R, and Peruzzi E (2000) Long term efficacy, safety, and tolerabilitity of low daily doses of isosmotic polyethylene glycol electrolyte balanced solution (PMF-100) in the treatment of functional chronic constipation, *Gut,* 46(4):522-6.

Corbett CL, Palmer KR, and Holdsworth CD (1980) Effect of loperamide, codeine phosphate and diphenoxylate on urgency and incontinence in chronic diarrhoea, *Gut,* 21(Suppl 1-12):A924.

Cornett JWD, Aspeling RL, and Mallegol D (1977) A double blind comparative evaluation of loperamide versus diphenoxylate with atropine in acute diarrhea, *Current Therapeutic Research, Clinical and Experimental,* 21(5):629-37.

Corney RH, Stanton R, Newell R, Clare A, and Fairclough P (1991) Behavioural psychotherapy in the treatment of irritable bowel syndrome, *Journal of Psychosomatic Research,* 35(4-5):461-9.

Corrao G, Corazza GR, Bagnardi V, Brusco G, Ciacci C, Cottone M, Sategna GC, Usai P, Cesari P, Pelli MA, Loperfido S, Volta U, Calabro A, and Certo M (2001) Mortality in patients with coeliac disease and their relatives: a cohort study, *Lancet,* 358(9279):356-61.

Corsetti M and Tack J (2004) Are symptom-based diagnostic criteria for irritable bowel syndrome useful in clinical practice?, *Digestion,* 70(4):207-9.

Creed F, Fernandes L, Guthrie E, Palmer S, Ratcliffe J, Read N, Rigby C, Thompson D, Tomenson B, and North of England IBS Research Group. (2003) The cost-effectiveness of psychotherapy and paroxetine for severe irritable bowel syndrome. *Gastroenterology,* 124(2):303-17.

Cremonini F and Talley NJ (2004) Diagnostic and therapeutic strategies in the irritable bowel syndrome, *Minerva Medica,* 95(5):427-41.

Curtin F, Morabia A, Bernstein M, and Dederding J-P (1998) A population survey of bowel habits in urban Swiss men, *European Journal of Public Health,* 8(2):170-5.

Czalbert HJ, Neder M, and Feher K (1990) Experiences with colpermintherapy (Tillots-England) at patients of irritable colon syndrome, *Gyogyszereszet,* 34(5):251-3.

Dancey CP and Backhouse S (1993) Towards a better understanding of patients with irritable bowel syndrome, *Journal of Advanced Nursing,* 18(9):1443-50.

Dancey CP, Hutton-Young A, Moye S, and Devins GM (2002) Perceived stigma, illness intrusiveness and quality of life in men and women with irritable bowel syndrome, *Psychology, Health and Medicine,* 7(4):381-95.

Dapoigny M, Bellanger J, Bonaz B, Des Varannes SB, Bueno L, Coffin B, Ducrotte P, Flourie B, Lemann M, Lepicard A, and Reigneau O (2004) Irritable bowel syndrome in France: A common, debilitating and costly disorder, *European Journal of Gastroenterology and Hepatology,* 16(10):995-1001.

Davis K (2006) Randomised double-blind placebo-controlled trial of aloe vera for irritable bowel syndrome, *International Journal of Clinical Practice,* 60(9):1080-6.

De Giorgio R, Barbara G, Stanghellini V, Cremon C, Salvioli B, De Ponti F, and Corinaldesi R (2004) Diagnosis and therapy of irritable bowel syndrome, *Alimentary Pharmacology and Therapeutics,* 20 Suppl 2:10-22.

De Schryver AM, Keulemans YC, Peters HP, Akkermans LM, Smout AJ, De Vries WR, and Berge-Henegouwen GP (2005) Effects of regular physical activity on defecation pattern in middle-aged patients complaining of chronic constipation, *Scandinavian Journal of Gastroenterology,* 40(4):422-9.

Dear KL, Elia M, and Hunter JO (2005) Do interventions which reduce colonic bacterial fermentation improve symptoms of irritable bowel syndrome?, *Digestive Diseases and Sciences,* 50(4):758-66.

Delbecq AL, Van de Ven AH, and Gustafson DH (1986) *Group Techniques For Program Planning: A Guide to Nominal Group and Delphi Processes*, Middleton, Wis: Green Briar Press.

Department of Health (2006) *Reference Costs 2005-06*, London: Department of Health.

Dettmar PW and Sykes J (1998) A multi-centre, general practice comparison of ispaghula husk with lactulose and other laxatives in the treatment of simple constipation, *Current Medical Research and Opinion,* 14(4):227-33.

Dettmer A (1994) Loperamide oxide in the treatment of acute diarrhea in adults, *Clinical Therapeutics,* 16(6):972-80.

Dogan UB and Unal S (1996) Kruis scoring system and Manning's criteria in diagnosis of irritable bowel syndrome: is it better to use combined?, *Acta Gastroenterologica Belgica,* 59(4):225-8.

Dom J, Leyman R, Schuermans V, and Brugmans J (1974) Loperamide (R 18 553), a novel type of antidiarrheal agent. Part 8: Clinical investigation. Use of a flexible dosage schedule in a double-blind comparison of loperamide with diphenoxylate in 614 patients suffering from acute diarrhea, *Arzneimittel-Forschung,* 24(10):1660-5.

Dretzke J, Cummins C, Sandercock J, Fry-Smith A, Barrett T, and Burls A (2004) Autoantibody testing in children with newly diagnosed type 1 diabetes mellitus, *Health Technology Assessment (Winchester, England),* 8(22):iii-ixi.

Dreverman JW and Van der Poel AJ (1995) Loperamide oxide in acute diarrhoea: a double-blind, placebo-controlled trial. The Dutch Diarrhoea Trialists Group, *Alimentary Pharmacology and Therapeutics,* 9(4):441-6.

Drossman DA, Thompson WG, and Whitehead WE (1992) Approaching IBS with confidence, *Patient Care,* 26(13):175-8.

Drossman DA and Thompson WG (1992) The irritable bowel syndrome: review and a graduated multicomponent treatment approach, *Annals of Internal Medicine,* 116(12 Pt 1):1009-16.

Drossman DA, Toner BB, Whitehead WE, Diamant NE, Dalton CB, Duncan S, Emmott S, Proffitt V, Akman D, Frusciante K, Le T, Meyer K, Bradshaw B, Mikula K, Morris CB, Blackman CJ, Hu Y, Jia H, Li JZ, Koch GG, and Bangdiwala SI (2003) Cognitive-behavioral therapy versus education and desipramine versus placebo for moderate to severe functional bowel disorders, *Gastroenterology,* 125(1):19-31.

Drossman DA, Morris CB, Hu Y, Toner BB, Diamant N, Leserman J, Shetzline M, Dalton C, Bangdiwala SI, Drossman DA, Morris CB, Hu Y, Toner BB, Diamant N, Leserman J, Shetzline M, Dalton C, and Bangdiwala SI (2005) A prospective assessment of bowel habit in irritable bowel syndrome in women: defining an alternator, *Gastroenterology,* 128(3):580-9.

Drossmsn DA (1984) Diagnosis of the irritable bowel syndrome. A simple solution, *Gastroenterology,* 87(1):224-5.

Drummond MF and Jefferson TO (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party, *British Medical Journal,* 313(7052):275-83.

Dubinsky MC, Johanson JF, Seidman EG, Ofman JJ, Dubinsky MC, Johanson JF, Seidman EG, and Ofman JJ (2002) Suspected inflammatory bowel disease--the clinical and economic impact of competing diagnostic strategies, *American Journal of Gastroenterology,* 97(9):2333-42.

Efskind PS, Bernklev T, and Vatn MH (1996) A double-blind placebo-controlled trial with loperamide in irritable bowel syndrome, *Scandinavian Journal of Gastroenterology,* 31(5):463-8.

Ehrenpreis ED (2004) Characterizing patients with IBS: how to make a confident diagnosis, *Family Practice Recertification,* 26(8):53-61.

El Serag HB, Olden K, Bjorkman D, El Serag HB, Olden K, and Bjorkman D (2002) Health-related quality of life among persons with irritable bowel syndrome: a systematic review, *Alimentary Pharmacology and Therapeutics,* 16(6):1171-85.

Ericsson CD, DuPont HL, Mathewson JJ, West MS, Johnson PC, and Bitsura JA (1990) Treatment of traveler's diarrhea with sulfamethoxazole and trimethoprim and loperamide, *Journal of the American Medical Association,* 263(2):257-61.

Fass R, Longstreth GF, Pimentel M, Fullerton S, Russak SM, Chiou CF, Reyes E, Crane P, Eisen G, McCarberg B, and Ofman J (2001) Evidence- and consensus-based practice guidelines for the diagnosis of irritable bowel syndrome, *Archives of Internal Medicine,* 161(17):2081-8.

Ferguson A, Culbert P, Gillett H, and Barras N (1999) New polyethylene glycol electrolyte solution for the treatment of constipation and faecal impaction, *Italian Journal of Gastroenterology and Hepatology,* 31 Suppl 3:S249-S252.

Fernandez C, Perez M, Amigo I, and Linares A (1998) Stress and contingency management in the treatment of irritable bowel syndrome, *Stress Medicine,* 14(1):31-42.

Fielding JF and Kehoe M (1984) Different dietary fibre formulations and the irritable bowel syndrome, *Irish Journal of Medical Science,* 153(5):178-80.

Forbes A, MacAuley S, and Chiotakakou-Faliakou E (2000) Hypnotherapy and therapeutic audiotape: effective in previously unsuccessfully treated irritable bowel syndrome? *International Journal of Colorectal Disease,* 15(5-6):328-34.

Forbes A, Jackson S, Walter C, Quraishi S, Jacyna M, and Pitcher M (2005) Acupuncture for irritable bowel syndrome: a blinded placebo-controlled trial, *World Journal of Gastroenterology,* 11(26):4040-4.

Fowell AJ, Thomas PW, Surgenor SL, and Snook JA (2006) The epidemiology of coeliac disease in East Dorset 1993-2002: an assessment of the 'coeliac iceberg', and preliminary evidence of case clustering, *QJM,* 99(7):453-60.

Fowlie S, Eastwood MA, and Prescott R (1992) Irritable bowel syndrome: assessment of psychological disturbance and its influence on the response to fibre supplementation, *Journal of Psychosomatic Research,* 36(2):175-80.

Francis CY, Duffy JN, Whorwell PJ, and Martin DF (1996) Does routine abdominal ultrasound enhance diagnostic accuracy in irritable bowel syndrome?, *American Journal of Gastroenterology,* 91(7):1348-50.

Francis CY and Whorwell PJ (1997) The irritable bowel syndrome, *Postgraduate Medical Journal,* 73(855):1-7.

Frigerio G, Beretta A, Orsenigo G, Tadeo G, Imperiali G, and Minoli G (1992) Irritable bowel syndrome. Still far from a positive diagnosis, *Digestive Diseases and Sciences,* 37(2):164-7.

Gade J and Thorn P (1989) Paraghurt for patients with irritable bowel syndrome. A controlled clinical investigation from general practice, *Scandinavian Journal of Primary Health Care,* 7(1):23-6.

Gaynes BN and Drossman DA (1999) The role of psychosocial factors in irritable bowel syndrome, *Best Practice and Research in Clinical Gastroenterology,* 13(3):437-52.

Gilbody JS, Fletcher CP, Hughes IW, and Kidman SP (2000) Comparison of two different formulations of mebeverine hydrochloride in irritable bowel syndrome, *International Journal of Clinical Practice,* 54(7):461-4.

Golechha AC, Chadda VS, Chadda S, Sharma SK, and Mishra SN (1982) Role of ispaghula husk in the management of irritable bowel syndrome (a randomized double-blind crossover study), *Journal of the Association of Physicians of India,* 30(6):353-5.

Gong Y-Y, Wen Y-H, and Qu T (2002) A study on the improvement of treatment effect to treat irritable bowel syndrome with psychological therapy, *Zhongguo Linchuang Kangfu,* 6(11):1706.

Griffith CJ (1994) Irritable bowel syndrome: pathophysiology, diagnosis, and management... recertification series, *Physician Assistant,* 18(11):25-6.

Grillage MG, Nankani JN, Atkinson SN, and Prescott P (1990) A randomised, double-blind study of mebeverine versus dicyclomine in the treatment of functional abdominal pain in young adults, *British Journal of Clinical Practice,* 44(5):176-9.

Gunn MC, Cavin AA, and Mansfield JC (2003) Management of irritable bowel syndrome, *Postgraduate Medical Journal,* 79(929):154-8.

Gwee K-A (2005) Irritable bowel syndrome in developing countries - A disorder of civilization or colonization? *Neurogastroenterology and Motility,* 17(3):317-24.

Hamilton-Miller JMT (2001) Probiotics in the management of irritable bowel syndrome: A review of clinical trials, *Microbial Ecology in Health & Disease,* 13(4):212-6.

Hamm LR, Sorrells SC, Harding JP, Northcutt AR, Heath AT, Kapke GF, Hunt CM, and Mangel AW (1999) Additional investigations fail to alter the diagnosis of irritable bowel syndrome in subjects fulfilling the Rome criteria, *American Journal of Gastroenterology,* 94(5):1279-82.

Hammer J and Talley NJ (1999) Diagnostic criteria for the irritable bowel syndrome, *American Journal of Medicine,* 107(5A):5S-11S.

Harford WV, Krejs GJ, Santa Ana CA, and Fordtran JS (1980) Acute effect of diphenoxylate with atropine (Lomotil) in patients with chronic diarrhea and fecal incontinence, *Gastroenterology,* 78(3):440-3.

Harvey RF, Mauad EC, and Brown AM (1987) Prognosis in the irritable bowel syndrome: a 5-year prospective study, *Lancet,* 1(8539):963-5.

Harvey RF, Hinton RA, Gunary RM, and Barry RE (1989) Individual and group hypnotherapy in treatment of refractory irritable bowel syndrome, *Lancet,* 1(8635):424-5.

Hatlebakk JG and Hatlebakk MV (2004) Diagnostic approach to suspected irritable bowel syndrome., *Best Practice and Research in Clinical Gastroenterology,* 18(4):735-46.

Haynes RB and Wilczynski NL (2004) Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey, *British Medical Journal,* 328(7447):1040.

Hebden JM, Blackshaw E, D'Amato M, Perkins AC, and Spiller RC (2002) Abnormalities of GI transit in bloated irritable bowel syndrome: effect of bran on transit and symptoms, *American Journal of Gastroenterology,* 97(9):2315-20.

Heymann-Monnikes I, Arnold R, Florin I, Herda C, Melfsen S, and Monnikes H (2000) The combination of medical treatment plus multicomponent behavioral therapy is superior to medical treatment alone in the therapy of irritable bowel syndrome, *American Journal of Gastroenterology,* 95(4):981-94.

Hin H, Bird G, Fisher P, Mahy N, and Jewell D (1999) Coeliac disease in primary care: case finding study, *British Medical Journal,* 318(7177):164-7.

Holten KB (2003) Irritable bowel syndrome: minimize testing, let symptoms guide treatment, *Journal of Family Practice,* 52(12):942-50.

House of Lords Select Committee on Science and Technology (2000) *Sixth Report (HL123): Complementary and Alternative Medicine*, London:

Hovdenak N (1987) Loperamide treatment of the irritable bowel syndrome, *Scandinavian Journal of Gastroenterology - Supplement,* 130:81-4.

Inauen W and Halter F (1994) Clinical efficacy, safety and tolerance of mebeverine slow release (200 mg) vs mebeverine tablets in patients with irritable bowel syndrome, *Drug Investigation,* 8(4):234-40.

Jaffe G (1977) A comparison of lomotil and imodium in acute non-specific diarrhoea, *Journal of International Medical Research,* 5(3):195-8.

Jalihal A and Kurian G (1990) Ispaghula therapy in irritable bowel syndrome: improvement in overall well-being is related to reduction in bowel dissatisfaction, *Journal of Gastroenterology and Hepatology,* 5(5):507-13.

Jeong H, Lee HR, Yoo BC, and Park SM (1993) Manning criteria in irritable bowel syndrome: its diagnostic significance, *Korean Journal of Internal Medicine,* 8(1):34-9.

Johnson KH and Millard PS (1996) Diagnosis of irritable bowel syndrome, *Journal of Family Practice,* 42(4):346-7.

Joint Formulary Committee (2007) *British National Formulary*, 53rd Edition. London: British Medical Association and Royal Pharmaceutical Society of Great Britain.

Jones J and Hunter D (1995) Consensus methods for medical and health services research, *British Medical Journal,* 311(7001):376-80.

Jones J, Boorman J, Cann P, Forbes A, Gomborone J, Heaton K, Hungin P, Kumar D, Libby G, Spiller R, Read N, Silk D, and Whorwell P (2000) British Society of Gastroenterology guidelines for the management of the irritable bowel syndrome, *Gut,* 47 Suppl 2:ii1-19.

Kajander K, Hatakka K, Poussa T, Farkkila M, and Korpela R (2005) A probiotic mixture alleviates symptoms in irritable bowel syndrome patients: a controlled 6-month intervention, *Alimentary Pharmacology and Therapeutics,* 22(5):387-94.

Keefer L and Blanchard EB (2001) The effects of relaxation response meditation on the symptoms of irritable bowel syndrome: results of a controlled treatment study, *Behaviour Research and Therapy,* 39(7):801-11.

Kennedy A, Robinson A, and Rogers A (2003) Incorporating patients' views and experiences of life with IBS in the development of an evidence based self-help guidebook, *Patient Education and Counseling,* 50(3):303-10.

Kennedy T, Jones R, Darnley S, Seed P, Wessely S, and Chalder T (2005) Cognitive behaviour therapy in addition to antispasmodic treatment for irritable bowel syndrome in primary care: randomised controlled trial, *British Medical Journal,* 331(7514):435.

Kennedy T, Chalder T, McCrone P, Darnley S, Knapp M, Jones R, and Wessely S (2006) Cognitive behavioural therapy in addition to antispasmodic therapy for irritable bowel syndrome in primary care: a randomised controlled trial, *Health Technology Assessment (Winchester, England),* 10(19)

Kienzle-Horn S, Vix J-M, Schuijt C, Peil H, Jordan CC, and Kamm MA (2006) Efficacy and safety of bisacodyl in the acute treatment of constipation: A double-blind, randomized, placebo-controlled study, *Alimentary Pharmacology and Therapeutics,* 23(10):1479-88.

Kienzle-Horn S, Vix JM, Schuijt C, Peil H, Jordan CC, and Kamm MA (2007) Comparison of bisacodyl and sodium picosulphate in the treatment of chronic constipation, *Current Medical Research and Opinion,* 23(4):691-9.

Kim HJ, Camilleri M, McKinzie S, Lempke MB, Burton DD, Thomforde GM, and Zinsmeister AR (2003) A randomized controlled trial of a probiotic, VSL#3, on gut transit and symptoms in diarrhoea-predominant irritable bowel syndrome, *Alimentary Pharmacology and Therapeutics,* 17(7):895-904.

Kim HJ, Vazquez Roque MI, Camilleri M, Stephens D, Burton DD, Baxter K, Thomforde G, and Zinsmeister AR (2005) A randomized controlled trial of a probiotic combination VSL# 3 and placebo in irritable bowel syndrome with bloating, *Neurogastroenterology and Motility,* 17(5):687-96.

Kim YJ and Ban DJ (2005) Prevalence of irritable bowel syndrome, influence of lifestyle factors and bowel habits in Korean college students, *International Journal of Nursing Studies,* 42(3):247-54.

Kruis W, Thieme C, Weinzierl M, Schussler P, Holl J, and Paulus W (1984) A diagnostic score for the irritable bowel syndrome. Its value in the exclusion of organic disease, *Gastroenterology,* 87(1):1-7.

Kruis W, Weinzierl M, Schussler P, and Holl J (1986) Comparison of the therapeutic effect of wheat bran, mebeverine and placebo in patients with the irritable bowel syndrome, *Digestion,* 34(3):196-201.

Kumar A, Kumar N, Vij JC, Sarin SK, and Anand BS (1987) Optimum dosage of ispaghula husk in patients with irritable bowel syndrome: correlation of symptom relief with whole gut transit time and stool weight, *Gut,* 28(2):150-5.

Kyosola K (1985) Positive diagnostic criteria for the irritable bowel syndrome, *Gastroenterology,* 89(1):235-6.

Lacy BE, Weiser K, Noddin L, Robertson DJ, Crowell MD, Parratt-Engstrom C, and Grau MV (2007) Irritable bowel syndrome: patients' attitudes, concerns and level of knowledge, *Alimentary Pharmacology and Therapeutics,* 25(11):1329-41.

Lambert JP, Brunt PW, Mowat NAG, Khin CC, Lai CKW, Morrison V, Dickerson JWT, and Eastwood MA (1991) The value of prescribed 'high-fibre' diets for the treatment of the irritable bowel syndrome, *European Journal of Clinical Nutrition,* 45(12):601-9.

Lavo B, Stenstam M, and Nielsen AL (1987) Loperamide in treatment of irritable bowel syndrome--a double-blind placebo controlled study, *Scandinavian Journal of Gastroenterology - Supplement,* 130:77-80.

Lea R, Hopkins V, Hastleton J, Houghton LA, and Whorwell PJ (2004) Diagnostic criteria for irritable bowel syndrome: utility and applicability in clinical practice, *Digestion,* 70(4):210-3.

Leahy ACB, Clayman C, Mason I, and Epstein O (1997) Dosed relaxation in resistant irritable bowel syndrome using gut directed biofeedback, *Gut,* 40 Suppl 1:A51.

Lech Y, Olesen KM, Hey H, Rask-Pedersen E, Vilien M, and Ostergaard O (1988) Treatment of irritable bowel syndrome with peppermint oil. A double-blind investigation with a placebo, *Ugeskrift for Laeger,* 150(40):2388-9.

Lee IK (1968) Comparative trial of Lomotil-with-neomycin and kaolin-and-morphine mixture in general practice, *British Journal of Clinical Practice,* 22(8):347-8.

Legorreta AP, Ricci J-F, Markowitz M, and Jhingran P (2002) Patients diagnosed with irritable bowel syndrome: Medical record validation of a claims-based identification algorithm, *Disease Management and Health Outcomes,* 10(11):715-22.

Lembo A (2004) Irritable bowel syndrome medications side effects survey, *Journal of Clinical Gastroenterology,* 38(9):776-81.

Leung WK, Wu JC, Liang SM, Chan LS, Chan FK, Xie H, Fung SS, Hui AJ, Wong VW, Che CT, and Sung JJ (2006) Treatment of diarrhea-predominant irritable bowel syndrome with traditional Chinese herbal medicine: a randomized placebo-controlled trial, *American Journal of Gastroenterology,* 101(7):1574-80.

Liu GZ (1997) [Treatment of acupuncture and moxibustion plus psychotherapy for irritable bowel syndrome], *Chinese Acupuncture and Moxibustion,* 10:611-2.

Liu JH, Chen GH, Yeh HZ, Huang CK, and Poon SK (1997) Enteric-coated peppermint-oil capsules in the treatment of irritable bowel syndrome: a prospective, randomized trial, *Journal of Gastroenterology,* 32(6):765-8.

Longstreth GF, Fox DD, Youkeles L, Forsythe AB, and Wolochow DA (1981) Psyllium therapy in the irritable bowel syndrome. A double-blind trial, *Annals of Internal Medicine,* 95(1):53-6.

Longstreth GF (1994) Irritable bowel syndrome and chronic pelvic pain, *Obstetrical and Gynecological Survey,* 49(7):505-7.

Longstreth GF and Drossman DA (2002) New developments in the diagnosis and treatment of irritable bowel syndrome, *Current Gastroenterology Reports,* 4(5):427-34.

Lucey MR, Clark ML, Lowndes J, and Dawson AM (1987) Is bran efficacious in irritable bowel syndrome? A double blind placebo controlled crossover study, *Gut,* 28(2):221-5.

Lustman F, Walters EG, Shroff NE, and Akbar FA (1987) Diphenoxylate hydrochloride (Lomotil(Reg.trademark)) in the treatment of acute diarrhoea, *British Journal of Clinical Practice,* 41(3):648-51.

Lustyk MKB, Jarrett ME, Bennett JC, and Heitkemper MM (2001) Does a physically active lifestyle improve symptoms in women with irritable bowel syndrome?, *Gastroenterology Nursing,* 24(3):129-37.

Lynch PM and Zamble E (1989) A controlled behavioral treatment study of Irritable Bowel Syndrome, *Behavior Therapy,* 20(4):509-23.

MacIntosh DG, Thompson WG, Patel DG, Barr R, and Guindi M (1992) Is rectal biopsy necessary in irritable bowel syndrome?, *American Journal of Gastroenterology,* 87(10):1407-9.

Madden JAJ and Hunter JO (2002) A review of the role of the gut microflora in irritable bowel syndrome and the effects of probiotics, *British Journal of Nutrition,* 88(SUPPL. 1):S67-S72.

Madisch A, Holtmann G, Plein K, and Hotz J (2004) Treatment of irritable bowel syndrome with herbal preparations: results of a double-blind, randomized, placebo-controlled, multi-centre trial, *Alimentary Pharmacology and Therapeutics,* 19(3):271-9.

Manning AP, Heaton KW, and Harvey RF (1977) Wheat fibre and irritable bowel syndrome. A controlled trial, *Lancet,* 2(8035):417-8.

Manning AP, Thompson WG, Heaton KW, and Morris AF (1978) Towards positive diagnosis of the irritable bowel, *British Medical Journal,* 2(6138):653-4.

Mascolo R and Saltzman JR (1998) Lactose intolerance and irritable bowel syndrome, *Nutrition Reviews,* 56(10):306-8.

Mason J, Nicolson D, and Wilson D (2002) Systematic Review Methods for National Guidelines (Unpublished Discussion Paper).

McIntyre A, Vincent RM, Perkins AC, and Spiller RC (1997) Effect of bran, ispaghula, and inert plastic particles on gastric emptying and small bowel transit in humans: The role of physical factors, *Gut,* 40(2):223-7.

Mearin F, Badia X, Balboa A, Baro E, Caldwell E, Cucala M, Diaz-Rubio M, Fueyo A, Ponce J, Roset M, and Talley NJ (2001) Irritable bowel syndrome prevalence varies enormously depending on the employed diagnostic criteria: comparison of Rome II versus previous criteria in a general population, *Scandinavian Journal of Gastroenterology,* 36(11):1155-61.

Mearin F, Baro E, Roset M, Badia X, Zarate N, Perez I, Mearin F, Baro E, Roset M, Badia X, Zarate N, and Perez I (2004) Clinical patterns over time in irritable bowel syndrome: symptom instability and severity variability, *American Journal of Gastroenterology,* 99(1):113-21.

Mearin F, Roset M, Badia X, Balboa A, Baro E, Ponce J, Diaz-Rubio M, Caldwell E, Cucala M, Fueyo A, and Talley NJ (2004) Splitting irritable bowel syndrome: from original Rome to Rome II criteria, *American Journal of Gastroenterology,* 99(1):122-30.

Medoff J, Katz S, Malik P, Pambianco D, Pruitt R, Poulos J, Rank J, and Rose M (2004) Open-label, dose-ranging pilot study of 4 weeks of low-dose therapy with sodium phosphate tablets in chronically constipated adults, *Clinical Therapeutics,* 26(9):1479-91.

Mein SM, Ladabaum U, Mein SM, and Ladabaum U (2004) Serological testing for coeliac disease in patients with symptoms of irritable bowel syndrome: a cost-effectiveness analysis, *Alimentary Pharmacology and Therapeutics,* 19(11):1199-210.

Miller V, Lea R, Agrawal A, and Whorwell PJ (2006) Bran and irritable bowel syndrome: The primary-care perspective, *Digestive and Liver Disease,* 38(10):737-40.

Mitchell SA, Mee AS, Smith GD, Palmer KR, and Chapman RW (2002) Alverine citrate fails to relieve the symptoms of irritable bowel syndrome: Results of a double-blind, randomized, placebo-controlled trial, *Alimentary Pharmacology and Therapeutics,* 16(6):1187-95.

Moreno-Osset E, Calvo B, Tomas-Ridocci M, Montero I, Minguez M, Mora F, and Benages A (1991) The irritable bowel syndrome: diagnostic strategies, *Italian Journal of Gastroenterology,* 23(8 Suppl 1):41-7.

Morgan T and Robson KM (2002) Irritable bowel syndrome: Diagnosis is based on clinical criteria, *Postgraduate Medicine,* 112(5):30-41.

Mortensen PB, Andersen JR, Arffmann S, and Krag E (1987) Short-chain fatty acids and the irritable bowel syndrome: the effect of wheat bran, *Scandinavian Journal of Gastroenterology,* 22(2):185-92. Myren J, Lovland B, Larssen SE, and Larsen S (1984) Psychopharmacologic drugs in the treatment of the irritable bowel syndrome. A double blind study of the effect of trimipramine, *Annales De Gastroenterologie Et D'Hepatologie.,*(3):117-23.

Myren J, Lovland B, Larssen SE, and Larsen S (1984) A double-blind study of the effect of trimipramine in patients with the irritable bowel syndrome, *Scandinavian Journal of Gastroenterology.,*(6):835-43.

National Collaborating Centre for Mental Health (2004) *National Clinical Practice Guideline Number 23: Depression: Management of depression in primary and secondary care,* Leicester: British Psychological Society.

National Institute for Health and Clinical Excellence (2006) *Public Health Intervention Guidance no. 2: Four commonly used methods to increase physical activity: brief interventions in primary care, exercise referral schemes, pedometers and community-based exercise programmes for walking and cycling,* London: National Institute for Health and Clinical Excellence.

National Institute for Health and Clinical Excellence (2007) *The guidelines manual,* London: National Institute for Health and Clinical Excellence.

Neff DF and Blachard EB (1987) A multi-component treatment for irritable bowel syndrome, *Behavior Therapy,* 18(1):70-83.

Netten A and Curtis L (2006) *Unit Costs of Health and Social Care 2006,* University of Canterbury, Kent: Personal Social Services Research Unit.

Newcastle University (2007) *http://www.ncl.ac.uk/nnp/teaching/management/dynamic/index.html [Accessed 6th August 2007],* Newcastle: Newcastle University.

NHS Health and Social Care Information Centre HC (2006) *Prescription Cost Analysis, England 2005*, London: NHS Health and Social Care Information Centre, Health CareStatistics.

Niec AM, Frankum B, and Talley NJ (1998) Are adverse food reactions linked to irritable bowel syndrome?, *American Journal of Gastroenterology,* 93(11):2184-90.

Niedzielin K, Kordecki H, and Birkenfeld B (2001) A controlled, double-blind, randomized study on the efficacy of Lactobacillus plantarum 299V in patients with irritable bowel syndrome, *European Journal of Gastroenterology and Hepatology,* 13(10):1143-7.

Nigam P, Kapoor KK, Rastog CK, Kumar A, and Gupta AK (1984) Different therapeutic regimens in irritable bowel syndrome, *Journal of the Association of Physicians of India,* 32(12):1041-4.

Niv E, Naftali T, Hallak R, and Vaisman N (2005) The efficacy of Lactobacillus reuteri ATCC 55730 in the treatment of patients with irritable bowel syndrome--a double blind, placebo-controlled, randomized study, *Clinical Nutrition,* 24(6):925-31.

Nobaek S, Johansson ML, Molin G, Ahrne S, and Jeppsson B (2000) Alteration of intestinal microflora is associated with reduction in abdominal bloating and pain in patients with irritable bowel syndrome, *American Journal of Gastroenterology,* 95(5):1231-8.

O'Leary C, Wieneke P, Buckley S, O'Regan P, Cronin CC, Quigley EM, and Shanahan F (2002) Celiac disease and irritable bowel-type symptoms, *American Journal of Gastroenterology,* 97(6):1463-7.

O'Mahony L, McCarthy J, Kelly P, Hurley G, Luo F, Chen K, O'Sullivan GC, Kiely B, Collins JK, Shanahan F, and Quigley EM (2005) Lactobacillus and bifidobacterium in irritable bowel syndrome: symptom responses and relationship to cytokine profiles, *Gastroenterology,* 128(3):541-51.

O'Sullivan MA, Mahmud N, Kelleher DP, Lovett E, and O'Morain CA (2000) Patient knowledge and educational needs in irritable bowel syndrome, *European Journal of Gastroenterology and Hepatology,* 12(1):39-43.

Odes HS and Madar Z (1991) A double-blind trial of a celandin, aloevera and psyllium laxative preparation in adult patients with constipation, *Digestion,* 49(2):65-71.

OECD (2006) *OECD Health Data 2006*, Paris: OECD.

Office of National Statistics (2006) *Interim Life Tables, Expectation of Life, England and Wales*, London: Office of National Statistics.

Olden KW (2003) The challenge of diagnosing irritable bowel syndrome, *Reviews in Gastroenterological Disorders,* 3 Suppl 3:S3-11.

Olesen M and Gudmand-Hoyer E (2000) Efficacy, safety, and tolerability of fructooligosaccharides in the treatment of irritable bowel syndrome, *American Journal of Clinical Nutrition,* 72(6):1570-5.

Ouyang H and Chen JD (2004) Review article: therapeutic roles of acupuncture in functional gastrointestinal disorders, *Alimentary Pharmacology and Therapeutics,* 20(8):831-41.

Oxman AD, Cook DJ, and Guyatt GH (1994) Users' guides to the medical literature: VI. How to use an overview, *Journal of the American Medical Association,* 272(17):1367-71.

Page JG and Dirnberger GM (1981) Treatment of the irritable bowel syndrome with Bentyl (dicyclomine hydrochloride), *Journal of Clinical Gastroenterology,* 3(2):153-6.

Paille F, Colombey N, Alleaume B, and Vicari F (1999) An open six-month study of the safety of Transipeg for treating constipation in community medicine, Journal of Clinical Research, 2:65-76.

Palmer KR, Corbett CL, and Holdsworth CD (1980) Double-blind cross-over study comparing loperamide, codeine and diphenoxylate in the treatment of chronic diarrhea, *Gastroenterology,* 79(6):1272-5.

Palsson OS, Turner MJ, Johnson DA, Burnett CK, and Whitehead WE (2002) Hypnosis treatment for severe irritable bowel syndrome: Investigation of mechanism and effects on symptoms, *Digestive Diseases and Sciences,* 47(11):2605-14.

Parisi G, Bottona E, Carrara M, Cardin F, Faedo A, Goldin D, Marino M, Pantalena M, Tafner G, Verdianelli G, Zilli M, and Leandro G (2005) Treatment effects of partially hydrolyzed guar gum on symptoms and quality of life of patients with irritable bowel syndrome. A multicenter randomized open trial, *Digestive Diseases and Sciences,* 50(6):1107-12.

Parisi GC, Zilli M, Miani MP, Carrara M, Bottona E, Verdianelli G, Battaglia G, Desideri S, Faedo A, Marzolino C, Tonon A, Ermani M, and Leandro G (2002) High-fiber diet supplementation in patients with irritable bowel syndrome (IBS): a multicenter, randomized, open trial comparison between wheat bran diet and partially hydrolyzed guar gum (PHGG), *Digestive Diseases and Sciences,* 47(8):1697-704.

Paterson WG, Thompson WG, Vanner SJ, Faloon TR, Rosser WW, Birtwhistle RW, Morse JL, and Touzel TA (1999) Recommendations for the management of irritable bowel syndrome in family

practice. IBS Consensus Conference Participants, *Canadian Medical Association Journal,* 161(2):154-60.

Payne A and Blanchard EB (1995) A controlled comparison of cognitive therapy and self-help support groups in the treatment of irritable bowel syndrome, *Journal of Consulting and Clinical Psychology,* 63(5):779-86.

Pelemans W and Vantrappen F (1976) A double blind crossover comparison of loperamide with diphenoxylate in the symptomatic treatment of chronic diarrhea, *Gastroenterology,* 70(6):1030-4.

Pimentel M, Chow EJ, and Lin HC (2000) Eradication of small intestinal bacterial overgrowth reduces symptoms of irritable bowel syndrome, *American Journal of Gastroenterology,* 95(12):3503-6.

Pimentel M, Chow EJ, and Lin HC (2003) Normalization of lactulose breath testing correlates with symptom improvement in irritable bowel syndrome. a double-blind, randomized, placebo-controlled study, *American Journal of Gastroenterology,* 98(2):412-9.

Poynard T, Couturier D, Frexinos J, Bommelaer G, Hernandez M, Dapoigny M, Buscail L, Benand-Agostini H, Chaput JC, and Rheims N (1992) French experience of Manning's criteria in the irritable bowel syndrome, *European Journal of Gastroenterology and Hepatology,* 4(9):747-52.

Prior A and Whorwell PJ (1987) Double blind study of ispaghula in irritable bowel syndrome, *Gut,* 28(11):1510-3.

Quah HM, Ooi BS, Seow-Choen F, Sng KK, and Ho KS (2006) Prospective randomized crossover trial comparing fibre with lactulose in the treatment of idiopathic chronic constipation, *Techniques in Coloproctology,* 10(2):111-4.

Quartero AO, Meineche-Schmidt V, Muris J, Rubin G, and de Wit N (2007) Bulking agents, antispasmodic and antidepressant medication for the treatment of irritable bowel syndrome., *Cochrane Database of Systematic Reviews,* 2005, Issue 2. Art. No.: CD003460. DOI: 10.1002/14651858.CD003460.pub2.

Quigley EMM (2007) Bacterial flora in irritable bowel syndrome: Role in pathophysiology, implications for management, *Chinese Journal of Digestive Diseases,* 8(1):2-7.

Radnitz CL and Blanchard EB (1989) A 1- and 2- year follow-up study of bowel sound biofeedback as a treatment for irritable bowel syndrome, *Biofeedback & Self Regulation.,* 14(4):333-8.

Rajagopalan M, Kurian G, and John J (1998) Symptom relief with amitriptyline in the irritable bowel syndrome, *Journal of Gastroenterology and Hepatology,* 13(7):738-41.

Rao KP, Gupta S, Jain AK, Agrawal AK, and Gupta JP (363) Evaluation of Manning's criteria in the diagnosis of irritable bowel syndrome, *Journal of the Association of Physicians of India,* 41(6):357-8. Rao SSC, Welcher K, Zimmerman B, and Stumbo P (1998) Is coffee a colonic stimulant?, *European Journal of Gastroenterology and Hepatology,* 10(2):113-8.

Rees GA, Trevan M, and Davies GJ (1994) Dietary fibre modification and the symptoms of irritable bowel syndrome - A review, *Journal of Human Nutrition & Dietetics,* 7(3):179-89.

Reid G (2001) *Regulatory and clinical aspects of dairy probiotics. Joint FAO/WHO Expert Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food Including Powder Milk with Live Lactic Acid Bacteria. Cordoba (Argentina), 1-4 Oct 2001*, Rome: Italy.

Ringel Y and Drossman DA (2001) Psychosocial Factors in Functional Gastrointestinal Disorders: Toward a More Comprehensive Understanding and Approach to Treatment, *Medscape,* http://www.medscape.com/viewarticle/418547 [Accessed 9th August 2007]

Ritchie JA and Truelove SC (1979) Treatment of irritable bowel syndrome with lorazepam, hyoscine butylbromide, and ispaghula husk, *British Medical Journal,* 1(6160):376-8.

Roberts L, Wilson S, Singh S, Roalfe A, and Greenfield S (2006) Gut-directed hypnotherapy for irritable bowel syndrome: piloting a primary care-based randomised controlled trial, *British Journal of General Practice,* 56(523):115-21.

Robinson A, Lee V, Kennedy A, Middleton L, Rogers A, Thompson DG, and Reeves D (2006) A randomised controlled trial of self-help interventions in patients with a primary care diagnosis of irritable bowel syndrome, *Gut,* 55(5):643-8.

Rogers A, Lee V, and Kennedy A (2007) Continuity and change?: Exploring reactions to a guided self-management intervention in a randomised controlled trial for IBS with reference to prior experience of managing a long term condition, *Trials [Electronic Resource],* 8, 2007. Article Number

Rothstein RD (2000) Irritable bowel syndrome, *Medical Clinics of North America,* 84(5):1247-57.

Rouse M, Chapman N, Mahapatra M, Grillage M, Atkinson SN, and Prescott P (1991) An open, randomised, parallel group study of lactulose versus ispaghula in the treatment of chronic constipation in adults, *British Journal of Clinical Practice,* 45(1):28-30.

Saggioro A (2004) Probiotics in the treatment of irritable bowel syndrome, *Journal of Clinical Gastroenterology,* 38(6 Suppl):S104-S106.

Saito YA, Locke GR, Talley NJ, Zinsmeister AR, Fett SL, and Melton LJ, III (2000) A comparison of the Rome and Manning criteria for case identification in epidemiological investigations of irritable bowel syndrome, *American Journal of Gastroenterology,* 95(10):2816-24.

Saito YA, Talley NJ, Melton J, Fett S, Zinsmeister AR, and Locke GR (2003) The effect of new diagnostic criteria for irritable bowel syndrome on community prevalence estimates, *Neurogastroenterology and Motility,* 15(6):687-94.

Sanders DS, Carter MJ, Hurlstone DP, Pearce A, Ward AM, McAlindon ME, and Lobo AJ (2001) Association of adult coeliac disease with irritable bowel syndrome: a case-control study in patients fulfilling ROME II criteria referred to secondary care, *Lancet,* 358(9292):1504-8.

Sanders DS, Patel D, Stephenson TJ, Ward AM, McCloskey EV, Hadjivassiliou M, and Lobo AJ (2003) A primary care cross-sectional study of undiagnosed adult coeliac disease, *European Journal of Gastroenterology and Hepatology,* 15(4):407-13.

Sanjoaquin MA, Appleby PN, Spencer EA, and Key TJ (2004) Nutrition and lifestyle in relation to bowel movement frequency: A cross-sectional study of 20 630 men and women in EPIC-Oxford, *Public Health Nutrition,* 7(1):77-83.

Schafer E and Ewe K (1990) [The treatment of irritable colon. Efficacy and tolerance of buscopan plus, buscopan, paracetamol and placebo in ambulatory patients with irritable colon], *Fortschritte Der Medizin,* 108(25):488-92.

Schaffstein W, Panijel M, and Luttecke K (1990) Comparative safety and efficacy of trimebutine versus mebeverine in the treatment of irritable bowel syndrome. A multicenter double-blind study, *Current Therapeutic Research, Clinical and Experimental,* 47(1):136-45.

Schneider A, Enck P, Streitberger K, Weiland C, Bagheri S, Witte S, Friederich HC, Herzog W, and Zipfel S (2006) Acupuncture treatment in irritable bowel syndrome, *Gut,* 55(5):649-54.

Schuster MM (1991) Diagnostic evaluation of the irritable bowel syndrome, *Gastroenterology Clinics of North America,* 20(2):269-78.

Schuster MM (2001) Defining and diagnosing irritable bowel syndrome., *American Journal of Managed Care,* 7(8 Suppl):S246-S251.

Scottish Nutrition and Diet Resources Initiative (2004) *Irritable Bowel Syndrome (IBS) Your Diet Can Help [Available from: http://www.caledonian.ac.uk/sndri/pdf/IrritableBowelSyndrome.pdf]*, Glasgow: Caledonian University.

Shrivastava RK and Siegel H (1984) The role of tricyclics and benzodiazepine compounds in the treatment of irritable gut syndrome and peptic ulcer disease, *Psychopharmacology Bulletin,* 20(4):616-21.

Skoog SM and Bharucha AE (2004) Dietary fructose and gastrointestinal symptoms: A review, *American Journal of Gastroenterology,* 99(10):2046-50.

Smith RC (1992) Diagnosing the irritable bowel syndrome, *Annals of Internal Medicine,* 117(12):1056-7.

Snook J and Shepherd HA (1994) Brain supplementation in the treatment of irritable bowel syndrome, *Alimentary Pharmacology and Therapeutics,* 8(5):511-4.

Soltoft J, Krag B, Gudmand-Hoyer E, Kristensen E, and Wulff HR (1976) A double-blind trial of the effect of wheat bran on symptoms of irritable bowel syndrome, *Lancet,* 1(7954):270-2.

Spiegel BM, DeRosa VP, Gralnek IM, Wang V, Dulai GS, Spiegel BMR, DeRosa VP, Gralnek IM, Wang V, and Dulai GS (2004) Testing for celiac sprue in irritable bowel syndrome with predominant diarrhea: a cost-effectiveness analysis, *Gastroenterology,* 126(7):1721-32.

Steinhart MJ, Wong PY, and Zarr ML (1981) Therapeutic usefulness of amitriptyline in spastic colon syndrome, *International Journal of Psychiatry in Medicine,* 11(1):45-57.

Steinmann B, Gitzelmann R, and Van den Berghe G (2007) 'Disorders of Fructose Metabolism', in: Scriver CR, (editor) *The Metabolic Basis of Inherited Disease. Vol 1.* New York; London: McGraw-Hill, 1511.

Suleiman S, Sonnenberg A, Suleiman S, and Sonnenberg A (2001) Cost-effectiveness of endoscopy in irritable bowel syndrome, *Archives of Internal Medicine,* 161(3):369-75.

Symons P, Jones MP, and Kellow JE (1992) Symptom provocation in irritable bowel syndrome. Effects of differing doses of fructose-sorbitol, *Scandinavian Journal of Gastroenterology,* 27(11):940-4.

Tabas G, Beaves M, Wang J, Friday P, Mardini H, and Arnold G (2004) Paroxetine to treat irritable bowel syndrome not responding to high-fiber diet: a double-blind, placebo-controlled trial., *American Journal of Gastroenterology,* 99(5):914-20.

Talley NJ, Phillips SF, Melton LJ, Mulvihill C, Wiltgen C, and Zinsmeister AR (1990) Diagnostic value of the Manning criteria in irritable bowel syndrome, *Gut,* 31(1):77-81.

Talley NJ (2003) When to conduct testing in patients with suspected irritable bowel syndrome, *Reviews in Gastroenterological Disorders,* 3 Suppl 3:S18-S24.

Taneja I, Deepak KK, Poojary G, Acharya IN, Pandey RM, and Sharma MP (2004) Yogic versus conventional treatment in diarrhea-predominant irritable bowel syndrome: a randomized control study, *Applied Psychophysiology and Biofeedback,* 29(1):19-33.

Tanum L and Malt UF (1996) A new pharmacologic treatment of functional gastrointestinal disorder. A double-blind placebo-controlled study with mianserin, *Scandinavian Journal of Gastroenterology,* 31(4):318-25.

Tarpila S, Tarpila A, Grohn P, Silvennoinen T, and Lindberg L (2004) Efficacy of ground flaxseed on constipation in patients with irritable bowel syndrome, *Current Topics in Nutraceuticals Research,* 2(2):119-25.

Thomas B (2001) 'Disorders of the colon: Irritable Bowel Synrome', in: Thomas B, (editor) *Manual of Dietetic Practice.* Oxford: Blackwell Scientific Publications, 398-400.

Thomas BJ (1992) British Diabetic Association's discussion paper on the role of 'diabetic' foods, *Diabetic Medicine,* 9(3):300-6.

Thompson WG, Dotevall G, Drossman DA, Heaton KW, and Kruis W (1989) Irritable bowel syndrome: Guidelines for the diagnosis, *Gastroenterology International,* 2(2):92-5.

Thompson WG (1997) Gender differences in irritable bowel symptoms, *European Journal of Gastroenterology and Hepatology,* 9(3):299-302.

Tijtgat GN, Meuwissen SG, and Huibregtse K (1975) Loperamide in the symptomatic control of chronic diarrhoea. Double-blind placebo-controlled study, *Annals of Clinical Research,* 7(5):325-30.

Tkachuk GA, Graff LA, Martin GL, and Bernstein CN (2003) Randomized controlled trial of cognitive-behavioral group therapy for irritable bowel syndrome in a medical setting, *Journal of Clinical Psychology in Medical Settings,* 10(1):57-69.

Tolliver BA, Herrera JL, and DiPalma JA (1994) Evaluation of patients who meet clinical criteria for irritable bowel syndrome, *American Journal of Gastroenterology,* 89(2):176-8.

Tovey P (2002) A single-blind trial of reflexology for irritable bowel syndrome, *British Journal of General Practice,* 52(474):19-23.

Tripathi BM, Misra NP, and Gupta AK (1983) Evaluation of tricyclic compound (trimipramine) vis-a-vis placebo in irritable bowel syndrome. (Double blind randomised study), *Journal of the Association of Physicians of India,* 31(4):201-3.

Tsuchiya J, Barreto R, Okura R, Kawakita S, Fesce E, and Marotta F (2004) Single-blind follow-up study on the effectiveness of a symbiotic preparation in irritable bowel syndrome, *Chinese Journal of Digestive Diseases,* 5(4):169-74.

Tynes LL and Gibson RL (1999) Irritable bowel syndrome: overview of diagnosis and treatment, *Journal of the Louisiana State Medical Society,* 151(2):76-81.

UK Council for Psychotherapy (UKCP) (1992) *Hypno-Psychotherapy Section [http://www.psychotherapy.org.uk/hypno_psychotherapy.html]*, London: UK Council for Psychotherapy.

Van Outryve M, Mayeur S, Meeus MA, Rosillon D, Hendrickx B, and Ceuppens M (1995) A double-blind crossover comparison study of the safety and efficacy of mebeverine with mebeverine sustained release in the treatment of irritable bowel syndrome, *Journal of Clinical Pharmacy and Therapeutics,* 20(5):277-82.

Vanner SJ, Depew WT, Paterson WG, DaCosta LR, Groll AG, Simon JB, and Djurfeldt M (1999) Predictive value of the Rome criteria for diagnosing the irritable bowel syndrome., *American Journal of Gastroenterology,* 94(10):2912-7.

Verhaegen H, De Cree J, and Schuermans V (1974) Loperamide (R 18 553), a novel type of antidiarrheal agent. Part 7: Clinical investigation. Efficacy and safety of loperamide in patients with severe chronic diarrhea, *Arzneimittel-Forschung,* 24(10):1657-60.

Vesa TH, Seppo LM, Martea PR, Sahi T, and Korpela R (1998) Role of irritable bowel syndrome in subjective lactose intolerance, *American Journal of Clinical Nutrition,* 67(4):710-05.

Vij JG, JR, and KN (1991) Effect of antidepressant drug (doxepin) on irritable bowel syndrome patients, *Indian Journal of Psychiatry,* 33:243-6.

Villagrasa M, Boix J, Humbert P, and Quer JC (1991) Aleatory clinical study comparing otilonium bromide with a fiber-rich diet in the treatment of irritable bowel syndrome, *Italian Journal of Gastroenterology,* 23(8 Suppl 1):67-70.

Vollmer A and Blanchard EB (1998) Controlled comparison of individual versus group cognitive therapy for irritable bowel syndrome, *Behavior Therapy,* 29(1):19-33.

Wang G, Li TQ, Wang L, Xia Q, Chang J, Zhang Y, Wan MH, Guo J, Cheng Y, Huang X, and Zhang RM (2006) Tong-xie-ning, a Chinese herbal formula, in treatment of diarrhea-predominant irritable bowel syndrome: a prospective, randomized, double-blind, placebo-controlled trial, *Chinese Medical Journal,* 119(24):2114-9.

Watson AR and Bowling TE (2005) Irritable bowel syndrome: diagnosis and symptom management, *British Journal of Community Nursing,* 10(3):118-22.

West J, Logan RF, Smith CJ, Hubbard RB, and Card TR (2004) Malignancy and mortality in people with coeliac disease: population based cohort study, *British Medical Journal,* 329(7468):716-9.

Whorwell PJ, Prior A, and Faragher EB (1984) Controlled trial of hypnotherapy in the treatment of severe refractory irritable-bowel syndrome, *Lancet,* 2(8414):1232-4.

Whorwell PJ, Altringer L, Morel J, Bond Y, Charbonneau D, O'Mahony L, Kiely B, Shanahan F, and Quigley EM (2006) Efficacy of an encapsulated probiotic Bifidobacterium infantis 35624 in women with irritable bowel syndrome, *American Journal of Gastroenterology,* 101(7):1581-90.

Wiklund IK, Fullerton S, Hawkey CJ, Jones RH, Longstreth GF, Mayer EA, Peacock RA, Wilson IK, and Naesdal J (2003) An irritable bowel syndrome-specific symptom questionnaire: development and validation, *Scandinavian Journal of Gastroenterology,* 38(9):947-54.

Wilhelmsen I (2000) The role of psychosocial factors in gastrointestinal disorders, *Gut,* 47(Suppl 4):73-5.

Wolf S (1981) The psyche and the stomach. A historical vignette, *Gastroenterology,* 80:605-14.

Wulkow R, Vix JM, Schuijt C, Peil H, Kamm MA, and Jordan C (2007) Randomised, placebo-controlled, double-blind study to investigate the efficacy and safety of the acute use of sodium picosulphate in patients with chronic constipation, *International Journal of Clinical Practice,* 61(6):944-50.

Yadav SK, Jain AK, Tripathi SN, and Gupta JP (1989) Irritable bowel syndrome: Therapeutic evaluation of indigenous drugs, *Indian Journal of Medical Research,* 90:496-503.

Zastrow C and Navarre R (1977) Nominal Group - New Tool for Making Social-Work Education Relevant, *Journal of Education for Social Work,* 13(1):112-8.