Provocation-Neutralisation Testing and Therapy for Food Allergy

A West Midlands Health Technology Assessment Collaboration Report

Authors: Janine Dretzke
Fujian Song

Correspondence to: Department of Public Health and Epidemiology
University of Birmingham
Edgbaston
Birmingham
B15 2TT

Report number: 44

ISBN No: 07044 24746

West Midlands Health Technology Assessment Collaboration (WMHTAC)

The WMHTAC produces rapid systematic reviews about the effectiveness of health care interventions and technologies, in response to requests from West Midlands Health Authorities. Reviews take approximately 6 months and aim to give a timely and accurate analysis of the available evidence, with an economic analysis (usually a cost-utility analysis) of the intervention accompanied by a statement of the quality of the evidence.

Contribution of authors

Janine Dretzke designed the protocol, undertook the searches, performed the quality assessment and data extraction and wrote the report.

Fujian Song advised on the protocol and the general writing of the report and checked all the extracted data.

Conflicts of interest

None.

Acknowledgements:

Sue Bayliss for giving advice on the search strategy.

Dr J R Mansfield, The Burghwood Clinic, Surrey; Dr M J Radcliffe, The Burghwood Clinic, Surrey, Dr HM Anthony, President of the British Society for Allergy, Nutritional and Environmental Medicine and Dr JA Monro, The Breakspear Hospital, Hertfordshire for commenting on the draft protocol and providing information.

West Midlands Regional Evaluation Panel Recommendation:

The recommendation for the use of provocation-neutralisation testing and neutralisation therapy for food allergy is:

No evidence identified to suggest provocation-neutralisation testing is useful for diagnosis of food allergy.

Insufficient evidence identified to recommend the use of neutralisation therapy for food allergy.

Further research is recommended.

Anticipated expiry date:

The searches on clinical effectiveness were completed in September and October 2002. The authors were not aware of any ongoing trials at the time. The report will require updating should new trial evidence become available.

Table of contents

Tables

Figures

Appendices

Summary

Objective: The objective of the report was to systematically review the available evidence regarding the clinical effectiveness and/or diagnostic test accuracy of provocation-neutralisation testing or treatment for food allergy.

Data sources: MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials and the Centralised Information Service for Complementary Medicine Database were searched, and citation searches performed. Ongoing or unpublished studies were sought through contact with experts and website searches.

Inclusion criteria: Parallel or crossover randomised controlled trials or diagnostic test accuracy studies; provocation-neutralisation with food allergens compared to placebo or other tests/treatment; any technique of provocation-neutralisation as defined by the authors (most common methodology is the Miller technique); patients with food allergy; outcomes as defined by authors.

Quality assessment, data extraction and synthesis: Data on study and patient characteristics, test and treatment protocols, outcomes, direction of effect and quality was extracted by one reviewer and checked by a second reviewer. Key quality criteria for randomised controlled trials and diagnostic test accuracy studies were assessed using checklists. Numerical pooling of data was not possible given the clinical heterogeneity.

Quantity of data: Six randomised controlled trials and two controlled trials (n=462) of provocation testing compared to placebo were identified, two of which (n=107) also investigated diagnostic test accuracy. Five trials (n=87) of neutralisation therapy compared to placebo were identified. Patients were clinically heterogeneous and displayed a variety of symptoms. Test and treatment protocols also varied between studies. The main outcomes assessed were the provocation of, or relief from, symptoms.

Quality of data & direction of effect: Five studies (n=365) showed no difference in symptom provocation between active extract and placebo. Two studies (n=60) found that more symptoms were provoked and one (n=132) found that more skin wheals were provoked. One study (n=37) was uninterpretable, as results for placebo were not stated. Three studies (n=61) showed a benefit from neutralisation therapy, one (n=11) showed no difference and one (n=15) showed either a benefit or no difference depending on outcome assessment. Most studies had methodological flaws, which made the interpretation of results difficult. These related, amongst others, to verification of allergy status, outcome assessment and presentation of results, randomisation and ability to identify the placebo. Two studies calculated sensitivities and specificities for small patient numbers, but design of the studies was poor and did not allow conclusions to be drawn.

Conclusion: There was no convincing evidence to suggest that more symptoms or wheals can be provoked with active extract compared to placebo by provocation-neutralisation. No evidence was identified to suggest that provocation-neutralisation is useful for the diagnosis of food allergy. There was some evidence, based on small patient numbers, to

suggest that neutralisation may be effective in the treatment of food allergies, although uncertainty remains around the outcome assessment and the initial diagnosis of food allergy in some of these studies. It should be noted that the absence of good evidence is not proof of ineffectiveness, and further well-designed studies are recommended for the assessment of the treatment aspect of this technique.

Abbreviations

BSACI	British Society for Allergy & Clinical Immunology
DBPCFC	double-blind placebo-controlled food challenge
CISCOM	Centralised Information Service for Complementary Medicine
MeSH	medical subject heading
IgE	Immunoglobulin E
ITT	intention-to-treat analysis
ND	neutralising dose
OFC	oral food challenge
RCT	randomised controlled trial

1 Aim of review

This report aims to systematically review the available evidence regarding the effectiveness of:

- provocation-neutralisation testing compared to placebo and/or the reference standard and/or other tests for diagnosing food allergies or sensitivities
- neutralisation therapy compared to placebo and/or conventional treatment and/or no treatment for the treatment of food allergy

2 Rationale for review

Provocation-neutralisation testing and neutralisation therapy (also known as the Miller technique) has been employed for diagnosis and treatment of food allergy or sensitivity since at least 1961.[1] It is a controversial technique, with studies of heterogeneous design reporting conflicting results regarding effectiveness.

The technique is not used within the National Health Service, nor is it supported by the British Society for Allergy and Clinical Immunology (BSACI), as stated in a 1992 report by the Royal College of Physicians Committee on Clinical Immunology[2] and Allergy and in a 1993 Position Paper by a BSACI working party.[3] Allergy UK (formerly the British Allergy Foundation) states that provocation-neutralisation testing is not regarded by conventional medical practitioners to be relevant and are considered to have no place in the diagnosis of true allergy.[4]

Similarly, the American Academy of Allergy and Immunology[5] and the National Centre for Health Care Technology[6] reported in 1981 that the procedure of both intra-and subcutaneous and sublingual provocation and neutralisation is unproven. In 1989 the American College of Physicians published a position paper[7], which reviewed evidence of effectiveness for provocation neutralisation testing. They concluded that all studies were seriously flawed in terms of study design and that results were therefore conflicting.

In contrast, there are many practitioners both in the UK and particularly in the USA who currently use the technique. A number of private clinics offer the treatment in the UK, for example the Burghwood Clinic[8], The Breakspear Hospital[9] and the Airedale Allergy Centre.[10] The technique is endorsed by the British Society for Allergy, Environmental and Nutritional Medicine and also by the American Academy of Environmental Medicine, the American Academy of Otolaryngic Allergy and the Australasian College of Nutritional and Environmental Medicine (personal communication Dr Anthony).

In support, a review by Gerdes (1989)[11], which looked at English language studies on provocation/neutralisation between 1969 and 1988, came to the conclusion that, although more research was needed to understand the nature of the technique, the evidence in favour of the technique outweighed the evidence against it.

Enquiries by West Midlands Health Authorities to the Aggressive Research Intelligence Facility (ARIF)[12] based at the University of Birmingham regarding the effectiveness of this technique reveal are ongoing. Given the overall rise in allergies in recent years[13], this interest is likely to increase.

3 Background

3.1 Food allergy and sensitivity

Allergy is a form of hypersensitivity to a substance (allergen), which can be inhaled, swallowed, injected or comes into contact with the skin or eye. Examples of allergens include pollens, house dust, fungal spores and foods. A number of terms are used to describe adverse reactions to food depending on the underlying mechanisms, including food allergy, food (hyper-) sensitivity, food intolerance and food aversion, and there is some confusion about terminology. Food sensitivity and intolerance are often used as general descriptive terms to include all types of adverse reactions to foods.[2] Food allergy is not a single disease but encompasses a range of disorders including acute potentially fatal reactions, chronic diseases affecting mainly the skin and gastrointestinal tract[14], also asthma and rhinitis.[15]

Typical food allergens of infancy and childhood are egg, milk, peanut, wheat and soya, whilst reactions in older children and adults are more often due to peanuts, tree nuts and seafood. Fruit and vegetable allergies are common, but are generally not severe.[14]
The terms listed below are used in accordance with a BMJ Clinical Review article[16] and a report of the Royal College of Physicians Committee on Clinical Immunology and Allergy.[2]

IgE mediated food allergy

In IgE mediated food allergy, atopic individuals produce increased amounts of IgE, which binds to mast cells. When the antigen (allergen) binds to IgE, it triggers a series of biochemical events resulting in the release of histamine and other substances from the mast cell, which can cause the symptoms of the allergy.[2] The timing of IgE mediated food reactions is closely associated with food intake and symptoms can arise within minutes, typically involving more than one organ (e.g. itching and swelling in the mouth, vomiting, abdominal pain, asthma, rhinitis, urticaria). In classic IgE mediated food allergy, a specific food can often be identified as the cause. Reactions can vary in the severity and discomfort they cause. Life threatening reactions can include exacerbation of asthma, laryngeal oedema and anaphylaxis with cardiovascular collapse. Foods, which may provoke severe reactions, are nuts, seeds, eggs, milk and shellfish.[16]

Non-IgE mediated food allergy

Non-IgE mediated food allergy is characterised by delayed reactions, which may take hours or days to develop. Foods provoking delayed reactions can include milk, eggs, fish, wheat, other cereals, yeast, soya, pork, chocolate, fruit and vegetables. IgE is not involved in these reactions, however there is strong evidence that the immune system plays a role. Coeliac disease, cow's milk enteropathy and dermatitis herpetiformis are examples of non IgE mediated food allergy. [14,16]

Non-allergic food intolerance

Non-allergic food intolerance can be due to pharmacological, metabolic or toxic causes.[16] Some foods have a histamine releasing action (for example egg whites, tomatoes, shellfish, strawberries and chocolate), which can result in symptoms similar to those resulting from an IgE mediated food allergy. Food intolerance can be linked to the absence of an enzyme, for example in the case of lactose intolerance caused by lactase deficiency. Other causes of adverse reactions include tyramine in wine and cheese (can provoke migraine) and monosodium glutamate in Chinese food (can provoke flushing, headache and abdominal symptoms).[2,16]

Psychological factors in food intolerance (food aversion)

Food may be avoided for psychological reasons, or there may be psychological intolerance where an adverse reaction is caused by emotions associated with the food rather than the food itself. Symptoms often do not manifest themselves when the suspected food is given in an unrecognisable form (e.g. as part of a double-blind food challenge protocol).[17]

Association of food allergy with specific diseases

There is uncertainty regarding the role of food allergy in certain (chronic) diseases such as chronic fatigue syndrome, migraine, rheumatoid arthritis, irritable bowel syndrome, psychological disturbances and others. The conventional viewpoint is that these diseases do not have an allergic basis[2,15], whilst practitioners of environmental medicine argue there is. Examples cited by environmental medicine practitioners are, amongst others, studies where symptoms of patients with conditions such as migraine[18] irritable bowel syndrome[19], rheumatoid arthritis[20] improved on elimination diets.

3.2 Burden of disease

The true burden of disease is difficult to estimate, not least due to the problems of diagnosis (see section 3.3.1) and the lack of consensus regarding the types of disease associated with food allergy (see section 3.1.). A further issue in determining the prevalence of food allergies is the lack of consensus around the terminology, the fact that many cases may be handled by the patient themselves by avoiding particular foods, and the fact that identifying offending foods is not easy as there may be a lack of objective signs or distinct symptomatology.[21]

Young *et al.* (1994)[22] conducted a survey of 7,500 households in the Wycombe Health Authority Area and 7,500 randomly selected households throughout the UK to identify the prevalence of reactions to eight foods (e.g. cow's milk, hen's egg, prawn, nuts) commonly perceived to cause sensitivity. It was estimated that 20.4% of the UK population perceive food intolerance as a problem, whilst the prevalence as confirmed by double-blind, placebo-controlled food challenge ranged between 1.4% and 1.8% depending on stringency of criteria. There is thus a discrepancy between perception of food intolerance and actual food intolerance. The British Nutrition Foundation[23]also quote an estimated prevalence of 1.4% for food allergy, with 20% of the adult population believing themselves to be allergic. A Dutch survey[24] found that 12.4% of the population believed they had an allergy or intolerance to specific foods, whilst a subsequent double-blind placebo-controlled food challenge revealed a minimum prevalence of 0.8%, with an estimated population prevalence of 2.4%. Similarly, polls carried out in the US[25] found that between 14% and 16% of responding households reported at least one individual with an allergy, which is not supported by studies using the double-blind placebo controlled food challenge. Food reactions are more common in children than adults and can often be transient.[22] The overall prevalence of IgE mediated food allergies in children has been estimated at 5-7%.[16]

There were several critical comments on the study by Young *et al*. (1994)[22], all suggesting that the prevalence is likely to be an underestimate. Anthony *et al*. (1994)[26] state that, amongst other issues, the study failed to take into account chronic symptoms apparently unrelated to food until 'unmasked' by a break in exposure (the controls chosen for the study - who were unaware of food intolerance - showed a higher rate than the estimate for the study as a whole); only a subset of patients undertook the double-blind challenge procedures, and fewer than 60% completed the study; patients were omitted due to the severity of their reported reactions but only a small subset of these were accepted as showing food intolerance. Finn (1994)[27] also states that the estimate is likely to be too low, because many patients do not recognise that they have a food intolerance. Moneret-Vautrin & Kanny (1994)[28] state that the foods used in the Young study represent only 60% of food allergies in children and 36.5% in adults. They also query the amount of fresh food used for the challenges.

Based on the most conservative estimate of an average of 1.6% by Young *et al*. (1994)[22], there are around 1 million food allergy sufferers in the UK (based on a population of roughly 60 million) bearing in mind that the real figure may be much higher.

It does however seem to be generally accepted that food allergies present a growing problem. A report from the Royal College of Physicians (June 2003)[13] states that there has been a dramatic increase in allergy over the last years (based on primary studies cited in the report on allergic rhinitis and asthma, nut allergy, anaphylaxis, occupational allergy and allergic reactions to drugs). This increase relates to the incidence but also to the severity (life-threatening anaphylaxis is becoming more common) and the complexity (disorders affecting several systems – 'multi-system allergic disease'). Food allergy has become increasingly widespread and is the most common cause of anaphylaxis in children.

3.3 Current service provision

3.3.1 Conventional diagnosis

Diagnosis of food allergies can present problems due to the large number of medical and psychosocial variables involved.[29] Food allergy has been described as being overdiagnosed by the public, often underdiagnosed by physicians and frequently misdiagnosed by allergists.[30] A correct diagnosis is important, as patients may unnecessarily restrict their diet.

The most commonly used diagnostic tests and some of the associated difficulties are briefly outlined in the following. The list is not exhaustive.

Medical history/physical examination
Assessment of food allergy begins with a history and physical examination to consider broad differential diagnoses.[14] Inquiry should cover factors such as age of onset, nature of symptoms, time of onset in relation to food intake and others. Medical history alone may be sufficient in some cases to diagnose or exclude food allergy. The findings in physical examination, particularly in a symptomatic patient, might support the diagnosis of food allergy or suggest a non-allergic disorder.[30]

Trials of elimination diets
If particular foods are suspected, then avoidance followed by reintroduction of that food can be attempted. If the patient remains asymptomatic during the avoidance period without taking medication, foods are slowly reintroduced to identify the offending ones. If severe symptoms are anticipated, the procedure should be done under medical supervision. The correct choice of food to eliminate and the degree of patient compliance can influence the success of the treatment.[30]

Double-blind placebo controlled food challenge (DBPCFC)
If true food sensitivity is present then symptoms should disappear when the food is eliminated from the diet, and should re-appear when the food is reintroduced, even if disguised.[2] The most commonly accepted diagnostic test is the double-blind placebo controlled elimination and challenge test, which is considered by many as the 'gold standard' as it is least prone to bias from the patients or investigators.[14,16,31] There are however practical difficulties in applying the tests, for example, regarding co-contaminants, synergistic food problems, food additive hypersensitivity and delayed reactions, and false negative results can be common.[31] There can be difficulties in finding an appropriate placebo or masking both the food and placebo, whilst also attempting to reproduce the natural exposure, particularly regarding the form and quantity of the food suspected of provoking the reaction (if the food is contained within a capsule, for example, the volume may not be large enough to provoke a reaction).[30,32] It has been suggested that negative challenges should be followed by supervised open feeding the following day in order to eliminate false-negative results.[33] Food challenges should be performed under strict medical supervision as anaphylactic reactions may occur and should be avoided in instances where a life-threatening reaction has previously

occurred.[30,34]

Skin prick testing

Skin prick testing can be performed to diagnose IgE mediated food allergy. A drop of liquid (undiluted) food or piece of solid food is placed on the forearm and pricked through the epidermis. If the antibody is present, a wheal and flare should occur within 15 minutes. Positive (histamine) and negative (saline-glycerine) controls are used to establish that the immune response is not blocked and to identify false positive responses due to local trauma.[14] A problem with this method is that standardised food (antigen) extracts are rarely available making the test less reliable compared to other allergens. [34]

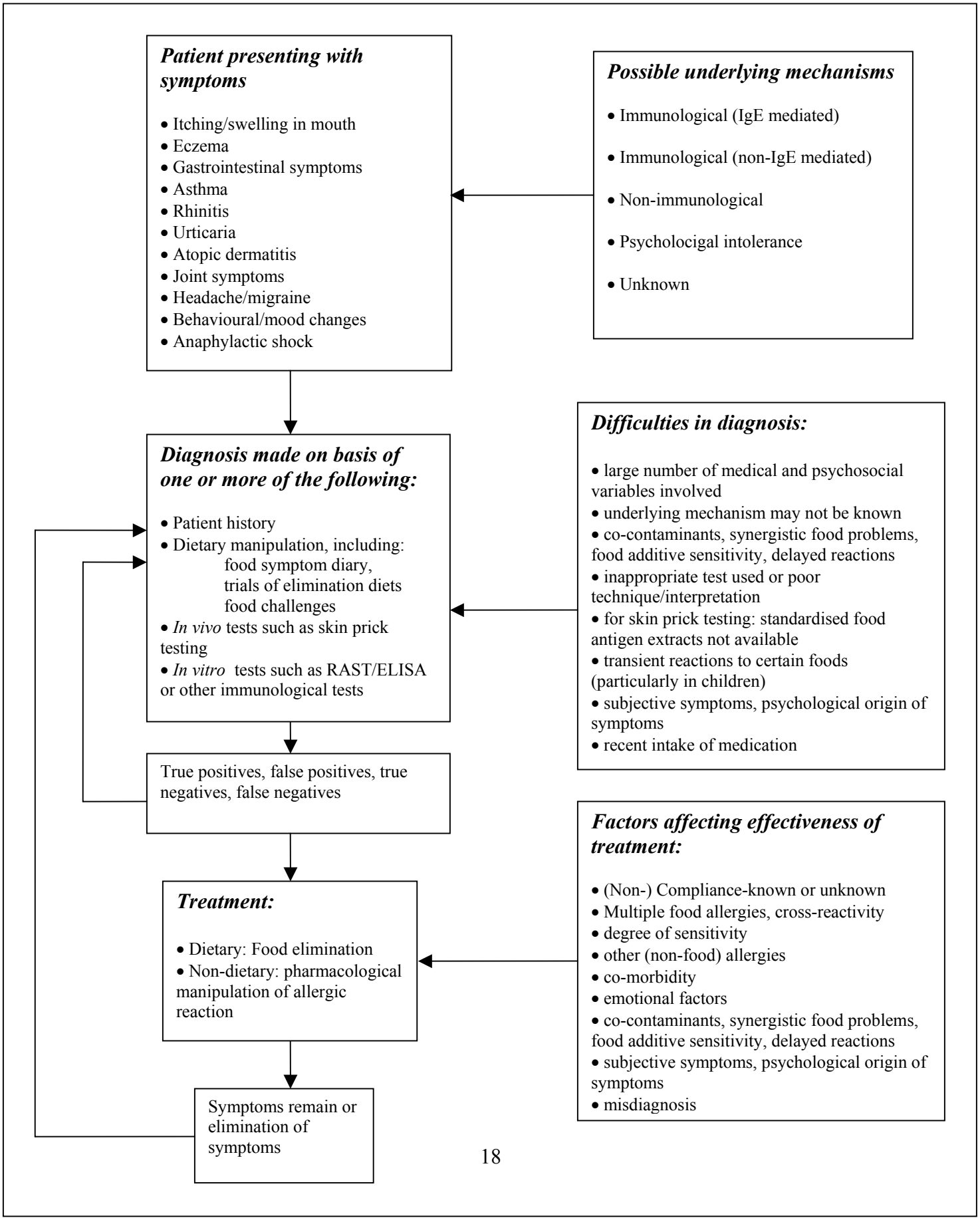
RAST

In cases where skin testing is inconclusive, serological tests may be carried out to identify allergen-specific IgE. The most widely used method is the radioallergosorbent test (RAST), a solid-phase immunoassay.[29] Allergens are bound to an insoluble matrix and patients' serum is added followed by radioactively labelled anti-IgE. The amount of bound radioactively labelled anti-IgE is then measured. Other immunological methods exist (ELISA, CAP-RAST etc.).[34]

Patient management

Management of individual patients will vary depending on their symptoms, the type of underlying mechanism of their allergy or sensitivity, the types of tests that are performed and their accuracy etc. 'Classic' IgE mediated food allergies may be easier to recognise due to the close time association between ingestion of the food and reaction.[16] Figure 1 below outlines a very broad patient pathway and some of the associated factors that can make diagnosis and treatment difficult. It refers to conventional diagnoses and treatment options.

Figure 1 Broad patient pathway and problems associated with diagnosis



3.3.2 Conventional treatment

Conventional treatment for food allergy is avoidance of the food in question, which involves training of the patient to avoid a specific food and advice on labelling of foods.[2,16,31] Patients (particularly children) may outgrow their reactivity to a food, so the diagnosis should be regularly re-evaluated.[16] Antiallergy drugs such as cromoglycate and glucocorticoids have been used in dietary treatment in patients with food allergy, but their effectiveness remains unproven.[2,16]

3.3.3 Provision of allergy services in the UK

The Royal College of Physicians Report (June 2003)[13] states that in view of the increasing incidence of allergies, there is a growing gulf between the need for effective advice and treatment and the availability of professional services. The report further states that there is an urgent need for specialist led allergy centres and for GPs to acquire the necessary knowledge and training in order to provide an effective primary care service. There are only six major centres across the country staffed by consultant allergists. A further nine centres offer a specialist service. The remaining allergy clinics (the majority) are run by part time consultants in other disciplines. Most allergy specialists are based in London and the South East. The provision of consultant allergists is approximately one per 2 million of the UK population compared to one per 100,000 for mainstream specialities such as gastroenterology, cardiology etc. It should be noted that these figures refer to all allergies, not just food allergy.

3.4 Provocation-neutralisation

3.4.1 Provocation-neutralisation testing and neutralisation therapy

Rinkel *et al* (1944)[35] introduced the concept of intracutaneous titration of aqueous inhalant allergens to determine an effective 'dose' for injection therapy of pollenosis. The modern form of the technique using food antigens dates back to the work carried out by Hugh Carlton Lee[36] and was subsequently modified by Joseph Miller.[37] The technique is now generally known as the Miller technique and is the technique most commonly used by practitioners of provocation-neutralisation. (personal communication Dr Mansfield)

In the Miller technique, 0.01 ml or 0.05 ml of consecutively stronger or weaker concentrations of an allergy extract in a 1:5 dilution series are injected intradermally (one at a time) every 7-10 minutes. Glycerine is used as the diluent. Appearance of a skin wheal within this time period is considered to be a positive response. A positive wheal will be blanched, hard and raised and typically grows 2mm or more in 10 minutes. Symptoms may or may not be additionally provoked or relieved, or there may be a delayed response. If a positive wheal response is obtained, successively weaker doses are given and the first dose that results in a negative wheal (no local reaction) is considered to be the neutralising dose (ND). The negative wheal is relatively soft, flat and may be neutral-coloured or erythematous, but not blanched, and typically grows less than 2mm. All concentrations weaker than this (underdose concentrations) will also produce negative wheals. The ND is usually also the concentration that relieves test symptoms should they have occurred. Overdoses (i.e. concentrations stronger than that of the neutralising dose) do not always produce symptoms, but should always produce a wheal.[38-40]

Once the ND has been determined, patients then go on to subcutaneously self-administer sets of 0.05 ml at this concentration in order to gain protection from the symptoms caused by ongoing exposures to a particular food. Individualised treatment doses for several foods can be combined into a single immunotherapeutic solution. Relief of symptoms can occur by the time testing is completed or within a few weeks. Errors in interpretation may occur when a ND is rejected or an underdose accepted as a ND by a physician due to symptoms persisting or appearing in response to a previous injection. Similarly overdoses may be accepted as a ND if symptoms clear very quickly.[38,39]

This focus on the appearance/disappearance of a wheal is in contrast to earlier methods for example the 1977 study by Miller[41], where it is stated that the test procedure consists of inducing mild symptoms and then giving successive intradermal injections of different dilutions until the dilution is found which completely relieves the symptoms. It is therefore possible that the results of studies performed in different years are interpreted differently depending on whether symptoms and/or wheals are being assessed.

Test substances are variable and include food and food additives, phenol, alcohol, glycerine, saline, histamine, tobacco, newsprint and inhalant allergens. A complete series of tests can take up to several days as only one dose of one item can be assessed at a time.[7]

There are many descriptions of provocation-neutralisation in the literature and not all

correspond to the Miller technique. Grieco (1982)[42] describes the technique according to Rinkel (1964)[43] and states that "increased concentrations of allergens are administered subcutaneously and intracutaneously to provoke symptoms within 10 minutes, corresponding to the patient's complaints; neutralising doses are initiated as soon as symptoms develop as weaker or stronger dilutions of the same antigen to relieve the provoked symptoms." Goldberg & Kaplan (1991)[44] state that: "food antigen extracts are injected intracutaneously or subcutaneously in increasing concentrations until significant whealing or reproduction of presenting symptoms is achieved; the ND is considered to be the dilution that produces a 7-8mm wheal and increases by 2mm in 10 minutes."

As can be seen in section 5, the trialists use different techniques for provocation testing and/or neutralisation therapy, and it is therefore important to note that not all of them are investigating the same technique. It is essential to establish, amongst others, the route used for administration, how the endpoint was defined (e.g. wheal or symptoms or both) and the dilutions used when comparing studies of provocation-neutralisation.

There is also uncertainty around whether provocation-neutralisation is or should be used predominantly for treatment or also as a diagnostic tool. Clearly, provocation-neutralisation testing is described in the literature as a diagnostic test [1,29], and trials have been performed to establish the diagnostic test accuracy [45,46]. Personal communication with a practitioner in the field (Dr Mansfield) revealed that in practice food allergy may be diagnosed by elimination diet trials and then neutralisation therapy used to enable the patient to consume those foods identified, as the technique is less reliable as a diagnostic test. Similarly a personal communication from Dr Anthony states that "experienced clinicians and their professional groups do not regard the neutralisation technique as a diagnostic method."

Provocation and neutralisation tests can also be carried out using sublingual techniques. Drops of diluted food antigen are administered sublingually and symptoms appearing within 20 minutes are neutralised by applications of weaker dilutions of the antigen. The neutralising dose can then be administered before ingestion of the food likely to cause a reaction.[44] It is not clear if there is one particular protocol that is more or less commonly used.

3.4.2 Use of provocation-neutralisation testing in the UK

It is difficult to estimate the exact number of patients being treated with this technique in the UK, as it performed only at some private clinics. A UK population based survey on the use of complementary and alternative medicines using a geographically stratified random sample of 5,010 adults found that this technique was not mentioned by any respondents under the option of 'other treatments' (although it should be noted that the questionnaire did not ask specifically about this treatment). (personal communication Kate Thomas, Medical Care Research Unit, University of Sheffield) A Scottish survey, which sampled 1,987 patients similarly found that this treatment was not listed by any respondent under 'other treatment', although again the questionnaire did not ask specifically about this treatment. (personal communication Cornelia Featherstone, Highlands and Islands Health Research Institute, Inverness)

Between 1996 and 2003 there were around ten enquiries made by West Midlands Health Authorities to the Aggressive Research Intelligence Facility (ARIF)[12] based at the University of Birmingham regarding the effectiveness of this technique (personal communication Dr Dave Moore and Sue Bayliss, ARIF). These enquiries relate to very small patient numbers. It is not possible to estimate the number of patients who seek treatment independently and it is not known how many GPs refer patients on to these clinics. Knowledge of the treatment, accessibility and cost are likely to be issues as the technique is not offered by the NHS. Dr Mansfield of the Burghwood Clinic, Surrey, states that he has tested around 12,000 patients using the intradermal technique since 1978 (around 500 per year). (personal communication)
It is not known to the authors of this report how many (if any) of these patients are NHS referrals.

4 Methodology

4.1 Search strategy

4.1.1 Existing reviews

The following databases were searched to identify any previous systematic reviews: Cochrane database of systematic reviews, the NHS Centre for Reviews and Dissemination databases (DARE, HTA), and MEDLINE and EMBASE using search terms relating to the intervention and to reviews. Several reviews were identified[7,11,47], however none followed a protocol for systematically identifying studies.

4.1.2 Primary studies

Searches were targeted to identify both those studies in which provocation-neutralisation was investigated or used as a diagnostic tool, and studies designed as trials to investigate the therapeutic benefit of neutralisation therapy:

In the literature, some studies assume diagnostic utility of provocation-neutralisation testing and go on to investigate if an association exists between apparent allergy to specific foods (based on provocation-neutralisation test) and symptoms (possible attributable to such allergy). These studies are performed as double-blind placebo-controlled RCTs. Each patient is injected with placebo or allergen extract in random order with the aim of provoking more symptoms on active extract compared to placebo in order to attempt a diagnosis of food allergy. Unfortunately, the subjective nature of symptoms experienced on testing and the unjustified assumption that a provocation-neutralisation test result is diagnostically valuable means such studies are flawed.

There are also studies in the literature that additionally compare test performance as just described with a gold standard (or other) test (also within an RCT design). These types of studies investigate the diagnostic test accuracy of provocation-neutralisation. The most commonly accepted gold standard test is the double-blind placebo-controlled food challenge (see section 3.3.1). Other tests that might be used as a comparison include non-controlled oral feeding tests.

Similarly, studies investigating the effectiveness of neutralisation therapy can do so within an RCT. Again, potentially subjective (placebo) reactions can thus be controlled for.

Sensitive search strategies using several alternative terms were used for searching the electronic databases. Databases were searched as far back as possible (1966 for MEDLINE, 1980 for EMBASE) as the technique was employed as early as 1961. There were no language restrictions.

Table 1 lists the data sources searched. Full details of search strategies are listed in Appendix 1.

Table 1 Data sources primary studies

Diagnostic test accuracy studies	MEDLINE (1966- 27th September 2002)	Search terms (MeSH and text words) relating to diagnostic tests (e.g. 'specificity', 'sensitivity' etc.) were combined with terms relating to allergies ('allergy', 'hypersensitivity' etc.) and food ('food', 'food additives' etc.)
	EMBASE (1980-27th September 2002)	
	CISCOM database	Form submitted to Research Council for Complementary Medicine
RCTs or controlled trials	MEDLINE (1966- 30th September 2002)	Search filters for RCTs and controlled trials were combined with MeSH and text words relating to the condition ('food allergy', 'hypersensitivity' etc.) and intervention ('provocation', 'neutralisation' etc.)
	EMBASE (1980-3rd October 2002)	
	Cochrane Central Register of Controlled Trials, 2002, Issue 3	MeSH and text words relating to the condition ('food allergy', 'hypersensitivity' etc.) and intervention ('provocation', 'neutralisation' etc.) were used.
	CISCOM database	Form submitted to Research Council for Complementary Medicine
Ongoing research, unpublished studies	Internet Search	Web sites of ongoing trial registers, professional associations' sites, charities' and patient group sites were searched using search terms relating to the intervention
	Reference list search	Reference lists from 20 reviews and primary research studies were scanned for additional relevant primary studies
	Expert contacts	Experts sent additional material, which was checked for relevance.

4.2 Inclusion and exclusion criteria

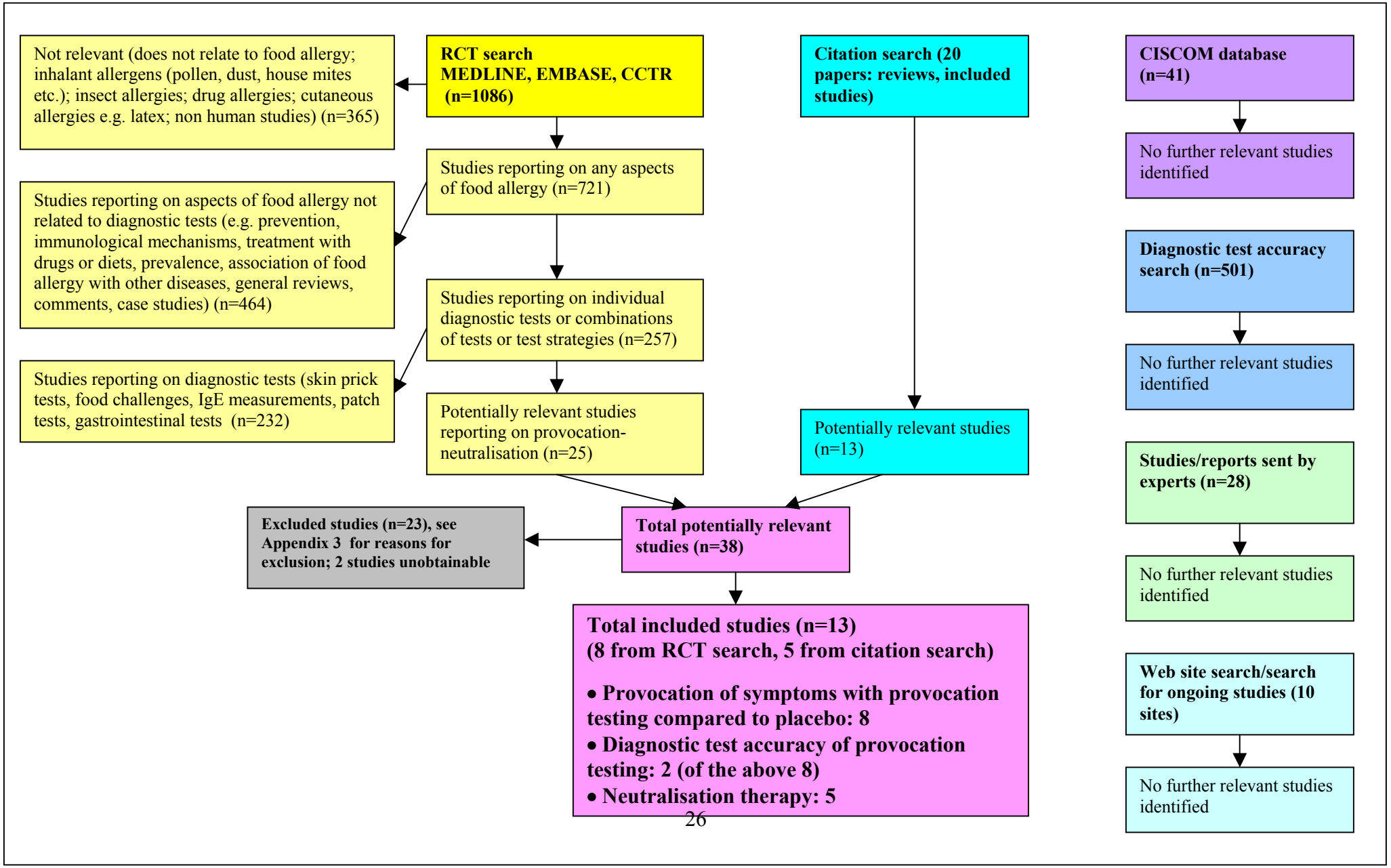
All identified studies were assessed on the basis of title and abstract by one reviewer. Where there was insufficient information to enable a decision on in- or exclusion, the full text study was retrieved. Table 2 lists the inclusion and exclusion criteria.

A flowchart of the inclusion and exclusion process is shown in Figure 2. A list of excluded studies and reasons for exclusion can be found in Appendix 3.

Table 2 Inclusion and exclusion criteria

Inclusion criteria:	
Study design	• parallel or crossover (randomised) controlled trial of: • provocation testing with active extract compared to placebo or another test (where diagnostic accuracy is investigated, sensitivity and/or specificity should be stated or calculable) *OR* • comparison of neutralisation therapy with placebo or another treatment
Population	Any population with suspected or confirmed food allergy as defined by the authors.
Intervention	Provocation-neutralisation testing and/or treatment using food allergens as defined by the authors
Outcomes	Responses as defined by the authors of the trials.
Exclusion criteria:	
Study Design	Observational studies, case reports, studies reported in the form of a letter or abstract only, animal models
Intervention	Any study testing non-food allergens only (e.g. pollen, insect venom, drugs etc.)

Figure 2 Study Identification Process



4.3 Data extraction

Data on the main study characteristics, patient characteristics, test/treatment protocols, outcomes (including adverse events) and quality was extracted directly into pre-piloted tables by one reviewer. All extracted data was checked by a second reviewer.

4.4 Quality assessment

Randomised controlled trials/controlled trials

The importance of randomisation of the test sequence during provocation testing is uncertain. One trialist stated that it is not important whether the sequence of tests is randomised[45] (presumably on the basis that all patients receive all tests and are blinded), however another[48] found that an inordinate number of positive reactions occurred when the first three or four extract were tested and concluded that the test sequence should have been randomised. It is conceivable that patients' reactions may change during the course of testing, for example if they feel initially apprehensive they may report stronger reactions, particularly as it is generally subjective symptoms that are assessed. Bias resulting from the order patients are tested in is also conceivable, however this issue has not been explored in any of the trials included in this review. Assuming the order can potentially influence results, it would also be important not to be able to predict which test sequence the next patient will receive (i.e. the test sequence should remain concealed).

The order in which active extract or placebo are used during neutralisation therapy may also have an effect on results and randomisation should ideally be used. In addition, there should be a washout period between the active extract and placebo periods, and ideally a period effect test should be performed, as treatment received in the first crossover period may influence the effect of treatment in the second period and vice versa.[49]

The key quality issue in these trials is blinding of patients, testers and outcome assessors and the inability to distinguish placebo from active extract as the outcomes assessed are generally subjective and could easily be biased by knowledge of the test or treatment substance.

Another issue is whether patients' food sensitivities are verified during the study. If they are not verified, there is no way of knowing whether a response to an active extract is a true response, a response to the test itself rather than a reflection of allergy or a placebo response.

Loss to follow-up and intention-to-treat analysis are also considered to be important as patients are more likely to drop out of a trial if the treatment is ineffective or unpleasant.

The following quality criteria were assessed for RCTs:

- randomisation
- concealment
- blinding of all relevant parties and inability to distinguish placebo from active extract

- loss to follow-up and intention-to-treat analysis
- presence of washout periods or period effect tests (for crossover trials of neutralisation therapy)
- allergy status of patients (under patient characteristics)

It was also assessed whether data was clearly presented.

Studies of diagnostic test accuracy
The most suitable design for determining test accuracy is one where a single cohort of consecutively or randomly recruited patients with unknown disease status is subjected independently and blindly to both the reference test and the test under evaluation. Selection of patients on the basis of known disease or test status or according to other pre-selection criteria can lead to bias in the estimation of test accuracy.[50,51] Due to the subjective nature of the outcomes assessed, it was additionally important to ensure that the test under evaluation (provocation-neutralisation) was performed as an RCT and that there were clear definitions of positivity or negativity.

The following quality criteria were assessed for studies of diagnostic test accuracy:
- selection of patients
- was test under evaluation performed as an RCT
- was the gold standard used as a reference test
- were the tests performed independently and blindly
- was receiving one test dependent on the results of the other
- were both tests performed in all patients
- were there clear definitions as to what constituted positive/negative test result
- were results for both tests clearly stated and was it clear how the sensitivity and specificity were calculated

4.5 Data analysis and synthesis

All extracted data was summarised in tables, and the direction of effect and quality described. Numerical pooling of data was not possible given the clinical heterogeneity (differences in type of patient, test protocol, food allergens, outcome assessment etc.). Results are presented separately for 1) provocation testing (active extract versus placebo), 2) diagnostic test accuracy analyses and 3) neutralisation therapy. It was also investigated whether results differed depending on the technique of provocation-neutralisation used (e.g. Miller technique or other).

5 Quantity of evidence identified and main study characteristics

5.1 Provocation testing

5.1.1 Number and type of studies

There were six RCTs and two non-randomised controlled trials (Breneman *et al*., 1973[48] and 1974[52] and Caplin, 1973[45]) of provocation testing (active extract compared to placebo). The main aim of these studies was to investigate the ability of active extracts to provoke symptoms and/or the association between food allergy and certain symptoms. One of these studies (detailed in Breneman *et al*., 1973[48] and 1974[52]) was composed of two parts conducted in 2 different years and reported in two papers. The studies by King *et al*., 1988 (part I[46] and part II[53], see section 5.3) investigate the reliability and accuracy of provocation testing and the effectiveness of neutralisation therapy respectively using the same patient group. All studies used a (double-blind) placebo controlled crossover design. The main study characteristics are shown in Table 4.

5.1.2 Patient characteristics

Patients were included on the basis of suspected food allergies, known food allergies, or a combination, although there were not always details on how food allergy had been verified. Another selection criterion was a previous active response to allergen and/or negative response to placebo (Jewett *et al*., 1990[54], Breneman *et al*., 1973[48]). In some cases, patients were specifically chosen by physicians to take part in the study (e.g. Breneman *et al*., 1973[48], Caplin, 1973[45] and King *et al*., 1988 part I[46]). Mandell & Conte's study (1982)[55] included individuals who volunteered or who responded to radio and newspaper.

Patient numbers varied between 15 and 132 (mean 58 patients). Ages were variable, with three studies including adults and children and three including adults only (ages not listed in Breneman *et al*., 1973[48] and 1974[52] and Caplin, 1973[45]). Based on five of eight studies where the information was given, the female:male ratio of patients was approximately 2.4:1.

Symptoms displayed by included patients were very variable, both within and between studies, and included asthma, rhinitis, gastrointestinal symptoms, arthritis and psychological symptoms. Symptoms were not always listed. The variable nature of the symptoms is in part due to the nature of food allergy or sensitivity, which manifests itself as many disorders. It should be noted that there is disagreement between conventional allergists and practitioners of environmental medicine regarding the type of disease (for example certain chronic diseases such as arthritis) or symptoms believed to be caused by food allergies or sensitivities (see section 3.1).

It is difficult to assess whether the patients included in these studies are representative of a general population with suspected or confirmed food allergies. The patients selected on the basis of positive or negative tests may differ from those patients where testing was unsuccessful. It is likely that where patients have been carefully selected or have volunteered, they will have a tendency to be more compliant with the tests.

A summary of patient characteristics is listed in Table 4. Full details of patient characteristics can be found in Appendix 4.

5.1.3 Test protocols

Four studies used sublingual testing, and four used intradermal or subcutaneous testing. The Miller technique was used by only two of eight studies and only one of these studies (Fox et al., 1999[56] reported theoccurrence of a wheal as an outcome measure (as specified by the Miller technique, see section 3.4.1). Other studies used sublingual methods or other versions of subcutaneous or intradermal testing. Several hundred provocation tests were usually performed (between 150 and over 2000) with each patient being tested with different dilutions of different allergens (between 4 and 31) over a period of a day or several days. Table 3 lists the techniques used in the trials. For full details of test protocols used, see Appendix 5.

5.1.4 Main outcome measures

The main outcome assessed was the provocation of symptoms (or relief from symptoms), with some studies using additional outcome measures. The manner in which symptoms were assessed was rarely detailed. Symptoms were generally defined as present/absent, with some studies using scales. It was not always clear if the symptoms provoked in individual patients were the same as ones previously experienced, or if they were in response to the food previously thought to be responsible for reactions.

The following outcomes were assessed:

- Provocation of any physical (objective or subjective) symptoms or relief of symptoms (Breneman *et al*., 1973[48] and 1974[52], Caplin, 1973[45], Fox *et al*., 1999[56], Jewett *et al*., 1990[54], King *et al*., 1988 (part I)[46], Mandell & Conte, 1982[55])
- Observed changes in degree of swelling and oedema of nasal mucosa (Lehman, 1980[57])
- Provocation of a wheal (Fox *et al*., 1999[56])
- Correct identification of active extract/placebo (Jewett *et al*., 1990[54])
- Range of outcome measures used, including pulse rate, signature size, Bender-Gestalt Test, cognitive and emotional self-report, somatic symptoms and others (King 1981[58]

A summary of outcomes assessed is listed in Table 4. Full details of outcomes assessed can be found in Appendix 8.

5.2 Diagnostic test accuracy

Two of the studies investigating provocation testing (Caplin, 1973[45] and King *et al*., 1988 (part I)[46]) additionally investigated the diagnostic test accuracy of provocation testing using (uncontrolled) oral feeding tests as a reference test. See section 5.1 for further details, or Table 4 for a summary.

5.3 Neutralisation therapy

5.3.1 Number and type of studies

There were five RCTs of neutralisation therapy compared to placebo. All studies used a (double-blind) placebo controlled crossover design. The main study characteristics are shown in Table 4.

5.3.2 Patient characteristics

Patients were included on the basis of a history of food allergy (either no details on how this was verified or use of single-blind or uncontrolled tests) or suspected food allergies. Another selection criterion was improvement on previous dietary management (Rapp, 1979[40]). In one study, patients were specifically chosen by physicians to take part in the study (King *et al*., 1988 part II[53]). School staff recommended for treatment those children who they thought would benefit most in the study by O'Shea & Porter (1981)[59]. In the study by Rea *et al*. (1984)[60], patients were included, who were recommended by the testing technician as appearing capable of obtaining accurate neutralising doses. Miller, 1977[41] selected patients based on symptoms judged to be caused by food allergy and on the basis of single-blind tests.

Patient numbers varied between 8 and 33 (mean 17 patients). Ages were variable, with two studies including only children (with hyperactivity), and three including adults and children. Based on three of five studies where the information was given, the female:male ratio of patients was approximately 1.3:1.

Symptoms displayed by included patients were very variable, both within and between studies, and included for three of the studies rhinitis, migraine, eczema and gastrointestinal symptoms and, for the other two studies, hyperactivity. Again, it should be noted that there is disagreement between conventional allergists and practitioners of environmental medicine regarding the type of disease (for example hyperactivity) or symptoms believed to be caused by food allergies or sensitivities (see 3.1).

It is also difficult to assess whether the patients included in these studies are representative of a general food allergy population. Patients selected on the basis of successful treatment during preliminary studies may differ from those patients where treatment was unsuccessful. Where patients have been specifically chosen it is likely that they will have a tendency to be more compliant.

A summary of patient characteristics is listed in Table 4. Full details on patient characteristics can be found in Appendix 4.

5.3.3 Test protocols

Three studies used subcutaneous neutralisation therapy, one used sublingual and one used either sublingual or subcutaneous therapy. All five studies used the Miller technique in establishing the neutralising dose subsequently used in the neutralising therapy (based on wheal and symptom provocation in three studies, symptom provocation in one and not described in one). Neutralisation therapy was evaluated over time periods of between 5-7 days and 3 weeks. Table 3 lists the techniques used in the trials. For full details of test protocols used, see Appendix 5.

5.3.4 Main outcome measures

The main outcome assessed was the improvement of symptoms.

Outcomes assessed during neutralisation therapy, and length of therapy were:
- Relief from or aggravation of symptoms (King *et al*., 1988 (part II)[53], two weeks on one treatment; Miller, 1977[41], 20 days on one treatment)
- Changes in behavioural and physical symptoms (O'Shea & Porter, 1981[59], 3 weeks on one treatment)
- Identification of neutralising dose (Rapp, 1979[40], 5-7 days on one treatment)
- Range of outcome measures used, including signs and symptoms, Visual Analogue Discomfort rating, apical heart rate and others (Rea *et al*., 1984[60], 12-18 days on one treatment)

It should be noted that no outcomes were assessed that specifically addressed (long-term) patient quality of life. It was also not clearly stated in the studies whether relief of symptoms related to those symptoms previously experienced by patients or was associated with their specific condition (e.g. asthma, eczema etc.), which food allergy was though to be a component of. Other useful outcome information might have included measures such as days off work or use of traditional health services (e.g. GP consultations).

A summary of outcomes assessed is listed in Table 4. Full details of outcomes assessed can be found in Appendix 8.

Table 3 Provocation-neutralisation techniques used by trialists

Study	Technique used
Breneman *et al*., 1973[48] and 1974[52] **Provocation testing**	1973: 1 dilution (1:40) given sublingually; 1974: 1 dilution (1:10) given sublingually
Caplin, 1973[45] **Provocation testing & diagnostic test accuracy**	1 dilution injected subcutaneously (1:100); no reference cited for methodology
Fox *et al*., 1999[56] **Provocation testing**	Miller technique (wheal and symptom provocation)
Jewett *et al*., 1990[54] **Provocation testing**	'underdose' or 'overdose' relative to ND given intradermally or subcutaneously (neutralising dose defined as dose than when injected in a volume of 0.1 ml resulted in a wheal of 7-8mm which enlarged by 2mm in 10 minutes)
King *et al*., 1988 (part I[46] and II[53]) **Provocation testing & diagnostic test accuracy (part I); neutralisation therapy (part II)**	Miller technique (wheal and symptom provocation)
King, 1981[58] **Provocation testing**	Sublingual testing using 3 different dilutions; no reference cited for method
Lehman, 1980[57] **Provocation testing**	Sublingual testing, method according to Morris*
Mandell & Conte, 1982[55] **Provocation testing**	Sublingual testing using 3 different dilutions; method according to Rinkel and Lee*, amended by authors of study
Miller, 1977[41] **Neutralisation therapy**	Miller technique (symptom provocation)
O'Shea & Porter, 1981[59] **Neutralisation therapy**	Intradermal: a 1:100 dilution tested; if test was positive, a ND was obtained using the Miller technique Sublingual: 1: 100 dilution given; if positive, a ND was obtained by giving succeedingly weaker strengths; no reference cited for method
Rapp, 1979[40] **Neutralisation therapy**	Miller technique (wheal and symptom provocation)
Rea *et al*., 1984[60] **Neutralisation therapy**	Miller technique (wheal and symptom provocation)

* for full details of methods see Appendix 5

Table 4 Main study characteristics

Study	Study description	Intervention(s)	Study design	Population/sample source	Outcomes assessed
Provocation-neutralisation testing					
Breneman *et al*., 1973, USA[48] (committee report)	Comparison of responses to active extracts and placebo	Sublingual provocative testing (food allergens)	Placebo-controlled double-blind crossover tests	In 1973: Patients (n=100) suspected of having food allergy and with a negative sublingual response to placebo; 10 patients each chosen by 10 allergists	Provocation of symptoms
Breneman *et al*., 1974, USA[52] (committee report)				In 1974: Patients (n=30) known to have an allergy to at least one of the test foods; 5 patients each chosen by 6 allergists	
Caplin, 1973, USA[45] (committee report)	Diagnostic test accuracy study (comparison of provocation testing results with those of oral challenge feeding test)	Subcutaneous provocative testing (food allergens) and oral challenge feeding test	Provocative testing: Placebo-controlled double-blind crossover tests Oral challenge feeding test: no details (appears to be uncontrolled, unblinded)	Protocol specifies three different types of patient groups (inhalant allergy, food sensitive and non-allergic); 8 physicians contributed 70 patients in total (55 atopic, 15 non-atopic patients)	Provocation of symptoms
Fox *et al*., 1999, Canada[56]	Comparison of responses to active extracts and placebo	Intradermal provocative testing (food and non-food allergens)	Placebo-controlled double-blind crossover tests	Patients (n=132) with possible/probable chemical sensitivity with symptoms that were disruptive to normal life or who were experiencing life-threatening anaphylaxis; referral by primary care physician or specialist	Provocation of wheal or symptoms
Jewett *et al*., 1990, USA[54]	Comparison of responses to active extracts and placebo	Intradermal or subcutaneous provocation testing; neutralisation therapy in	Placebo-controlled double-blind crossover tests	Patients (n=18) included on basis of previous consistent active responses to injections of allergens (no response to diluent alone) during	Provocation of symptoms and ability of patient to

Study	Study description	Intervention(s)	Study design	Population/sample source	Outcomes assessed
		7/18 patients (food allergens and mould)		unblinded testing	identify allergen/placebo
King *et al*., 1988, USA (part I)[46]	Test of the reliability of provocation testing (3 series of identical provocation tests); Diagnostic test accuracy study (comparison of provocation testing results with those of oral challenge food test)	Intracutaneous provocative testing (food allergens) and oral challenge food test	Provocation testing: 3 series of placebo-controlled double-blind crossover food tests; Oral food challenge: open (uncontrolled, unblinded)	Patients (n=37) with a variety of symptoms (gastrointestinal, bronchopulmonary, skin problems and others), who had never undergone provocation-neutralisation testing or allergy treatment before; Physicians experienced with the intracutaneous provocative food test technique asked to contribute patients with food sensitivity	Provocation of symptoms
King, 1981, USA[58]	Comparison of responses to active extracts and placebo	Sublingual provocation testing (food and non-food allergens)	Placebo-controlled double-blind crossover tests	Patients (n=30) with at least one psychological symptom (e.g. anxiety, depression, confusion, difficulty in concentrating); patients were new allergy outpatients	Self-report of cognitive-emotional or somatic symptoms and other measures (e.g. pulse rate, signature size, Bender-Gestalt Test, Mood Affect Adjective Checklist (MAACL) etc.)
Lehman, 1980, USA[57]	Comparison of responses to active extracts and placebo	Sublingual provocative testing (food allergens)	Placebo-controlled double-blind crossover tests	Patients (n=15) with a history of a variety of symptoms (e.g. allergic rhinitis, gastroenteritis, atopic eczema) after eating certain foods; patients selected based on clinical history	Observed changes in degree of swelling and oedema of nasal mucosa (nasocyrogram)
Mandell & Conte, 1982, USA[55]	Comparison of responses to active extracts and placebo	Sublingual provocation testing (food and non-food allergens)	Placebo-controlled double-blind crossover tests	Patients with arthritis and rheumatism (n=30); patients were volunteers from Dr. Conte's practice, referrals from Easter Seal Centre and respondents to radio and newspaper	Provocation of symptoms

Study	Study description	Intervention(s)	Study design	Population/sample source	Outcomes assessed
Neutralisation Therapy					
King *et al*., 1988, USA (part II)[53]	Follow-up study to King *et al*., 1988 (part I); Comparison of neutralisation therapy with placebo	Subcutaneous neutralisation therapy	Placebo-controlled double-blind crossover study	As King part I; 7/8 physicians contributed 33 patients; neutralisation therapy performed in those patients where it was thought to be appropriate	Response of symptoms to treatment
Miller, 1977, USA[41]	Identification of neutralising dose; Comparison of neutralisation therapy with placebo	Intradermal provocation testing (food allergens); subcutaneous neutralisation therapy	Provocation testing: Single-blind Neutralisation therapy: double-blind placebo controlled crossover trial	Patients (n=8) with a history of food allergy (with symptoms including migraine, hyperactivity, gastrointestinal problems and others); Patients who presented with symptoms and syndromes judged to be caused by food sensitivities and who responded well to food injection therapy were included	Relief from or occurrence of symptoms
O'Shea & Porter, 1981, USA[59]	Identification of neutralising dose; Comparison of neutralisation therapy with placebo	Intradermal and sublingual provocation testing (food and non-food allergens); sublingual neutralisation therapy	Provocation testing: not detailed, appears to be uncontrolled Neutralisation therapy: Double-blind placebo controlled crossover study	Children (n=15) who met the clinical criteria of hyperkinetic syndrome; selected by staff at the Lawrence public schools according to who they felt was in most need of treatment	Provocation of behavioural symptoms
Rapp, 1979, USA[40]	Identification of neutralising dose; Comparison of neutralisation therapy with placebo	Intradermal provocation testing (food allergens); sublingual or subcutaneous neutralisation therapy	Provocation testing: uncontrolled Initial neutralisation therapy: uncontrolled Subsequent trial of neutralisation therapy: placebo-controlled double-blind crossover trial	Hyperactive patients (n=11) who responded favourable to dietary management; 11/23 patients selected who were part of a preliminary study on the efficacy of diet in the management of hyperactivity and who showed improvement on food omission diet	Sore derived from Parent Abbott Hyperkinesis Index sheet (variation of Connor's Child Behaviour Rating Scale); Identification of coded food solution by parents

Study	Study description	Intervention(s)	Study design	Population/sample source	Outcomes assessed
Rea *et al*., 1984, USA[60]	Identification of neutralising dose; Comparison of neutralisation therapy with placebo	Intracutaneous provocation testing (food allergens); subcutaneous neutralisation therapy	Open diagnostic food challenge Provocation tests: Method according to Miller, no further details Neutralisation therapy: placebo-controlled double-blind crossover trials	Patients (n=20) with suspected food or inhalant allergies with a wide range of symptoms (e.g. eczema, diarrhoea, myalgia and others) and who exhibited symptoms at least one hour after oral food challenge; patients selected from 150 consecutive admissions, who fulfilled the selection criteria and were willing to participate	Signs and symptoms; Visual Analogue Discomfort rating; Symbol-Digit Modalities Test; Aaron Smith Symbol-Digit Modalities Subtest; apical heart rate; subject's signature; laboratory studies (T-lymphocyte, B-cell, complement, total IgE count)

6 Direction of effect and quality of evidence

6.1 Provocation testing

6.1.1 Study quality

Randomisation & Concealment
The 1st part of the study by Breneman *et al*. (1973)[48] and the study by Caplin, 1973[45] did not use randomisation. The study by Lehman, 1980[57] and the 2nd part of the study by Breneman *et al*. (1974)[52] did not mention randomisation. Three studies were described as randomised, but gave no details on the method (Fox *et al*., 1999[56], King, 1981[58], King *et al*. 1988 (part I)[46]). The remaining two studies stated that randomisation was achieved by die and coin toss (Jewett *et al*., 1990[54]) or that a code was selected for the vials by random selection from the alphabet (Mandell & Conte, 1982[55]).

It is important to note that randomisation as a method of achieving baseline equivalence is compromised by small study sizes (<50). Five of the eight studies contained less than 50 patients.

One study stated that the sequence was unknown (Caplin, 1973[45]) and one that the order of tests was kept concealed (Jewett *et al*., 1990[54]). In the study by Breneman *et al*. (1973)[48] and (1974)[52] it was stated that the code was kept by laboratories until the testing was complete.

Blinding & Ability to distinguish placebo
Four studies had details on blinding (patients and testers and/or outcome assessors) and on the appearance of the placebo or on attempts to minimise recognition (Caplin, 1973[45], Fox *et al*., 1999[56], Jewett *et al*., 1990[54], King *et al*. 1988 (part I)[46]). The study by Lehman, 1980[57] had details on the placebo, but none on blinding, and the study by Mandell & Conte, 1982[55] had details on blinding but none on the placebo. The study by Breneman *et al*.(1973[48] and 1974[52]) did not explicitly comment on blinding of patients or testers, but stated that the code was kept by the laboratories until testing was complete. There were no details on the placebo (1974[52]study) and some concerns that the active allergen extract containing chocolate was probably identifiable (1973[48] study). The study by King, 1981[58] relied on experimenters administering solutions with their eyes closed in order to maintain blinding, which appears to be a very weak method. The author also stated that placebo aware or experimenter suspicious trials were removed from the analysis, which suggests that the placebo or active extract could be identified by some patients or experimenters.

Loss to follow-up, ITT & outcome assessment
Five studies had no loss to follow-up (Caplin, 1973[45], Fox *et al*., 1999[56], Jewett *et al*., 1990[54], King *et al*. 1988 (part I)[46] and Mandell & Conte, 1982[55]). The study by Breneman *et al*. (1973[48] and 1974[52]) lost 39/100 patients to follow-up in the first part of the study and 10/30 in the second part; there was no ITT (intention-to-treat) analysis. Placebo aware or experimenter suspicious trials were removed from the analysis of the study by

King, 1981[58], again there was no ITT analysis. In the study by Lehman, 1980[57], 2/15 patients were lost to follow-up and 5 additional patients were added to the analysis six years later.

Outcome assessment consisted of the occurrence of (any) symptoms being reported by patients and recorded by observers. There were few other details. Scales of severity were not used. In the study by Breneman *et al*., 1973[48], it was stated that patients were shown a list of suggested symptoms, which may have biased results.

The study by King *et al*., 1988 (part I)[46] did not list results for placebo tests and the study by Lehman, 1980[57] assessed an outcome (observed changes in degree of swelling and oedema of nasal mucosa), which is described by the author himself as difficult to interpret. In the study by Breneman *et al*., 1974[52], it was stated that a feeding challenge was to be conducted, however results for this were not reported. The study by King, 1981[58] reports in more detail those results where a significant effect is seen, and gives few details on those results where there is no difference.

Verification of allergy status
There was variation between the studies regarding how or whether food sensitivities had been verified. There were no details regarding verification in the studies by Breneman *et al*., 1973[48], Fox *et al*., 1999[56] or Mandell & Conte, 1982[55]. The study by Jewett *et al*., 1990[54] stated that only patients where symptoms had been consistently provoked during unblinded testing were included, whilst the study by King, 1981[58] included patients who responded positively to one of 5 test foods when orally challenged. The studies by Breneman *et al*., 1974[52] and King *et al*. 1988 (part I)[46] included patients who responded positively to at least one of the included foods, and the study by Caplin, 1973[45] included patients with and without known food sensitivities to at least one food and patients with known inhalant allergies. Lehman, 1980[57] included patients with symptoms after eating certain foods and/or improving during an elimination period.

Summary
Based on the information reported in the publication only, the study with the best internal validity appears to be Jewett *et al*., 1990[54], whilst the studies with the poorest internal validity appear to be Breneman *et al*. (1973[48] and 1974[52]), King, 1981[58] and Lehman, 1980[57]. The study by Caplin, 1973[45] appears adequately conducted, however there was no randomisation, which may have compromised the whole study. The study by King *et al*., 1988 (part I)[46] appears to be well conducted, however it is not possible to interpret the results as there is no placebo data.

Full details of the quality assessment can be found in Appendix 6**.**

6.1.2 Direction of effect

The eight studies comparing provocation testing with active extract compared to placebo found that:

- in four cases there was no difference in provocation of symptoms on active solution or placebo (Breneman *et al*. (1973[48]and 1974[52], Caplin, 1973[45], Jewett *et al*., 1990[54], Lehman, 1980[57]; based on the occurrence of symptoms there was no difference in the number of times an active or placebo solution were judged to be active (Jewett *et al*., 1990[54])

- in two cases a statistically significant greater number of psychological and somatic (King, 1981[58]) and physical symptoms (Mandell & Conte, 1982[55]; no statistical tests performed) were provoked on active solution

- in one case there was no difference regarding provocation of symptoms, but more wheals were provoked on active solution (Fox *et al*., 1999[56]; no statistical tests performed)

The study by King *et al*., 1988 (part I)[46] states that there was a good correlation between positive and negative symptom provocation for consecutive series of provocation tests, however it is not stated whether these positive and negative reactions occur on active extract or placebo. There is therefore no way of knowing whether this could be attributed solely to the placebo effect (i.e. someone with one positive reaction is more likely to expect a second one) or whether there was any difference in symptom provocation on active extract or placebo.

All studies showing no difference used techniques other than the Miller technique. The two studies showing a benefit used sublingual testing. The study showing an increased number of wheals on active extract (Fox *et al*., 1999[56]) was the only interpretable study using the Miller technique.

No attempt has been made to quantitatively synthesise results, as studies were heterogeneous in terms of patients' initial symptoms, symptoms provoked, test protocols used and the way (if detailed) in which outcomes were assessed.

Two studies listed adverse effects. The study by Caplin, 1973[45] stated that one patient suffered severe asthma in response to placebo, and King, 1981[58] stated that there were 18 requests for relief from uncomfortable symptoms.

A summary of results can be found in Table 5. Full details of the results can be found in Appendix 8.

6.2 Diagnostic test accuracy

6.2.1 Study quality

The two studies (Caplin, 1973[45] and King *et al*. 1988 (part I)[46]), which also investigated diagnostic test accuracy were of poor methodological quality. The selection criteria for patients and their disease status were unclear (selection of a particular patient group can bias the sensitivity and specificity). An open food challenge was used in both studies as the reference standard. As the assessed symptoms are of a subjective nature, this is likely to result in biased results depending on patients' and outcome assessors' preconceptions. It was not clear if the two tests (food challenge and provocation-neutralisation testing) were performed blindly (i.e. without knowledge of any test results). In the study by Caplin, 1973[45], only 48/70 patients received the food challenge test, reasons for this are not stated. In the same study, a milkshake containing all the test foods was used for the oral food challenge. It is not clear how the authors established which food the patient was reacting to or if it was the same food that provoked symptoms during provocation testing. It is not clear in the study by King *et al*. 1988 (part I)[46] how the sensitivities and specificities were calculated (false positives and false negatives not stated).

Full details of the quality assessment can be found in Appendix 7.

6.2.2 Diagnostic test accuracy

Based on results in 48 patients in the study by Caplin, 1973[45], the sensitivity was 0.80 (95% CI 0.52-0.96, calculated by author of this review) and the specificity was 0.64 (95% CI 0.54-0.73, calculated by author of this review).

Based on results from 33 patients in the study by King *et al*., 1988 (part I)[46], the sensitivity and specificity based on skin response was 79.7% and 72.4% respectively, and the sensitivity and specificity based on symptom provocation was 59.6% and 92.1% respectively. Confidence intervals were not calculable, as the number of false positives and false negatives was not stated.

6.3 Neutralisation therapy

6.3.1 Study quality

Randomisation & Concealment
Studies were described as random (O'Shea & Porter, 1981[59]) or it was stated that the treatment set was randomly coded (Rapp, 1979[40]), or that the order was determined arbitrarily (Rea *et al*., 1984[60]) or arbitrarily by lot (King *et al*., 1988 (part II[53]), but there were no further details on the methodology. One study (Miller, 1977[41]) stated that the first extract was chosen by coin flip. Rea *et al*., 1984[60] additionally stated that the observer did not know the order of trials.

All five studies included less than 50 patients, which could potentially result in uneven distribution of baseline characteristics during randomisation (the total number of patients from all five studies was n=87).

Blinding & Ability to distinguish placebo
All studies gave some indication that they were blinded, with 3 studies giving more explicit information about who was blinded (patients (King *et al*., 1988 (part II)[53]), patients and physicians (Miller, 1977[41]), and patient, technicians and observers (Rea *et al*., 1984[60])).

All studies stated that the placebo was identical to or indistinguishable from the active extract. Two studies stated that tests were performed to establish whether the placebo was identifiable: Rea *et al*., 1984[60] in a separate double-blind study using volunteers and King *et al*., 1988 (part II)[53] at the beginning of their study.

Loss to follow-up, ITT & outcome assessment
Three studies had no loss to follow-up (King *et al*., 1988 (part II)[53], Miller, 1977[41] and Rea *et al*., 1984[60]), the other two studies lost 1/15 (O'Shea & Porter, 1981[59]) and 3/11 patients (Rapp, 1979[40]), there was no ITT analysis. The study by Miller, 1977[41] stated that it was a preliminary study, with results presented for the first 8 patients. No follow-up studies were identified despite extensive searching. Patient numbers in the studies were small (between 8 and 33, mean 17) and no power calculations were performed to establish what sample size would be needed to show a difference in effect.

Outcome assessment was variable in terms of being quantitative or mainly qualitative. Assessment in the study by Rapp, 1979[40] was based on the correct identification of the active extract by parents based on behavioural improvements in their children. Similarly, children's behavioural improvements were noted in a diary kept by parents, and parents and teachers were interviewed weekly by a psychologist in the study by O'Shea & Porter, 1981[59]. Patients in the study by Miller, 1977[41] used a scale from 0 to +4 to grade symptoms according to intensity, duration and frequency. The study by King *et al*., 1988 (part II)[53] also used a symptom scale from 1 (much worse) to 6 (excellent relief). The study by Rea *et al*., 1984[60] used six outcome measures. There was some concern around how some of the results were quantified. For the 'signs and symptoms' outcome assessment, a scale from 0-4 was used (increasing numbers equalling more symptoms). However, results were not stated in a numerical manner but divided into positive or negative, with negative changes (i.e. a protective effect of the ND) being grouped together with slight increase in symptoms. A cut-off point for positivity or negativity is not stated. Similarly, for the Visual Analogue Scale, results are not stated numerically, but again divided into positive or negative depending on whether there was a change of 1cm or more from baseline. There is no information on absolute values for discomfort on placebo or active extract. The outcome 'subject's signature' was classified as positive or negative depending on the presence or absence of substantial deterioration in quality.

Washout periods
There was a one week gap between treatment/placebo periods (study by King *et al*., 1988 (part II) [53]and O'Shea & Porter, 1981[59]) and a 2-4 day gap in the study by Rea *et al*., 1984[60]. There were no details on washout periods for the studies by Miller, 1977[41] or Rapp, 1979[40]. No period effect tests were performed.

Verification of allergy status
There was some uncertainty around the verification of individual food allergies. King *et al*., 1988 (part II)[53] and Rea *et al*., 1984[60]included patients who responded to foods one hour after oral food challenge. Miller, 1977[41] selected patients on the basis of symptoms judged to be caused by food allergy and on the basis of single-blind tests. Rapp, 1979[40] included patients, who had previously responded to omission diets and there were no details on verification of allergy status in the study by O'Shea & Porter, 1981[59].

Summary

Overall the studies were of similar (adequate) study quality, although there is some uncertainty regarding the verification of allergy status. The study by Rea *et al*., 1984[60] appears to have the best internal validity, there were however concerns around the outcome assessment in this study. The study by Miller, 1977[41] appears to be well conducted, it is however only a preliminary report and no follow-up reports were identified.

Full details of the quality assessment can be found in Appendix 6**.**

6.3.2 Direction of effect

Three studies showed a better response on active treatment compared to placebo:

• Rea *et al*., 1984[60] found a statistically significant protective effect against adverse reactions with active treatment compared to placebo (statistically significant for 6 different outcome measures, using number of positive or negative reactions rather than numerical data: signs and symptoms, Visual Analogue Discomfort rating, Symbol Digit Modalities Test, Aaron-Smith Symbol-Digit Modalities subtest, apical heart rate, subject's signature)

• Miller, 1977[41] found in a preliminary study of 8 patients, a statistically significant mean symptom improvement on neutralising dose compared to placebo

• King *et al*., 1988 (part II)[53] found that the symptoms of a greater number of patients improved on active solutions compared to placebo (65.2% improved versus 34.8% worsening or no improvement; statistical tests were not performed)

One study showed little or no difference in the ability to correctly identify the active solution from placebo:

• Rapp, 1979[40] found that 5/8 parents correctly identified the active extract, whilst 3/8 parents incorrectly identified the placebo as active extract (statistical tests not performed)

One study found a beneficial effect or no effect from neutralisation therapy, depending on how the outcome was assessed:

• O'Shea & Porter, 1981[59] found that 11/14 extracts were correctly identified as active based on an improvement in behaviour in 11/14 children (parents' assessment); 7/13 children were rated by teachers as having improved behaviour; statistical tests not performed

No attempt has been made to quantitatively synthesise results or to calculate an overall effect size as outcomes were assessed in different ways (e.g. mean symptom scores in one study compared to correct identification of active extract in another) and could not be directly compared, and studies were heterogeneous in terms of patients' symptoms, allergens used in neutralising dose and method of neutralisation therapy.

All studies used the Miller technique (symptom only or symptom and wheal provocation), the study by O'Shea & Porter, 1981[59]additionally used a sublingual technique in some patients.

Two studies listed adverse events. The study by O'Shea & Porter, 1981[59] stated that one child's behaviour deteriorated to such an extent that Ritalin had to be reinstated, and Rapp, 1979[40] stated that two children developed behavioural problems and parents refused to complete the study.

A summary of the results is listed in Table 6. Full details of the results can be found in Appendix 8.

6.4 Summary of effectiveness and quality

Table 5 and Table 6 show an overview of the direction of effect and methodological quality. It should be noted that unless 'Miller' is stated, the technique does not correspond to the Miller technique. Full details of outcomes and study quality (RCTs and diagnostic test accuracy) can be found in Appendix 8, Appendix 6 and Appendix 7 respectively.

Table 5 Overview of direction of effect and study quality: provocation testing

Provocation testing				
Study	**Technique**	**Study Size**	**Direction of effect**	**Comment**
Breneman *et al*., 1973[48]and 1974[52]	Sublingual	n= 130	No difference in symptom provocation	No randomisation; no intention-to-treat analysis; no details on the verification of allergy status
Caplin, 1973[45]	Subcutaneous	n=70	No difference in symptom provocation	No randomisation; no details on the verification of allergy status
Fox *et al*., 1999[56]	Miller	n=132	No difference in symptom provocation; more wheals provoked on active extract	No details on the verification of allergy status
Jewett *et al*., 1990[54]	Intradermal or subcutaneous	n=18	No difference in symptom provocation	Verification of allergy status unclear, otherwise good methodological quality
King *et al*., 1988 (part I)[46]	Miller	n=37	Good correlation between successive positive and negative provocation tests	Uninterpretable, as results for placebo not detailed
King, 1981[58]	Sublingual	n=30	More psychological and somatic symptoms provoked on active extract (statistically significant); no difference for 10 other outcome measures	No details on the verification of allergy status; no intention-to-treat analysis; blinding potentially compromised
Lehman, 1980[57]	Sublingual	n=15	No difference in swelling and oedema of nasal mucosa	Outcome measure difficult to interpret; additional patients added to analysis 6 years after study
Mandell & Conte, 1982[55]	Sublingual	n=30	More symptoms provoked on active extract (no statistical tests performed)	No details on the verification of allergy status

Diagnostic test accuracy				
Study	**Technique**	**Study Size**	**Direction of effect**	**Comment**
Caplin, 1973[45]	Subcutaneous versus open oral food challenge	n=70 (provocation testing); n=48 (open food challenge)	Sensitivity 0.80 (95% CI 0.52-0.96); specificity 0.64 (95% CI 0.54-0.73), confidence intervals calculated by JD	Poor methodological quality (gold standard not used as reference test, only 48 patients received feeding test; for full details see Appendix 7)
King *et al*., 1988 (part I)[46]	Miller versus open oral food challenge	n=37 for both tests	Sensitivity 79.7%, specificity 72.4% (based on skin response); sensitivity 59.6%, specificity 92.1% (based on symptom provocation); confidence intervals not calculable	Poor methodological quality (gold standard not used as reference test, patients selected on basis of food sensitivity; for full details see Appendix 7)

Table 6 Overview of direction of effect and study quality: neutralisation therapy

Neutralisation therapy				
Study	**Technique**	**Study Size**	**Direction of effect**	**Comment**
King *et al*., 1988 (part II)[53]	Miller	n=33	Improvement of symptoms on active extract (65.2% versus 34.8%); no statistical tests	Concerns around allergy verification: patients included on the basis of response to foods when orally challenged and results of a previous study (King *et al*., 1988 part I[46]) where results are uninterpretable (included in this review)
Miller, 1977[41]	Miller	n=8	Mean improvement in symptoms (statistically significant)	Preliminary study results in 8 patients; no follow-up study identified despite extensive searching; allergy verification relied on single-blind tests (no details on type of test)
O'Shea & Porter, 1981[59]	Miller or sublingual	n=15	Improvement in behaviour of 11/14 children (parents assessment) or 7/13 children (teachers' assessment); no statistical tests	No details on the verification of allergy status
Rapp, 1979[40]	Miller	n=11	Similar correct/incorrect identification of active extract (5/8 versus 3/8 parents)	21% (3/11) patients lost to follow-up, no intention-to-treat analysis; patients included on basis of response to previous omission diets
Rea *et al*., 1984[60]	Miller	n=20	Protective effect with ND using 6 outcome measures (all statistically significant)	Concerns around outcome assessment: assessment and scoring systems not defined for all outcome measures; some results not stated numerically divided into positive or negative results with no cut-off stated

7 Economic evaluation

A search was performed to identify studies relating to the cost or cost-effectiveness of provocation-neutralisation (testing or therapy). The full strategy is detailed in Appendix 2. No relevant studies were identified.

Dr Mansfield provided information on costs of treatment at the Burghwood Clinic, Surrey (personal communication). A first one-hour consultation costs £95.00. Two and a half hours of skin testing performed by a trained allergy nurse costs £75.00. Total costs will vary depending on how many consultations or treatment sessions are needed. A patient with several food allergies may for example need one primary consultation, four follow-up consultations and two skin testing sessions, amounting to around £500 in total. On occasions, a patient with multiple food sensitivities, chemical sensitivities and biological inhalant sensitivities may have to pay in excess of £2000 over a 1 or 2 year period.

Costs for other tests vary depending on the type of test and the provider, how many tests are needed etc. The following prices are provided as examples only. A full consultation and allergy work-up, which includes skin prick tests and lung function tests costs £100, with a follow-up consultation fee of £50 at the Surrey Allergy Clinic.[61] An ELISA test can cost around £200 for a full screening test.[62] YORKTEST Laboratories charge between £135 or £260 for test kits based on the detection of elevated levels of IgG.[63] Conventional treatment in the form of an elimination diet (avoidance of the relevant food) would entail no drug costs, but may require dietary advice to be given to the patient.

Prices of individual tests or test kits are difficult to compare directly, as patients will differ regarding the types of allergies or sensitivities they have, the types of tests they need, how many test sessions they require and how effective the tests are.

Data for a number of parameters would be required for a decision analytical model to determine the cost-effectiveness of a particular test/treatment strategy. These include the prevalence and incidence of food allergy or sensitivity, the test accuracy of the different tests or test strategies (number of true and false positives and negatives), the cost of the different tests or combinations of tests and of treatment, the utility of the health states associated with untreated and treated food allergy, the disutility of the tests and treatment and the degree of compliance with treatment. Given that much of this data would be difficult to obtain or estimate, such an analysis was beyond the scope of this review.

Based on the studies identified for this report, there is at present no evidence to suggest that provocation-neutralisation would be useful as a diagnostic tool, therefore an economic evaluation would not be appropriate. For neutralisation therapy there is some evidence to suggest effectiveness, however there is some uncertainty around the validity of the results. There is also a lack of data on the impact of neutralisation therapy on long term quality of life. It should be noted that even modest reductions in the use of GP services might offset the relatively modest costs of testing.

8 Discussion

8.1 Discussion of effectiveness

Based on the results of this review, there is little evidence to suggest that more symptoms are provoked in response to active extract compared to placebo. The majority of studies suggest there is no difference. Most of the evidence is of poor quality and there are design issues, which make the interpretation of results difficult. One of the major issues is the fact that the allergy status of patient is not always verified, which means that it is not possible to know whether a response to an active extract is actually a placebo response. Any conclusions regarding the association of food allergy with a specific disease must therefore be viewed cautiously.

All results must be interpreted in the context of the test method used. Only one interpretable study looking at provocation testing (Fox *et al*., 1999[56]) used the Miller technique and there is some evidence to suggest that more wheals are provoked on active extract. Again there is uncertainty around the verification of patients' allergy status, which raises questions around the clinical relevance of results.

No study, including the two looking specifically at diagnostic test accuracy, which were both of poor methodological quality, provides evidence to suggest that the technique may be of useful for diagnosis of specific food allergies (although it should be noted that most studies did not have this as an explicit aim). Again, the results must be interpreted in the context of the test methodologies used. One criticism of the two studies was the fact that an unblinded food challenge was used as a reference test. Whilst the double-blind placebo-controlled food challenge is considered by many to be an appropriate reference test, there may also be problems with this for example with regard to delayed reactions or the appropriate way to mask foods.

There is some evidence to suggest that neutralisation therapy can be effective in the relief or prevention of symptoms, particularly from the studies by Rea *et al*. 1984[60], Miller, 1977[41] and King *et al*., 1988 (part II)[53], with a total of 61 patients. There are however a number of issues, which should be taken into account. These include the fact that the study by Miller, 1977[41] was a preliminary study in 8 patients only (no follow-up study was identified), and the concerns around the outcome assessment in the study by Rea *et al*. 1984[60]. A further issue is the initial verification of the food allergy or sensitivity to be treated. Unless there is good evidence that the treatment relates to a food the patient is definitely sensitive to, and that the symptoms experienced definitely relate to food allergy, any subsequent results are meaningless. Miller, 1977[41] stated that single-blind tests (no details on the type of test) were used to test for those foods suspected to be the cause of symptoms. King *et al*., 1988 (part II)[53] included patients who responded to foods when orally challenged, and on the basis of the results of a previous study, where the results are uninterpretable (see King *et al*., 1988 (part I)[46]). Rea *et al*., 1984[60] selected patients on the basis of exhibiting one of a variety of symptoms one hour after an oral food challenge. The results of these studies must be interpreted in the context of whether

it is likely that a definitive diagnosis was established.

Many of the patients included in the studies were selected on the basis of previous good responses to tests/treatment or recommended by physicians, particularly for those studies looking at neutralisation therapy, which appeared to be effective. In the study by Rea *et al*. 1984[60], patients were included, who were recommended by the testing technician as appearing capable of obtaining accurate neutralising doses. These patients are unlikely to be representative of a general food allergy population, and it is therefore possible that the technique may be more or less effective depending on the type of patient. Included patients were also generally very heterogeneous in terms of their symptoms, whilst it was not always clear which foods (or how many) they were allergic to. It is uncertain whether the trials would be reproducible in other patient groups.

Due to the subjectivity of outcome assessments, one of the key factors in maintaining the internal validity is ensuring that all relevant parties are blinded and that the placebo cannot be distinguished from the active extract. Only two studies included in this review conducted tests to investigate whether it was possible to distinguish between the two solutions, whilst in two further studies there was a problem in that patients appeared to became placebo aware. It is conceivable that this may have been an issue in other included studies, although generally attempts at making the placebo identical to active extracts were described. Outcome assessment involved the use of mainly subjective outcome measures (i.e. presence or absence of symptoms, relief from symptoms). Whilst blinding of all parties involved should eliminate bias, if the blinding were compromised in any way, the measurement of subjective symptoms is more likely to result in bias.

There were several issues around outcome assessment and presentation of results that should be considered. Data was often not presented numerically (where numerical scales were used) or was not presented separately for active extracts and placebo (or for individual patients), but was instead described in terms of a positive or negative test or a patient's relative improvement compared to placebo. This made it difficult to assess the size of the difference in effect. The study by King *et al*., 1988 (part I)[46] does not report results for placebo tests at all. Rea *et al*., 1984[60] does not always clarify where the cut-off point for classifying a test as positive or negative is. The study by Breneman *et al*., 1973[48] states that a feeding challenge was performed, however the results are not reported, whilst the study by King, 1981[58] focuses the reporting on those outcomes that show a difference in effect. The study by Lehman, 1980[57] added additional patients to the analysis after a six year period, whilst Miller, 1977[41]reports the results of a preliminary report only (no follow-up report was identified). Statistical tests were also not always performed, with no reasons given as to why this might be appropriate. Poor or selective reporting can clearly weaken the credibility of studies. King (1988)[64] states that dichotomous ratings (reactions present or absent) are less reliable than a graded scale (e.g. slight reaction to severe one). At least six of the included studies used only a present/absent classification for the occurrence of symptoms. Similarly there was a lack of detail on whether the same symptoms as previously experienced by the patient were being provoked or whether they were different symptoms.

An editorial in the BMJ has discussed the effect conflict of interest might have on results.[65] It has been shown that authors are more likely to be supportive of a technique if they have financial or other interests. Some of the included studies were conducted at the private clinics of the authors (Mandell & Conte, 1982[55]) or sponsored by organisations (Rea *et al*., 1984[60]), which support the techniques, however it is impossible to say what the effect of this might be.

Whilst no study included in this review reported very serious side effects, provocation-neutralisation may pose risks to some patients with IgE mediated food allergy. Teuber & Vogt[66] report a case of a patient with systemic mastocytosis who experienced potentially life-threatening anaphylactic reactions during provocation-neutralisation treatment. The authors also refer to a report of angio-oedema following application of sublingual drops where the patient had an IgE mediated food hypersensitivity.

8.2 Limitations of review

Whilst an attempt was made to identify all relevant studies, it is possible that some studies may have been missed, particularly as there are no specific MeSH terms associated with this technique and indexing terms used were very variable. However, additional searches were performed (citation searching and web searching) and no additional studies were identified through expert contacts.

Any systematic review can be subject to publication related biases, e.g. grey literature bias, language bias, time lag bias or publication of predominantly positive studies. It was not possible due to the nature of the results to formally assess publication bias. Contact with experts may have helped to limit grey literature bias.

Ideally, the inclusion and exclusion criteria would have been applied independently by another reviewer, and disagreements resolved through discussion, however this was not possible due to limited resources. Similarly, data extraction would ideally have been performed independently by a second review, again this was not possible. Extracted data was however checked by a second reviewer.

Studies were excluded if they were reported in the form of an abstract or letter only. As there was insufficient information to assess the internal and external validity of the studies, it was felt that including them would not further inform the discussion. The results do however generally either support the conclusions of this review (provocation testing) or do not further strengthen the results (neutralisation therapy): Bronsky *et al*., 1971[67] and Crawford *et al*., 1976[68] found that the technique was not useful for diagnosing food allergy, whilst Kailin & Collier, 1971[69] found no more relief of symptoms from the neutralising dose than from a placebo.

Assessment of quality was based on the reported studies only. Therefore the poor quality may in some cases be a reflection of the reporting rather than the study quality itself.

It would have been of interest to investigate the validity of the outcome measures used, as this would have helped in the assessment of the validity of the results. This was not undertaken due to time constraints.

The review was limited to food allergies. It would potentially be of interest to investigate the evidence base surrounding the application of the technique for inhalant or other allergens.

8.3 Recommendations for future research

A number of methodological issues have been identified, which make the interpretation of the current study results difficult. Further well-designed RCTs are therefore essential. These should ideally avoid a cross-over design, which can be difficult to interpret (one of the drawbacks of a cross-over design is that the effect of the treatment received in the first study period may carry over into the next treatment period).

The following issues should be considered when designing any future studies:

- Test or treatment sequences should be randomised and the sequence kept concealed in order to avoid any potential bias resulting from the test order.

- In order to avoid potential identification of the placebo, future studies should perform preliminary tests to ensure there is a system whereby neither patients nor testers or observers are able to distinguish between placebo and active extract at any point during the study.

- There should be agreement on the model of provocation-neutralisation to be used, otherwise the conclusions drawn from any future trial may be contested on the basis that a different technique was used to that most commonly employed. The Miller technique itself appears to have changed over time, as interpretation previously relied on the provocation of symptoms, but a more recent description focuses on the appearance /disappearance of a wheal. Similarly, a well-defined model is needed for sublingual testing. This should include agreement on the type of food extracts used, which should ideally be standardised.

- Patients' allergy status should be verified if comparisons to placebo are to be made, otherwise there is no way of knowing if a response to an active extract is actually a placebo response. It should be made clear which patient is sensitive to which food and whether reactions subsequently occurred to the same food or to other foods.

- There should be agreement on how the patient's allergy status is to be verified, which could be problematic given the uncertainty surrounding the accuracy of the current gold standard or other tests, and the uncertainty surrounding the link between certain symptoms or diseases with food allergy or sensitivity.

- Outcome assessment should be as objective and reproducible as possible. If very subjective outcome measures are used, then ideally there should be two independent outcome assessors. If a standard outcome measure was used for different studies, this may allow results to be quantitatively combined. Longer-term, more clinically relevant outcome measures (for example relating to patient quality of life) should also be used.

- Patient groups should be well defined in terms of demographics, symptoms and allergy status.

- All results (for treatment and placebo groups) should be clearly reported.

- A power calculation should be performed to identify the sample size that would allow a difference in effect to be seen.

Other related topics of interest, which would help to clarify the overall evidence base surrounding diagnosis of food allergy include:

- The evidence base for possible immunological mechanisms that could account for the reactions induced by provocation-neutralisation.

- The evidence base for the association of food allergies with certain (chronic) diseases or symptoms (e.g. arthritis, psychological symptoms, multiple chemical sensitivity, chronic fatigue syndrome, total allergy syndrome). This is particularly important as some of these diseases are difficult to treat using conventional therapy and consume a lot of NHS resources.

- The evidence base for the reliability and accuracy of the current gold standard (double-blind, placebo-controlled food challenge) or other tests for food allergies and sensitivities.

9 Conclusion

There was no convincing evidence to suggest that more symptoms or wheals can be provoked with active extract compared to placebo by provocation-neutralisation and conclusions regarding the association of food allergy with certain symptoms or diseases cannot be drawn from the studies included in this review. No evidence was identified to suggest that provocation-neutralisation is useful for the diagnosis of food allergy.

There was some evidence to suggest that neutralisation may be effective in the treatment of food allergies, based on 61 patients from three studies. One of these studies was a preliminary report in eight patients only, and there are some queries around the outcome assessment in one further study. Some uncertainty around the initial diagnosis of food allergy in these studies also remains. It is also unclear whether the results are applicable to other patient groups, as the included populations were highly selected.

It should be noted that the absence of good evidence is not proof of ineffectiveness, and further well-designed studies are recommended for the assessment of the treatment aspect of this technique. In particular there should be clarity around the verification of allergy status and the presentation of all results.

Appendix 1Search Strategies

Search for diagnostic test accuracy studies (MEDLINE, EMBASE)

MEDLINE (OVID) (1966-present), search 27th September 2002

1. exp "Sensitivity and Specificity"/
2. sensitivit$.mp.
3. specificit$.mp.
4. predictive value$.mp.
5. likelihood ratio$.mp.
6. false positive$.mp.
7. true positive$.mp.
8. false negative$.mp.
9. true negative$.mp.
10. reference test.mp.
11. reference standard.mp.
12. reference value.mp.
13. gold standard.mp.
14. diagnostic accuracy.mp.
15. diagnostic performance.mp.
16. test performance.mp.
17. test accuracy.mp.
18. ROC.mp.
19. SROC.mp.
20. or/1-19
21. exp allergy/
22. exp allergens/
23. allerg$.mp.
24. sensitivit$.mp.
25. hypersensitiv$.mp.
26. exp hypersensitivity/
27. intolerance$.mp.
28. adverse reaction$.mp.
29. exp anaphylaxis/
30. anaphyla$.mp.
31. or/21-30
32. food$.mp.
33. exp food/
34. exp food additives/
35. or/32-34
36. 31 and 35
37. exp food hypersensitivity/
38. 36 or 37
39. provo$.mp.
40. neutrali$.mp.
41. exp injections, intradermal/
42. exp injections, subcutaneous/
43. intradermal$.mp.
44. intracutaneous$.mp.
45. subcutaneous$.mp.
46. sublingual$.mp.
47. sub-lingual$.mp.
48. exp administration, sublingual/

49. allergy vaccin$.mp.
50. exp immunotherapy/
51. immunotherap$.mp.
52. or/39-51
53. 20 and 38 and 52
54. limit 53 to human

EMBASE (OVID) (1980-present), search 27th September 2002

1. specificit$.mp.
2. sensitivit$.mp.
3. predictive value$.mp.
4. likelihood ratio$.mp.
5. false positive$.mp.
6. true positive$.mp.
7. false negative$.mp.
8. true negative$.mp.
9. reference test.mp.
10. reference standard.mp.
11. reference value.mp.
12. gold standard.mp.
13. diagnostic accuracy.mp.
14. diagnostic performance.mp.
15. test performance.mp.
16. test accuracy.mp.
17. ROC.mp.
18. SROC.mp.
19. or/1-18
20. exp allergy/
21. exp allergen/
22. allerg$.mp.
23. exp hypersensitivity/
24. hypersensitiv$.mp.
25. sensitivit$.mp.
26. intolerance$.mp.
27. adverse reaction$.mp.
28. exp anaphylaxis/
29. anaphyla$.mp.
30. or/20-29
31. exp food/
32. exp food additives/
33. food$.mp.
34. or/31-33
35. 30 and 34
36. exp food allergy/
37. exp food allergen/
38. exp food antigen/
39. or/36-38
40. 35 or 39
41. provo$.mp.
42. neutrali$.mp.
43. exp injection/
44. intradermal$.mp.
45. intracutaneous$.mp.

46. subcutaneous$.mp.
47. sublingual$.mp.
48. sub-lingual$.mp.
49. exp sublingual drug administration/
50. exp immunotherapy/
51. immunotherap$.mp.
52. allergy vaccin$.mp.
53. exp provocation test/
54. or/41-53
55. 19 and 40 and 54
56. limit 55 to human

Search for controlled trials of diagnosis/treatment strategies (MEDLINE, EMBASE, CCTR)

MEDLINE (OVID) (1966-present), search 30th September 2002

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized controlled trials/
4. random allocation/
5. double blind method/
6. single blind method/
7. or/1-6
8. (animal not human).sh.
9. 7 not 8
10. clinical trial.pt.
11. exp clinical trials/
12. (clin$ adj25 trial$).ti,ab.
13. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
14. placebos/
15. placebo$.ti,ab.
16. random$.ti,ab.
17. research design/
18. or/10-17
19. 18 not 8
20. 19 not 9
21. comparative study/
22. exp evaluation studies/
23. follow up studies/
24. prospective studies/
25. (control$ or prospectiv$ or volunteer$).ti,ab.
26. or/21-25
27. 26 not 8
28. 26 not (9 or 20)
29. 9 or 20 or 28
30. exp allergy/
31. exp allergens/
32. allerg$.mp.
33. sensitivit$.mp.
34. hypersensitiv$.mp.
35. exp hypersensitivity/
36. intolerance$.mp.
37. adverse reaction$.mp.

38. exp anaphylaxis/
39. anaphyla$.mp.
40. or/30-39
41. food$.mp.
42. exp food/
43. exp food additives/
44. or/41-43
45. 40 and 44
46. exp food hypersensitivity/
47. 45 or 46
48. provo$.mp.
49. neutrali$.mp.
50. exp injections, intradermal/
51. exp injections, subcutaneous/
52. intradermal$.mp.
53. intracutaneous$.mp.
54. subcutaneous$.mp.
55. sublingual$.mp.
56. sub-lingual$.mp.
57. exp administration, sublingual/
58. allergy vaccin$.mp.
59. exp immunotherapy/
60. immunotherap$.mp.
61. Miller.mp.
62. or/48-61
63. 29 and 47 and 62
64. limit 63 to human

EMBASE (OVID) (1980-present), search 3rd October 2002

1. randomized controlled trial/
2. exp clinical trial/
3. exp controlled study/
4. double blind procedure/
5. randomization/
6. placebo/
7. single blind procedure/
8. (control$ adj (trial$ or stud$ or evaluation$ or experiment$)).mp.
9. ((singl$ or doubl$ or trebl$ or tripl$) adj5 (blind$ or mask$)).mp.
10. (comparison group$ or control group$).mp.
11. (clinical trial$ or random$).mp.
12. (quasiexperimental or quasi experimental or pseudo experimental).mp.
13. exp double blind procedure/
14. exp crossover procedure/
15. or/1-14
16. exp allergy/
17. exp allergen/
18. allerg$.mp.
19. exp hypersensitivity/
20. hypersensitiv$.mp.
21. sensitivit$.mp.
22. intolerance$.mp.
23. adverse reaction$.mp.
24. exp anaphylaxis/

25. anaphyla$.mp.
26. or/16-25
27. food$.mp.
28. exp food/
29. exp food additives/
30. or/27-29
31. 26 and 30
32. exp food allergy/
33. exp food allergen/
34. exp food antigen/
35. or/32-34
36. 31 or 35
37. provo$.mp.
38. neutrali$.mp.
39. exp injection/
40. exp immunotherapy/
41. immunotherap$.mp.
42. intradermal$.mp.
43. intracutaneous$.mp.
44. subcutaneous$.mp.
45. sublingual$.mp.
46. sub-lingual$.mp.
47. exp sublingual drug administration/
48. allergy vaccin$.mp.
49. exp provocation test/
50. Miller.mp.
51. or/37-50
52. 15 and 36 and 51
53. limit 52 to human

Cochrane Controlled Trials Register, 2002, Issue 3 (search 3rd October 2002)

1 HYPERSENSITIVITY*:ME
2 ANAPHYLAXIS*:ME
3 ALLERG*
4 SENSITIVIT*
5 HYPERSENSITIV*
6 INTOLERANCE*
7 (ADVERSE and REACTION*)
8 ANAPHYLA*
9 (((((((#1 or #2) or #3) or #4) or #5) or #6) or #7) or #8)
10 FOOD*:ME
11 FOOD-ADDITIVES*:ME
12 FOOD*
13 ((#10 or #11) or #12)
14 FOOD-HYPERSENSITIVITY*:ME
15 (#9 and #13)
16 (#14 or #15)
17 PROVOC*
18 NEUTRALI*
19 INTRADERMAL*
20 SUBCUTANEOUS*
21 INTRACUTANEOUS*
22 INJECTIONS-INTRADERMAL*:ME

23 INJECTIONS-SUBCUTANEOUS*:ME
24 SUBLINGUAL*
25 SUB-LINGUAL*
26 ADMINISTRATION-SUBLINGUAL*:ME
27 IMMUNOTHERAPY*:ME
28 IMMUNOTHERAP*
29 (ALLERGY and VACCIN*)
30 MILLER
31 ((((((((((((#17 or #18) or #19) or #20) or #21) or #22) or #23) or #24) or #25) or #26) or #27) or #28) or #29) or
32 (#16 and #31)

CISCOM-Database of the Research Council for Complementary Medicine (www.rccm.org.uk)

On-line search form (www.rccm.co.uk/ciscom/CISCOMform.htm) submitted 10th October 2002. Searching carried out by organisation for a fee of £20. Search question as required by form: 'randomised controlled trials or controlled trials or diagnostic test accuracy studies regarding provocation-neutralisation testing (intradermal, subcutaneous or sublingual; also known as Miller technique) for the diagnosis and/or treatment of food allergies or sensitivities in any population with suspected food allergy compared to placebo and/or other tests strategies.'
Suggested search terms as required by form: 'food allergy/sensitivity/intolerance, provocation, neutralisation, immunotherapy, allergy vaccine, Miller technique, intradermal, subcutaneous, sublingual'.

Web sites/Ongoing trials

The following web sites were searched for completed or ongoing trials, and background information (by looking at any links to publications, research or references and searching for 'provocation', 'neutralisation', 'food allergy' or 'immunotherapy' where search facilities were available):

- Allergy UK (The British Allergy Foundation) (www.allergyfoundation.com), 28th October 2001
- The British Nutrition Foundation (www.nutrition.org.uk), 28th October 2002
- The British Society for Allergy & Clinical Immunology (www.bsaci.ston.ac.uk-web page under development 28th October)
- The American Academy of Allergy, Asthma and Immunology (www.aaaai.org), 28th October 2002
- The food allergy and anaphylaxis network (www.foodallergy.org), 28th October 2002
- www.allergytofood.co.uk (food allergy consultancy), 28th October 2002
- www.allergyaction.org (food allergy consultancy), 28th October 2002

Ongoing trials sites:

• The meta-register of Controlled Trials (UK based; 20 registers of ongoing/completed trials, including the National Research Register and the Medical Research Council; http://controlled-trials.com), 29th October 2002

• Trials Central (US and international meta-register; www.trialscentral.org), 29th October 2002-10-29

• http://clinicaltrials.gov (US based register), 29th October 2002

Contact with experts

Studies and additional information were provided by Dr HM Anthony, President of the British Society for Allergy, Nutritional and Environmental Medicine; Dr JA Monro, The Breakspear Hospital, Hertfordshire; Dr J R Mansfield, The Burghwood Clinic, Surrey; Dr M J Radcliffe, The Burghwood Clinic, Surrey.

Search of citation lists

The following reviews and primary studies were scanned for additional relevant studies:

Clinical ecology. American College of Physicians. [see comments.]. [Review] [91 refs]. *Annals of Internal Medicine* 1989;**111**:168-78.

Dixon HS. Dysphonia and delayed food allergy: a provocation/neutralization study with strobovideolaryngoscopy. *Otolaryngology - Head & Neck Surgery* 1999;**121**:418-29

Forman R. A critique of evaluation studies of sublingual and intracutaneous provocative tests for food allergy. *Medical Hypotheses* 1981;**7**:1019-27.

Fox RA, Sabo BMT, Williams TPW, Joffres MR. Intradermal testing for food and chemical sensitivities: A double-blind controlled study. *Journal of Allergy & Clinical Immunology* 1999;**103**:907-11.
Gerdes KA. Provocation/Neutralization testing: A look at the controversy. *Clinical Ecology* 1989;**6**:21-3.

Goldberg BJ,.Kaplan MS. Controversial concepts and techniques in the diagnosis and management of food allergies. *Immunology & Allergy Clinics of North America* 1991;**11**:863-84.

Grieco MH. Controversial practices in allergy. *JAMA* 1982;**247**:3106-11.

Jewett DL, Fein G, Greenberg MH. A double-blind study of symptom provocation to determine food sensitivity. *New England Journal of Medicine* 1990;**323**:429-33.

King DS. The reliability and validity of provocative food testing: a critical review. [Review] [56 refs]. *Medical Hypotheses* 1988;**25**:7-16.

King WP, Fadal RG, Ward WA, Trevino RJ, Pierce WB, Stewart JA *et al*. Provocation-neutralization: A two-part study. Part II. Subcutaneous neutralization therapy: A multi-center study. *Otolaryngology - Head & Neck Surgery* 1988;**99**:272-7.

King WP, Rubin WA, Fadal RG, Ward WA, Trevino RJ, Pierce WB *et al*. Provocation-neutralization: A two-part study. Part I. The intracutaneous provocative food test: A multi-center comparison study. *Otolaryngology - Head & Neck Surgery* 1988;**99**:263-71.

Lehman CW. A double-blind study of sublingual provocative food testing: a study of its efficacy. *Annals of Allergy* 1980;**45**:144-9.

Miller JB. A double-blind study of food extract injection therapy: a preliminary report. *Annals of Allergy* 1977;**38**:185-91.

Monro J. Provocation tests for food allergy. [letter; comment]. *Lancet* 1991;**338**:1204.

Podell RN. Food extract injection for food sensitivity. Valid technique or 'black magic'? *Postgraduate Medicine* 1984;**76**:59-62.

Ranheim P. Provocation-neutralization in the treatment of food allergy. [letter; comment]. *American Family Physician* 402;**60**:392.

Rapp DJ. Food allergy treatment for hyperkinesis. *Journal of Learning Disabilities* 1979;**12**:608-16.

Rapp DJ. Diagnostic testing and immunotherapy for allergy. *JAMA* 1988;**260**:341-2.

Rea WJ, Podell RN, Williams ML. Elimination of oral food challenge reaction by injection of food extracts. A double-blind evaluation. *Archives of Otolaryngology* 1984;**110**:248-52.

Van MT, Jr. Unproven procedures for diagnosis and treatment of food allergy. [Review] [52 refs]. *New England & Regional Allergy Proceedings* 1987;**8**:17-21.

Appendix 2 Cost effectiveness search strategy

The following databases were searched: MEDLINE (Ovid, 1966-8th July 2003) and EMBASE (Ovid, 1980-8th July 2003) using search filters designed to identify cost related studies combined with relevant text and MeSH words, the Cochrane Library, Issue 2, 2003 (DARE, Database of Abstracts of Reviews of Effectiveness and NHS EED, the NHS Economic Evaluation Database) and the OHE Office of Health Economics Database (OHE HEED, updated July 2003). There were no language restrictions. Full details of the search strategies are shown below.

MEDLINE (OVID) (1966- 8th July 2003)

1. exp allergy/
2. exp allergens/
3. allerg$.mp.
4. sensitivit$.mp.
5. hypersensitiv$.mp.
6. exp hypersensitivity/
7. intolerance$.mp.
8. adverse reaction$.mp.
9. exp anaphylaxis/
10. anaphyla$.mp.
11. exp food hypersensitivity/
12. provo$.mp.
13. neutrali$.mp.
14. exp injections, intradermal/
15. exp injections, subcutaneous/
16. intradermal$.mp.
17. intracutaneous$.mp.
18. subcutaneous$.mp.
19. sublingual$.mp.
20. sub-lingual$.mp.
21. exp administration, sublingual/
22. allergy vaccin$.mp.
23. exp immunotherapy/
24. immunotherap$.mp.
25. Miller.mp.
26. or/12-25
27. or/1-11
28. 26 and 27
29. economics/
30. exp "costs and cost analysis"/
31. cost of illness/
32. exp health care costs/
33. economic value of life/
34. exp economics medical/
35. exp economics hospital/
36. economics pharmaceutical/
37. exp "fees and charges"/
38. (econom$ or cost or costs or costly or costing or price or pricing or

pharmacoeconomic$).tw.
39. (expenditure$ not energy).tw.
40. (value adj1 money).tw.
41. budget$.tw.
42. or/29-41
43. 28 and 42
44. limit 43 to human

EMBASE (OVID) (1980- 8th July 2003)
1. exp allergy/
2. exp allergen/
3. allerg$.mp.
4. exp hypersensitivity/
5. hypersensitiv$.mp.
6. sensitivit$.mp.
7. intolerance$.mp.
8. adverse reaction$.mp.
9. exp anaphylaxis/
10. anaphyla$.mp.
11. exp food allergy/
12. exp food allergen/
13. exp food antigen/
14. provo$.mp.
15. neutrali$.mp.
16. exp injection/
17. exp immunotherapy/
18. immunotherap$.mp.
19. intradermal$.mp.
20. intracutaneous$.mp.
21. subcutaneous$.mp.
22. sublingual$.mp.
23. sub-lingual$.mp.
24. exp sublingual drug administration/
25. allergy vaccin$.mp.
26. exp provocation test/
27. Miller.mp.
28. or/14-27
29. or/1-13
30. 28 and 29
31. cost benefit analysis/
32. cost effectiveness analysis/
33. cost minimization analysis/
34. cost utility analysis/
35. economic evaluation/
36. (cost or costs or costed or costly or costing).tw.

37. (economic$ or pharmacoeconomic$ or price or pricing).tw.
38. (technology adj assessment).tw.
39. or/31-38
40. 30 and 39
41. limit 40 to human

The Cochrane Library, Issue 2, 2003: DARE and NHS EED

The two databases were searched using each search term individually.

Provo*
Neutrali*
Clinical ecology*
Allerg*
Immunotherap*
Hypersensitivity*
Hypersensitivity*;ME
Environmental illness*;ME

OHE EED (July 2003 update)

The database was searched using each term individually.

Provo*
Neutrali*
Immunotherapy
Food allergy
Food sensitivity
Food hypersensitivity

Appendix 3 Excluded studies

Study	Reason(s) for exclusion
Provocation testing and food sensitivity, 1991[70]	Comment on included study
Bronsky *et al*., 1971[67]	Abstract only (no further studies by the same author identified)
Crawford *et al*., 1976[68]	Abstract only (no further studies by the same author identified)
Dixon, 1999[71]	Study is not placebo-controlled in allergic subjects
Draper, 1972[72]	No RCT of provocation testing; comparison of different test methods but not possible to calculate sensitivity or specificity
Duncan *et al*., 1977[73]	Uncontrolled study
Endicott & Stucker, 1977[74]	Preliminary report; uncontrolled provocative food testing; no later reports identified
Forman, 1981[75]	Review
Green, 1974[76]	Provocation testing is uncontrolled, no comparison of tests
Jewett & Greenberg, 1985[77]	Abstract only -appears to report same study as publication from 1990 included in this report
Kailin & Collier, 1971[69]	Letter only, insufficient information (no further studies by the same author identified)
King, 1988[64]	Review
King, 1978[78]	Not possible to calculate sensitivity or specificity
Mandell & Conte, 1980[79]	Abstract only -appears to report same study as publication from 1982 included in this report
McGovern, 1981[80]	Comment
Podell, 1983[81]	Comment
Podell, 1984[47]	Review
Rapp, 1978[82]	Case-study
Rapp, 1981[83]	Comment
Rapp, 1988[84]	Comment
Shaver, 1975[85]	Uncontrolled study
Teuber & Vogt, 1999[66]	Case study
Van Metre, 1987[1]	Review

The following studies were unobtainable and could therefore not be assessed: Rapp, 1982[86] and Heyse, 1981[87].

Appendix 4 Patient characteristics

Study	Inclusion and exclusion criteria for participants	Age, sex	Allergy history (suspected, confirmed allergies, previous tests/ treatment etc)	Current symptoms/duration of symptoms	Co-morbidity/ medication
Provocation-neutralisation testing					
Breneman *et al*., 1973[48]	Patients suspected of having food allergy and a negative sublingual response to placebo	Not stated	Not stated	Not stated	Not stated
Breneman *et al*., 1974[52]	Patients with a known allergy to at least one of the foods tested and a negative sublingual response to placebo	Not stated	See inclusion criteria	Not stated	Not stated
Caplin, 1973[45]	Protocol for patient selection: Inclusion: 10 patients meeting the following criteria (primary group): atopic patients, who have positive intradermal skin test to inhalants correlating with an allergic history and symptomatology; without concomitant chronic conditions; no steroid therapy for at least 3 months, no regular medication; no prior knowledge of food sensitivity to the test foods 5 additional patients who have clinical sensitivity to one or more foods being tested and who have had symptoms to the food during the preceding week (secondary group) 5 additional new patients who have been classified by history, physical examination and skin tests as non-allergic Exclusion: patients with known fixed severe reactions to any food on each ingestion of the food	Not stated	See inclusion criteria	Atopic patients with perennial primary symptoms of asthma, rhinitis and /or eczema and other allergic manifestations	See inclusion criteria
Fox *et al*., 1999[56]	Only those patients thought to have chemical sensitivity after initial consultation were included, with symptoms severe enough to justify early consultation	103 female, 29 male; 9-78 years	Possible or probable chemical sensitivity	Patients with symptoms that were disruptive to normal life, unable to work or attend school, suicidal or experiencing life-threatening anaphylaxis	Not stated

Study	Inclusion and exclusion criteria for participants	Age, sex	Allergy history (suspected, confirmed allergies, previous tests/ treatment etc)	Current symptoms/duration of symptoms	Co-morbidity/ medication
Jewett *et al*., 1990[54]	No history of anaphylactic or anaphylactoid reactions, fainting, cardiac irregularities, severe laryngeal oedema, severe asthma, epileptic or epileptoid seizures; recommendation for study by physician on basis of previous active responses to intradermal or subcutaneous injections of 'active' substances and no response to unblended injections of diluent alone	15 female, 3 male; 18-60 years	Symptoms consistently provoked during previous unblinded testing	Not stated	Not stated
King *et al*., 1988, (part I)[46]	Included: patients aged between 5-50 years, not pregnant or nursing, not severely reactive (as provocation testing was begun with a maximum provoking dose), never have experienced previous provocation-neutralisation testing or any form of food allergy treatment, have read and signed a standardised release form before participation, respond positively to at least one of the foods when orally challenged	13 female, 24 male; 5-50 years	Patients should respond positively to at least one of the foods being tested when orally challenged	Mainly rhinitis, headache, gastro-intestinal or bronchopulmonary symptoms and fatigue, also urticaria, throat problems, dizziness, ear discomfort, skin itch/rash, oedema, conjunctiva, palpitations, arthralgia, hot flash, mood change	Patients should not have used 2 weeks before study or during: antihistamines, cromolyn sodium, nasal sprays and aerosols, oral decongestants, tranquillisers, antidepressants; should not have used for 4 weeks before study or during: hydroxyzine, steroids, theophylines, beta blockers, beta agonists, nonsteroidal anti-inflammatory medications

Study	Inclusion and exclusion criteria for participants	Age, sex	Allergy history (suspected, confirmed allergies, previous tests/ treatment etc)	Current symptoms/duration of symptoms	Co-morbidity/ medication
King, 1981[58]	Inclusion criterion: the presence of at least one psychological symptom (e.g. anxiety, depression, confusion, difficulty in concentrating)	20 female, 10 male range 17-56 years,	Not stated	Not stated	Not stated
Lehman, 1980[57]	Inclusion criteria; history of symptoms occurring repeatedly after eating a given food or symptoms improving during an elimination period and recurring when reintroduced or history of allergic symptoms during infancy/early childhood	Mean 13.7 (range 3-31; year of birth only stated)	History of food allergy; food allergy established clinically in 12/15; 10/15 allergic to one of the foods tested (9 milk, 1 corn); symptoms after eating a given food or symptoms improving during a 4-5 day elimination period and recurring when test food reintroduced	Allergic rhinitis, asthma, otitis media, allergic bronchitis, allergic gastroenteritis, atopic eczema; varying durations (1-18 years)	Not stated
Mandell & Conte, 1982[55]	Volunteers with arthritis or rheumatism	23 female, 7 male, mean 57 (range 35-79),	Not stated	Not stated	Not stated
Neutralisation Therapy					
King *et al*., 1988, (part II)[53]	As part I; additionally, patients should not be on any injection treatment other than for inhalant allergy	33 of the 37 patients from part I	As part I	As part I	As part I; necessary symptom relieving medications
Miller, 1977[41]	Patients who presented with symptoms and syndromes judged to be caused by food sensitivities and who responded well to food injection therapy	5 female, 3 male, mean 30.9 (range 4-57)	History of food sensitivity as detailed in case descriptions in study; symptoms and syndromes judged to be food allergy; use of single-blind tests	Migraine, hyperactivity, nausea, vomiting, diarrhoea, abdominal cramps, nasal congestion,	No patients receiving inhalant immunotherapy

Study	Inclusion and exclusion criteria for participants	Age, sex	Allergy history (suspected, confirmed allergies, previous tests/ treatment etc)	Current symptoms/duration of symptoms	Co-morbidity/ medication
				allergic rhinitis and others	
O'Shea & Porter, 1981[59]	Children who met the clinical criteria of hyperkinetic syndrome (mean score on the Abbot's Hyperkinetic Index of 20), includes the following characteristics: hyperactivity, disruptiveness, compulsiveness, low frustration tolerance, short attention pan, unable to sustain play or work projects, academic difficulties and who respond to a trial of methylphenidate hydrochloride (Ritalin); children were excluded who exhibited hyperkinetic syndrome-like behaviour primarily associated with psychosis, overanxious reactions, organic brain damage and mental retardation	Mean 7.7 years (range 5-13), sex not stated	7/15 positive family history of allergy (no details on type of allergy or how verified); 10/15 personal history of allergy (no details on type of allergy or how verified); 11/15 positive physical signs of respiratory allergy	Hyperkinetic syndrome) see inclusion criteria)	Not stated
Rapp, 1979[40]	Hyperactive patients who responded favourable to dietary management (food omission; preliminary study)	4 female; 7 male; range 6-15 years	10/11 patients positive allergy history; previous treatment with food omission diet (individualized diets omitting frequently ingested foods)	Original symptoms stated for 8/11 patients who completed study: behavioural problems, depression, eye circles, gastrointestinal symptoms, urinary complaints, headaches, hyperactivity, muscle aches, nose or chest symptoms, skin eczema	Not stated

Study	Inclusion and exclusion criteria for participants	Age, sex	Allergy history (suspected, confirmed allergies, previous tests/ treatment etc)	Current symptoms/duration of symptoms	Co-morbidity/ medication
Rea *et al*., 1984[60]	Inclusion criteria: not pregnant; not menstruating; be eating foods to be tested at least three times weekly and exhibit within one hour one of the following symptoms: asthma, urticaria, oedema, gastrointestinal upset (diarrhoea), eczema, rhinitis, spastic vascular phenomena, bone or muscle tenderness or swelling; be receiving no medication or food injection therapy for two weeks; exhibit positive whealing to histamine and negative whealing to saline; be recommended by the testing technician as appearing capable of obtaining accurate neutralising doses; willing and able to provide informed consent	13 female, 7male, mean 36 +/- 3 years (range 20-69)	Suspected food and/or inhalant allergies; symptoms at least 1 hour after oral food challenge	Admitting chief complaints: eczema, diarrhoea, myalgia, arthralgia, rhinitis, fatigue, optic neuritis, vasculitis, dermatitis, arthritis, headache, mental dysfunction, asthma, irritable bowel, psoriasis, Sjorgen's syndrome, myositis, seizure disorder	No medication (was inclusion criterion); no co-morbidity stated

Appendix 5 Test Protocols

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
Provocation-Neutralisation Testing		
Breneman *et al*., 1973[48]	Allergens: milk, wheat, egg, corn, chocolate Placebo: normal saline-glycerine	PROVOCATION TESTING: All foods in 1:40 extracts of food in glycerine vehicle; each patient, after a negative control test, started on sequence of unknowns; two drops of extract placed under tongue and patients observed for 20 minutes; reactions were recorded on a standard provided form; if there was a reaction, testing was stopped for 24 hours; neutralisation could be attempted
Breneman *et al*., 1974[52]		PROVOCATION TESTING: All foods in 1:10 extracts of food in glycerine vehicle; each patient, after a negative control test, started on sequence of unknowns; two drops of extract placed under tongue and patients observed for 20 minutes; if there was no reaction, the investigator could proceed to the next test substance; reactions were recorded on a standard provided form; neutralisation could be attempted
Caplin, 1973[45]	Allergens: milk, wheat, egg, corn, chocolate Placebo: normal saline (phenol as preservative)	PROVOCATION TESTING: Food allergenic extracts in a 1:100 solution; 0.1cc of unknown allergenic extract injected subcutaneously in the lateral area of the arm, observed for 20 minutes; if no changes occur in 20 minutes, another unknown food may be tested; an effort should be made to test as many substances in one day as possible; if delayed reactions occur, then a retesting of the individual foods on the consecutive day is completed; 840 tests performed in 70 patients (12 each) ORAL FOOD CHALLENGE: The patient should be on a diet avoiding all test foods for at least 4 days; the patient is observed for a preliminary period of 30 minutes, then a test meal is administered (milkshake containing canned corn kernels, egg, chocolate syrup , flour, milk, ice cubes); 48 patients had feeding tests
Fox *et al*., 1999[56]	Allergens: banana, beef, baker's yeast, brewer's yeast, cane sugar, chicken, chocolate, corn, cow's milk, egg, soy, tomato, wheat; cigarette smoke extract, ethanol, formaldehyde, orris root, phenol, grass terpene, pine terpene, unleaded/diesel and either fireplace ash or women's/men's cologne or cedar terpene Placebo: saline solution	PROVOCATION TESTING: Each allergen set up as series of 1 in 5 dilutions (1 in 5 to 1 in 3125); patients tested with single allergen at a time, response assessed over 10 minutes; testing over several days, average 2.5 days to complete full panel of allergens; positive response if growth of size of wheal by 2mm; endpoint recorded as the dilution for the first negative wheal; if the patient developed symptoms they were recorded; incremental doses of 0.05 ml injected until symptoms disappeared (neutralising dose); the neutralisation point was recorded as the total volume injected at the final dilution when symptoms disappeared; delayed responses not recorded as positive; panels changed 5 times during 17 month study period (Miller technique)

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
Jewett *et al*., 1990[54]	1 allergen chosen for each patient by their physician on basis of previous unblended tests: chocolate, wheat, baker's yeast, potato, ethanol, brewer's yeast, apple, milk, corn, beef, mould A, orange or chicken Placebo: saline solution	PROVOCATION TESTING: Concentration and method of injection chosen by physician; 17 given 'underdoses' relative to neutralising dose, 3 'overdoses'; (neutralising dose defined as dose than when injected in a volume of 0.1 ml resulted in a wheal of 7-8mm which enlarged by 2mm in 10 minutes); injections given intradermally in 16 patients, subcutaneously in 4; 12 injections per patients (3 active and 9 placebo), with observation period of 10 minutes; 240 tests in total (60 active, 180 placebo); only 1 dilution per patient used
King *et al*., 1988 (part I)[46]	Allergens: wheat, corn, beef, white potato, milk Placebo: 50% glycerin	PROVOCATION TESTING: Patients were instructed to consume the five foods to be tested the day before, as such priming is thought to elicit clearer skin responses; five-fold dilutions of a 1:10 food extract concentrate were used; five-fold dilution of 50% glycerin were used as a control; testing was begun by placing a maximal provoking 0.05 cc intracutaneous dose of the #1 dilution of the related food extract on the upper arm; if no provocation of symptoms and no positive whealing of 2mm or larger than the #1 dilution of the control occurred within 20 minutes, test was recorded as negative; if the whealing response was positive without symptom provocation within 20 minutes, an application of identical sized wheals using consecutively weaker dilutions were applied every 10 minutes until the strongest dilution that produced a negative wheal was identified (final endpoint dose); if the whealing response was positive and symptoms were provoked within the 20 minutes, the tester proceeded to neutralise the provoked symptoms by applying identically sized wheals via weaker dilutions at 10 minute intervals 3 identical trials performed at 7 day intervals; 185 comparisons in total (Miller technique) COMPARISON ORAL FOOD CHALLENGE AND PROVOCATION TESTING: 5 day open oral challenge food test-patients instructed to eat usual diet; subsequent double-blind intracutaneous provocative food tests of five test foods, 3 times at 7 day intervals; if positive whealing response resulted without symptom provocation, the testing for this allergen was stopped without identifying the strongest dilution that produced a negative wheal (in order to ensure that the second provocation tests are not influenced by the previous delivery of a neutralising dose); when symptoms were provoked on the second or third provocation test, appropriate neutralisation was accomplished and recorded; when positive whealing occurred without symptom provocation, the strongest dilution that produced a negative wheal identified as the final endpoint, thereby any effect on the third provocation test responses of providing a neutralising or endpoint dose 7 days earlier could be observed

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
King, 1981[58]	Food allergens: wheat, beef, milk, cane sugar, egg, chocolate, potato, lettuce, tea, apple, orange, peanut, cheddar cheese, corn, rice, tomato, yeast (baker's), yeast (brewer's), almond, banana, carrot, chicken, cucumber, grape, pork, rye, sesame, turkey Other allergens: tobacco smoke, ethanol (petrochemical), auto exhaust, chlorine, altenaria) Placebo: triple distilled water 4 allergens tested based on intake information	PROVOCATION TESTING: Standard allergenic extracts administered sublingually by disposable syringes (dilutions of 0.2 cc of 1:5, 0.1 cc of 1:125 and 0.2cc of undiluted); (1) 4 substances (allergen) each in 3 doses, (2) 2 placebo tests, each in 3 doses, (3) 3 base rate trials, (4) 3 screening trials to assess biological reactivity to placebo; tests 10 minutes duration; order: 2 screening trials, 2 base rate, 6 randomised allergen/placebo, final base rate and screening trial 540 allergen (360)/placebo (180) trials in total (30 patients x 6 trials x 3 doses) Data collection: patient lists any symptoms present, expectation marked, 1 minute later 3 pulse readings were taken, followed by administration of the dose, guess of the substance's identity and certainty of guess recorded, 3 pulse readings followed by other measures
Lehman, 1980[57]	Frequently ingested foods, which commonly produce food sensitivity Allergens: egg, corn, milk, yeast Placebo: distilled water (phenolated saline in first 2 patients, subsequently changed)	PROVOCATION TESTING: Testing with sublingual food drops or placebo; details described elsewhere (Morris DL. Use of sublingual antigen in diagnosis and treatment of food allergy. Ann Allergy 1969. 27:289-294) 9 patients tested in 1972, 1 in 1973, 5 in 1978; baseline testing, same tests repeated 1 month later; appear to be 5 tests (1 placebo, 4 food) per patient at baseline and after 1 month
Mandell & Conte, 1982[55]	Allergens: altenaria, apple, auto exhaust, beef, chicken, chocolate, coffee, corn, egg, ethanol, house dust, lettuce, milk, natural gas, orange, pork, potato, peanut, soy, sugar, tea, tobacco, tomato, wheat, yeast (baker's) Placebo: sterile distilled water	PROVOCATION TESTING: 25 x 3-dose sets of sublingual challenges were administered to each subject; provocative tests conducted at 10-15 minute intervals; the first dose was 0.2cc of a 1:100 solution, followed by 0.1 cc of a 1:2500 solution; the third dose was 0.2cc of either a 1;10 or 1:20 extract; 7 or 8 3-dose sublingual challenges were performed during each of four 3.4-4-hour provocation testing sessions; total number of allergens tested: 2250, total number of placebos tested: 360. (method according to Rinkel, HJ, Randolph, TG and Zeller, M. *Food Allergy*. Charles Thomas, Publishers, Springfield, 1950; Rinkel, HJ. Food Allergy: II. The Technique and Clinical application of Individual Food Tests. *Ann*

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
		Allergy, 2:504, 196)
Neutralisation Therapy		
King *et al*., 1988 (part II)[53]	Allergens: wheat, corn, beef, white potato, milk; additional foods used for treatment at discretion of physician Placebo: 50% glycerine	NEUTRALISTION THERAPY: Treatment over 8 weeks (3 2-week sessions with 1 week gaps, 1 2-week session is placebo); total number of foods used for treatment between 1 and 13, majority of patients treated with 3-5 foods; each injection of the active treatment material contained the exact amount of allergen present in the neutralising dose or the endpoint final dose; injections were administered once a day; self-administration from the provided 20-dose multiple treatment vial was allowed after the first injection was given in the office under direct observation to assure compatibility between the vial and the patient
Miller, 1977[41]	Wide range of foods in diet normally consumed and those foods that cause symptoms Placebo: phenolated saline coloured with mushroom extract to similar colour as extracts (no patients in study significantly sensitive to mushroom) Neutralisation: foods not specified (combination of those foods causing symptoms) in a single dose	PROVOCATION TESTING: Intradermal injection of 0.05 ml of one dilution of a specific extract, successive intradermal injections of 0.05ml of different dilutions of the same food extract, usually at 10 minute intervals, until the dilution is found which completely relieves the symptoms (neutralising dilution) NEUTRALISATION THERAPY: Treatment solution injected daily for duration of study (self-administration by patient) 4x 20 days (2 periods of active treatment, 2 of placebo) 0.05ml of treatment dilution combining all neutralising doses
O'Shea & Porter, 1981[59]	Allergens (in phenolated saline): food: milk, peanuts, tomato, apple, cane sugar, corn, grape, orange, chocolate, wheat, egg; dye: red dye, yellow dye, blue dye; inhalants: dust, mould, tree Placebo: phenolated saline	PROVOCATION TESTING: Intradermal testing: positive reaction was a wheal growth of 2mm or greater in 10 minutes after whealing 0.05 ml of a No. 1 dilution (1:100) of the food or pollen tested with or without symptom responses; if the test was positive a neutralising dose was obtained by Miller's technique Sublingual testing: dyes were tested sublingually starting with 0.05 ml (2 drops) of a 1:5 dilution; positive test determined by behavioural symptom response and a neutralising dose was obtained by an improved behaviour response; neutralising doses were achieved by giving the patient 0.05 ml (2 drops) sublingually of succeedingly weaker strengths of the dye which were serially diluted in multiples of five until normal behaviour response was obtained

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
		NEUTRALISATION THERAPY: Sublingual treatment with multiallergen extract at neutralising dose (number/doses of treatment not stated) 7 weeks (3 weeks of allergen extract and 3 weeks of placebo, with 1 week vacation in between)
Rapp, 1979[40]	Frequently ingested foods, which commonly produce food sensitivity Allergens include: artificial food colourings, sugar, wheat, milk, cocoa, corn, egg	PROVOCATION TESTING: Neutralising dose determined according to Miller method; food antigen (1:20 W/V) is diluted by successive negative powers of five; intracutaneous wheals from progressively weaker dilutions are applied at 10 minute intervals; neutralising dose is defined as most concentrated dose where there is a negative wheal and signs and symptoms associated with the oral food challenge are substantially relieved within 10 minutes NEUTRALISATION THERAPY: Subcutaneous (once daily) or sublingual (3 times daily) administration of treatment dilution (neutralising dose) Coded solutions (A, B, C) were given by parent for duration of trial; artificial food colouring avoided during this time Neutralisation treatment over 1-3 months; coded solutions trial 5-7 days Sublingual: 0.1 ml in 3 drops; subcutaneous: 0.1 ml per injection; individual; dilutions not stated For coded trial: 3 opaque bottles, 2 with diluent (buffered saline), 1 with food dilution
Rea *et al*., 1984[60]	Foods selected on basis of clinical history of sensitivity (n=9) or those to which the patient was believed to be sensitive based on a previous oral food challenge Allergens: chicken, wheat, potato, white potato, beef, corn, milk, egg, lamb; food extracted in modified Coca's solution and suspended in saline (no phenol or glycerin preservatives) Placebo: sterile saline	PROVOCATION TESTING: Neutralising dose determined according to Miller method; food antigen (1:20 W/V) is diluted by successive negative powers of five; intracutaneous wheals from progressively weaker dilutions are applied at 10 minute intervals; neutralising dose is defined as most concentrated dose where there is a negative wheal and signs and symptoms associated with the oral food challenge are substantially relieved within 10 minutes NEUTRALISATION THERAPY: 2-4 days after OFC (where neutralising dose was established), patient received subcutaneous injection with ND or placebo; OFC repeated 30 minutes later; observations made 30 and 15 minutes prior to the injection and 15 and 30 minutes after the injection; process repeated after 2-4 days and again after 2-4 days; within a day of the last injection, observational assessment is completed 3 trials run per patient, 60 trials in total for each outcome; 12-18 days from admission to completion of data gathering

Study (1st author, year, country)	Allergens & placebo	Test protocol (include details on dilutions, no of tests, duration of testing)
		3 syringes for each patient (1 with ND, 2 with placebo); Each patient re-challenged with the same food that provoked symptoms during the initial OFC.

Appendix 6 Study quality RCTs

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
Provocation-Neutralisation Testing						
Breneman *et al*., 1973[48]	The authors state that the testing sequence should have been randomised (each patient started with same set of materials and progressed in the same sequence); not clear if sequence concealed	Code kept by laboratories until testing was complete; 4/5 foods tested were not identifiable, however the authors state that chocolate was identifiable probably due to colour, odour and flavour	Code was broken after 78/100 patients had been tested; 61 records suitable for computer analysis; not clear why testing not completed or why records not suitable	The authors state that patients should not have been shown a suggested list of symptoms they might experience	Clinicians chosen by Food Allergy Committee on the sublingual method of provocative testing for food allergy	No details on verification of allergy status of patients; method of testing does not correspond to Miller technique
Breneman *et al*., 1974[52]	No details on randomisation or concealment	Code kept by laboratories until testing was complete; no details on inability to distinguish allergen/placebo	20/30 patients considered satisfactory for computer analysis; not clear why other records not suitable	No details on method of outcome assessment		It was stated that a feeding challenge was to be conducted to confirm the presence of food allergy, there were however no details of this; method of testing does not correspond to Miller technique
Caplin, 1973[45]	Authors state that it makes no difference whether extracts are tested in alphabetical order or randomly; sequence of vials unknown to investigator	Neither patient nor investigator knew contents of vials; to minimise identification of placebo by the physician, the technician and nurse should draw the extract for testing and record the	No loss to follow up	Attempts made to ensure patients were not biased during outcome reporting (e.g. not suggesting symptoms to the patient, or informing patients that symptom is anticipated)	Financial support from the Women's Auxiliary of the College of Allergists	Method of testing does not correspond to Miller technique; s*ee Appendix 7 for quality assessment of diagnostic test accuracy part of study*

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
		physicians observations				
Fox *et al*., 1999[56]	Study described as randomised, no details on randomisation or concealment of test order	Codes of solutions were not known by patients or the testing nurse; appears that outcomes were assessed by testing nurse; code broke for physician when testing was complete; not clear if analysis was performed blindly; likely that placebo was indistinguishable from allergen (vials identified only by letter and number)	No loss to follow-up	Lack of detail on how symptoms were defined or assessed; wheal measure was more objective; data presentation mainly graphical, not very easy to determine total reactions in the form of a wheal or symptom only to a given food	Nova Scotia Environmental Health Centre, Fall River	No details on verification of allergy status of patients; Miller technique used
Jewett *et al*., 1990[54]	Order of tests determined by combined die and coin toss; order of tests kept concealed	Patients, technicians and observer blinded; likely that placebo was indistinguishable from allergen as syringes were unmarked	No loss to follow-up, 2 patients tested twice	Not stated clearly how often/which symptoms occurred with allergen and placebo solutions; details only on ability to identify allergen or placebo	Offices of seven clinical ecologists in private practice, who were proponents of the technique with 5 years experience	Method of testing does not correspond to Miller technique

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
King *et al*., 1988 (part I)[46]	The coder randomly changed the order of allergens and a placebo; order of testing was also randomised	A coder, tester and a collector used to maintain double-blinding; codes only known by coder until completion of study; all food extracts were clear and presented no sight recognition problems	No loss to follow-up	Results for placebo not stated	Offices of eight different physicians	Miller technique used; *See Appendix 7 for quality assessment of diagnostic test accuracy part of study*
King, 1981[58] Provocation testing	Study described as randomised, no details on randomisation or concealment of test order	Patients not informed that placebos were in use; experimenters administering solutions closed their eyes when placing solutions under patients' tongues; outcome assessors blinded; syringe containing solutions masked with opaque paper	Placebo aware or experimenter suspicious or aware trials were removed from analysis; not always clear which data included in analyses; no intention-to-treat analysis	It was not always clear which data was included in the analyses; only measures that showed an effect were reported in detail Judges independently scored reported symptoms into categories (88% agreement) or as old/new symptoms (86% agreement)	Alan Mandell Centre for Bio-Ecologic Disease, Norwalk, Connecticut; research conducted at clinic of Dr Marshall Mandell (clinical ecologist)	No details on verification of allergy status of patients; sublingual method
Lehman, 1980[57]	No details on randomisation or concealment	No details on blinding; food drops and placebo looked alike (placebo was changed after testing the two initial patients to	2/15 patients did not return for 2nd test; 5 additional patients tested 6 years after other	Not clear if outcome measure validated; author stated that measure is difficult to interpret; environmental	Department of Allergy & Clinical Immunlogy, Straub Clinic and Hospital, Hawaii	Sublingual testing

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
		eliminate the possibility of a reaction to the phenol)	patients and added to analysis to determine whether the difference in reactions to corn/placebo were significant	inhalants were not controlled for		
Mandell & Conte, 1982[55]	Each vial identified by a code letter hat had been assigned by random selection from the alphabet by an individual not participating in the investigation	Patients and technicians blinded; patients and technicians not informed that some of the solutions were placebos; no further details on placebo	No loss to follow-up	Self-reported outcome measures; no details on assessment methodology	Not stated; appears to be author's own clinic	No details on verification of allergy status of patients; sublingual method
Neutralisation Therapy						
King *et al*., 1988 (part II)[53]	Order of treatment sessions determined by lot by an out-of-office individual; at the beginning of each treatment session, the same individual was provided with a vial of active material and placebo to provide one to the patient	Patients were blinded; tests were performed to verify that the placebo and active solution appeared identical to the patient (patients who could distinguish would not have been eligible for the study, but this did not occur); symptom response diaries were forwarded to a central collector	There appeared to be no loss to follow-up	Symptoms scored 1 (much worse) to 6 (excellent relief)	Offices of seven different physicians (of the original eight, see King *et al*., 1988, part I)	Miller technique used

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
Miller, 1977[41]	First extract chosen was determined by coin flip by a third party	For neutralisation treatment: neither patient not physician were aware of the extract; placebo was coloured to prevent recognition (the authors state that mushroom extract, one of the darkest coloured extracts available, was used)	No loss to follow-up	Symptom improvement graded 0-+4 (in terms of intensity, duration and frequency)	Not stated	Miller technique was used; preliminary report of 8 patients, no follow-up study identified
O'Shea & Porter, 1981[59]	Study described as random, but no details on randomisation or concealment	All treatment solutions coded by chief pharmacist; solutions identical in colour and taste; code not broken until study was completed	1/15 children lost to follow-up	Outcome assessment based on diaries kept by parents and interviews with a psychologist	Not stated	1 week washout period between treatment and placebo; no details on verification of allergy status of patients; Miller technique and sublingual
Rapp, 1979[40]	It was stated that each patient's treatment set was randomly coded in a double-blind manner	Stated that active extracts and placebo were identical in colour and taste	3/11 children did not complete study; excluded from analysis (no intention-to-treat analysis)	Outcome assessment in small sample; no numerical data (correct guessing of active extract); the ability of children to identify solution stated only for placebo solution incorrectly identified as active extract by parents (not for active extract correctly identified by parents)	Not stated	Miller technique

Study	Randomisation/ concealment	Blinding/placebo	Loss to follow-up/intention-to-treat analysis	Outcome assessment/Data presentation	Setting/ Sponsor	Other
Rea *et al*., 1984[60]	Order of trials determined arbitrarily by pharmacist; observer did not know order of trials	Patients, technicians and observers blinded; syringes covered in tape so that colour of antigen was not visible; separate double-blind study using different volunteers demonstrated inability to distinguish antigens from saline	No loss to follow-up	Not clear if all outcome measures validated; assessment and scoring systems not defined for all outcome measures	Environmental Care Unit, The Lutheran Brookhaven Medical Systems Hospital, Dallas; study supported in part by research grants from the Human Ecology Research Foundation of the Southwest, Hillcrest Foundation and the American Academy of Otolaryngic Allergies	Miller technique used

Appendix 7 Study quality diagnostic test accuracy studies

Caplin, 1973[45]

Quality criterion	Study	Criterion met?
Was the selection method appropriate? (ideally a random or consecutive cohort of individuals with unknown disease status)	Patients selected according to atopic status (inhalant and/or food sensitivity); unclear if food sensitivity status was known in final patient group	Unclear
Test under evaluation: was it performed as part of a randomised controlled trial and was the trial quality adequate?	Double-blind placebo-controlled crossover trial, but not randomised.	Not met
Reference test: was the gold standard used? (double-blind placebo-controlled oral food challenge)	Tests appeared to be an open food challenge	Not met
Was the provocation test measured independently (blindly) of the gold standard test? And vice versa	No details	Unclear
Was receiving one test dependent on results of other test?	No details	Unclear
Were both tests performed in both patients?	70/70 patients underwent provocative testing; 48/70 patients received the feeding test (not clear why not all patients tested)	Not met
Clear definitions as to what constituted positive/ negative test result?	Symptoms are subjective/self-reported or observed by assessor; attempts were made to ensure unbiased assessment	Met
Were results for both tests clearly stated and was it clear how the sensitivity and specificity were calculated?	Yes	Met

King *et al*., 1988 (part I)[46]

Quality criterion	**Study**	**Criterion met?**
Was the selection method appropriate? (ideally a random or consecutive cohort of individuals with unknown disease status)	Food sensitive individuals selected by 8 physicians on basis of food sensitivity to at least one of the test foods; not clear how many individuals were sensitive to which foods; not clear how this food sensitivity was established	Not met
Test under evaluation: was it performed as part of a randomised controlled trial and was the trial quality adequate?	Double-blind placebo-controlled crossover trial; quality appears acceptable (see Appendix 6)	Met
Reference test: was the gold standard used? (double blind placebo controlled oral food challenge)	Gold standard test was performed in open fashion (i.e. no blinding)	Not met
Was the provocation test measured independently (blindly) of the gold standard test? And vice versa	Unclear	Unclear
Was receiving one test dependent on results of other test?	No	Met
Were both tests performed in both patients?	All patients received both tests	Met
Clear definitions as to what constituted positive/ negative test result?	Few details on outcome assessment	Unclear
Were results for both tests clearly stated and was it clear how the sensitivity and specificity were calculated?	No; not clear which individuals were used for calculations	Not met

Appendix 8 Study outcomes

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Provocation-Neutralisation testing						
Breneman *et al*., 1973[48]	Occurrence of objective or subjective symptoms; reactions between +1 (light) and +3 (marked)	224 positive tests of 732 tests total – not clear how many tests used active extract 192 positive responses to active extract; 345 pluses	224 positive tests of 732 tests total – not clear how many tests used placebo 32 positive responses to active extract; 32 positive expected by chance); 45 pluses (57 expected by chance)	Results were similar to those expected by random sampling	Code was broken after 78/100 patients had been tested; 61 records suitable for computer analysis	Not stated
Breneman *et al*., 1974[52]		74 positive tests of 240 tests total – not clear how many tests used active extract 53 positive (not stated of how many tests; not stated how many positive expected by chance)	74 positive tests of 240 tests total – not clear how many tests used placebo 21 positive (not stated of how many tests; 24 positive expected by chance)		20/30 patients considered satisfactory for computer analysis	
Caplin, 1973[45] *Provocation testing*	Occurrence of subjective symptoms or objective signs associated with allergy or relief of symptoms; whealing not recorded	Between 23 and 28 positive tests to each allergen (not stated how many patients)	26 positive tests to placebo (not stated how many patients)	Similar number of reactions to placebo and active substance (although number of patients not stated)	Not stated	One patient suffered severe asthma in response to placebo

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Diagnostic test accuracy	Not stated ('demonstrable clinical sensitivity') – assume is occurrence of symptoms	15 patients had a positive feeding test (used as reference standard);of those, 12 had a positive subcutaneous provocation test, and 3 a negative one 33 patients had a negative feeding test; of those, 21 had a negative subcutaneous provocation test, and 12 a positive one		Sensitivity: 0.80 Specificity: 0.64	48/70 patients received feeding test only (not stated why patients were not tested)	Not stated
Fox *et al*., 1999[56]	Occurrence of a wheal and/or symptoms Wheal: positive reaction if growth of 2mm within 10 minutes	Any positive reaction (wheal and/or symptoms): Wheat: 52.5% Tomato: 52.0% Soy: 41.6% Egg: 48.4% Cow's milk: 55.2% Corn: 35.5% Chocolate: 60.9% Chicken: 56.4% Sugar: 88.2% Brewer's yeast: 86.8% Beef: 65.9% Banana: 89.5% Baker's yeast: 69.1% NB numbers estimated from graph	Any positive reaction (wheal and or symptoms): Placebo1: 39% Placebo2: 42.9% Placebo3: 35.6% Placebo4: 35.6% Positive reaction by wheal to any of the placebos: 15% Positive reaction by symptoms to any of the placebos: 70% NB numbers estimated from graph	There appear to be more reactions in the form of a wheal in response to allergens compared to placebo; reactions in the form of symptoms were similar between allergens and placebo; overall a higher percentage of patients reacted to allergen compared to placebo, although 70% of patients reacted to at least one placebo with symptoms; there were more positive reactions to allergens in those patients who reacted to saline Statistical tests not performed	No loss to follow-up	Not stated

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Jewett *et al*., 1990[54]	Occurrence of symptoms and identification of active substance by patients	16/60 (27%) active substances judged to be active	44/180 (24%) placebo substances judged to be active	P=0.87 Fisher's exact test	No loss to follow-up, 2 patients tested twice	Not stated
King *et al*., 1988 (part I)[46] *Reliability of provocation testing*	Number and type of symptom provoked during successive trial of provocation testing	Results not listed separately for allergens and placebo tests; data not listed for positive/negative responses to individual foods or for individual patients; no data on responses to placebo		Average correlation of 2nd and 3rd provocation trial (skin response) with 1st summed across 5 foods: 0.68 (significant beyond 0.01 level); average correlation (symptom provocation): 0.40 (significant beyond 0.01 level); individual food correlations ranged from 0.60 to 0.74 (significant beyond 0.01 level); consistency of positive or negative responses was 83% (skin responses) and 75% (symptom provocation; same symptoms half the time); overall consistency of ND (within one dilution) was 82.2%	34/370 comparisons not used due to extract lot change	Not stated
Diagnostic test accuracy	Symptoms provoked	Raw data (number of patients responding/not responding to allergen/placebo for both tests) not stated For skin response and symptom provocation, the average correlation coefficients were 0.78 and 0.61respectively (significant beyond 0.01 level); coefficients for the five foods ranged from 0.95-0.49 (skin response) and 0.50-0.69 (symptom provocation), (significant beyond 0.01 level and between the 0.05 and 0.01 level respectively) Positive and negative agreement approximately 80% for wheat, corn and beef, 70% for white potato and 65% for milk (skin response or symptom provocation)		Provocation testing skin response: Sensitivity: 79.7% Specificity: 72.4% Provocation testing symptom provocation: Sensitivity: 59.6% Specificity: 92.1% PPV and NPV not calculable	No loss to follow-up	Not stated

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
King, 1981[58]	-Pulse rate -Signature size -Bender-Gestalt Test -'Uses-of' test -Estimation of 1 minute -Cognitive-emotional self-report (also self-report of somatic and mixed symptoms) -Mood Affect Adjective Checklist (MAACL) -Number cancellations -Digit Symbol Substitution Test (from the Wechsler Adult Intelligence Scale, WAIS, with 3 parallel forms) -Block Design Test (also from WAIS) -Graphic Constriction Expansion Test -Expectation to react to test, guessing test solution & certainty of guess	**Cognitive-emotional self-report (severity scores):** Trial A1: 0.22 A2: 1.17 A3: 1.67 B1: 0.63 B2: 1.16 B3: 2.79 C1: 0.38 C2: 0.58 C3: 1.38 D1: 0.19 D2: 1.38 D3: 0.43 E1: 0.71 E2: 0.50 E3: 1.94 F1: 0.27 F2: 0.50 F3: 0.23 Patient 1: 0.33 Patient 2: 0.36 Patient 3: 0.67 Patient 4: 1.50 Patient 5: 0.45 Patient 6: 0 Patient 7: 3.88 Patient 8: 0.30 Patient 9: 0.42 Patient 10: 0.36 Patient 11: 1.73	**Cognitive-emotional self-report (severity scores):** Trial A1: 0 A2: 0.50 A3: 0.22 B1: 0 B2: 0.25 B3: 0 C1: 0 C2: 0 C3: 0 D1: 0 D2: 0 D3: 1.00 E1: 0.50 E2: 0.67 E3: 0 F1: 0.10 F2: 0.50 F3: 0 Patient 1: 0.33 Patient 2: 0 Patient 3: 0 Patient 4: 0 Patient 5: 0 Patient 6: 0 Patient 7: 2.0 Patient 8: 0.83 Patient 9: 0 Patient 10: 0.20 Patient 11: 0	Scores by trial: Based on n=17, a sign test showed that greater psychological symptoms were reported following exposure to allergens (p=0.001) Scores by patient: Based on n=19, a sign test showed that greater psychological symptoms were reported following exposure to allergens (p=0.002) Mixed symptoms (e.g. headache, fatigue) greater on allergen trials (for severity and number of symptoms; analysed by patient and trial, sign test statistically significant); Somatic symptoms (e.g. itching, nasal symptoms) greater on allergen trials when analysed by trial (sign test statistically significant, but not be patient (for severity and number of symptoms) No difference in pre- or post dose heart rates (sign test) For all remaining outcome measures, no difference was found (results not listed in full)	All placebo-aware or experimenter aware or suspicious trials or patients were removed in the analysis (e.g. where patient was aware of taste of test substance) 9/30 patients excluded from analysis 21/540 trials removed from most analyses; not clear exactly how many trials/patients analyses are based on	18 requests for relief from uncomfortable symptoms

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
		Patient 12: 2.88 Patient 13: 0.17 Patient 14: 0 Patient 15: 0.33 Patient 16: 0.50 Patient 17: 0.42 Patient 18: 1.33 Patient 19: 3.22 Patient 20: 0.08 Patient 21:1.71	Patient 12: 0.20 Patient 13: 0 Patient 14: 0.67 Patient 15: 0 Patient 16: 0 Patient 17: 0 Patient 18: 0 Patient 19: 0.50 Patient 20: 0.17 Patient 21: 1.00			
Lehman, 1980[57]	Observed changes in degree of swelling and oedema of nasal mucosa (nasocyrogram)	24/28 positive reactions egg; 14/28 corn; 22/28 milk; 20/28 yeast	21/28 positive reaction placebo	Similar number of reactions to allergens and placebo; no statistical tests performed	2/15 patients available for baseline measurement only	Not stated
Mandell & Conte, 1982[55]	Symptoms provoked in response to substance (rheumatic, nervous system, respiratory tract, gastrointestinal, vascular or eye symptoms)	Between 30% (9) and 73.3% (22) patients reacted to each active substance	6.6% (2) patients reacted to placebo	A greater number of patients reacted to active substance compared to placebo; no statistical tests performed	None	Not stated

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Neutralisation therapy						
King *et al*., 1988 (part II)[53]	Symptom response diaries Symptom scores between 1 and 6 (1=much worse, 2=some worse, 3=no relief, 4=some relief, 5=good relief, 6=excellent relief)	Responses of individual treatment periods: 65% improvement of symptoms compared to placebo, 12% no change, 23% aggravation Patients' overall combined treatment result (total response of both treatment sessions): 64% improvement of symptoms compared to placebo, 12% no change, 24% aggravation		Statistically significant result (p= 0.001) using Chi-square test	No loss-to follow up	Not stated
Miller, 1977[41]	Symptom survey during first office visit and during and at end of each 20 day course (0=baseline symptomatology, improvements 1-4; 4=most improvement)	Mean improvement Patient 1: 3.75 Patient 2: 2.83 Patient 3: 3.00 Patient 4: 3.88 Patient 5: 4.00 Patient 6: 3.50 Patient 7: 3.64 Patient 8: 4.00	Mean improvement Patient 1: 1.75 Patient 2: 0.42 Patient 3: 1.17 Patient 4: 3.50 Patient 5: 3.56 Patient 6: 2.25 Patient 7: 2.43 Patient 8: 2.90	Mean difference 1.33, significant at 99.8% level (t-test)	None	Not stated

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
O'Shea & Porter, 1981[59]	Uncontrolled provocation testing: Behaviour provoked: hyperactivity, restlessness, crying, hostility, facial grimaces, lethargy, aggressiveness, defiance, boisterousness, irrationality, physical abusiveness	No. of children (of 15) who had behaviour changes when tested with the following foods: milk (73%), peanuts (47%), tomato (47%), apple (40%), cane sugar (40%), corn (40%), grape (40%), orange (40%), chocolate (33%), wheat (27%), egg (20%); dye: red dye (87%), yellow dye (80%), blue dye (80%); inhalants: dust (27%), mould (13%), tree (33%)	It appears that no placebo was tested for this part of the study	Not performed	No loss to follow up	Not stated
	Neutralisation: Behavioural changes & physical symptoms (daily diary card kept by parents); parents and teachers interviewed weekly by psychologist (using as a guide Abbot's Hyperkinetic index)	Parents' evaluation: Improvement in behaviour on treatment compared to placebo: 11/14 Grading of behaviour: marked worsening (2/14), slightly worse (1/14), slight improvement (1/14), good improvement (7/14), marked improvement (3/14); inconclusive (1/15, lost to follow-up) Teachers' evaluation: Improvement in behaviour on treatment compared to placebo: 7/13; agreement with parents that child was worse on allergy extract compared to placebo: 3/13; no change observed during entire study programme: 3/13 NB 1 child excluded from evaluation as Ritalin was reinstated during placebo phase Evaluations for extract and placebo not listed separately, only comparative improvement or deterioration		Not performed	1/15	1 child's behaviour deteriorated to such an extent during the placebo phase that Ritalin had to be reinstated

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Rapp, 1979[40]	Uncontrolled part of study: Sore derived from Parent Abbott Hyperkinesis Index sheet (variation of Connor's Child Behaviour Rating Scale) Maximum score 30; scores above 17 tend to reflect major activity problems at home and in school; parents rarely complain of activity problems if score is below 10	Baseline Abbott score in 8 children at baseline, after 1 week on diet only, 2-12 weeks on diet, 4-24 weeks with neutralisation treatment and 18-36 months with neutralisation treatment: Patient 1: 22; 12; 8; 5; 5-7 Patient 2: 18; 1; 0; 0; 2-6 Patient 3: 26; 19; 14; 9 ;7 Patient 4: 19; 0;10;5; - Patient 5: -; -; 10; 4; 4 Patient 6: 14; 0; 1; 1; 4 Patient 7: 11; 13; 9; 6; 3 Patient 8: 19; 1; 2; 2; 3 (- =not measured)		Improvement in score over time (NB-uncontrolled study); no statistical tests performed	3/11 children.	Not stated
	Controlled part of study: Correct/incorrect identification of coded food solution by parents	5/8 food solutions identified by parents correctly (no details on children)	3/8 placebo solutions identified incorrectly by parents as food solution (2/3 children identified food solution correctly)	Correct identification of active solution similar to incorrect identification of placebo; no statistical tests performed	2/11 parents gave coded solutions at weekends only; 3/11 did not complete study (2 developed behavioural problems and parents refused to complete study, 1 child was ingesting artificial food colouring)	2 children developed behavioural problems and parents refused to complete study

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
Rea *et al*., 1984[60]	Signs and symptoms (each sign/symptom rated 0-4) compared to signs and symptoms after baseline OFC D is hypothesized to protect against reactions after OFC compared to placebo (=negative response; failure to protect against symptoms =positive response; negative and slight changes grouped together as negative; moderate and strong changes grouped together as positive)	Negative response (protection) in 12/20 trials (0.6+/-0.11, SD=0.5)	Negative response (protection) in 5/40 trials (0.13 +/- 0.05, SD=0.33)	Chi-square test; p<0.001	None	Not stated
	Visual Analogue Discomfort rating (degree of discomfort marked on 10 cm line) Highest score during baseline period compared to highest score after the OFC; significant change defined as increase by 1 cm	Negative response in 12/20 trials (0.6+/-0.11, SD=0.5)	Negative response in 6/40 trials (0.15 +/-0.06, SD=0.36)	Chi-square test; p<0.001	None	
	Symbol-Digit Modalities Test (ability to translate geometric figures into numbers)	Negative response in 16/20 trials (0.8+/-0.09, SD=0.41)	Negative response in 20/40 trials (0.5 +/-0.08, SD=0.51)	Chi-square test; p<0. 01	None	

Study	Outcomes measured/ Assessment Tool/ Scale	Size of effect (CI), SD Intervention (Allergen)	Size of effect (CI), SD Placebo	Direction of effect/ Summary measure/ Statistical tests/p-values	Loss to follow up	Side effects (other than expected symptoms)
	A baseline to post-OFC change in the number of errors was considered substantial if it exceeded 8					
	Aaron Smith Symbol-Digit Modalities Subtest (ability to translate numbers into geometric figures) A baseline to post-OFC change in the number of errors was considered substantial if it exceeded 8	Negative response in 16/20 trials (0.8+/- 0.09, SD=0.41)	Negative response in 26/40 trials (0.65 +/- 0.08, SD=0.48)	Chi-square test; p<0. 05	None	
	Apical Heart rate (measured for 1 minute) A change of eight beats or more per minute up or down was defined as substantial	17/20 trials (0.85+/- 0.08, SD=0.37)	25/40 trials (0.63 +/- 0.08, SD=0.49)	Chi-square test; p<0. 05	None	
	Subject's Signature (subjects asked to write names) Presence or absence of substantial deterioration in the quality of the signature as noted by the observer	15/20 trials (0.75+/- 0.10, SD=0.44)	18/40 trials (0.45 +/- 0.07, SD=0.08)	Chi-square test; p<0. 001	None	

10 References

1. Van MT, Jr. Unproven procedures for diagnosis and treatment of food allergy. [Review] [52 refs]. *New England & Regional Allergy Proceedings* 1987;**8**:17-21.

2. Kay AB,.Lessof MH. Allergy. Conventional and alternative concepts. A report of the Royal College of Physicians Committee on Clinical Immunology and Allergy. [Review] [104 refs]. *Clinical & Experimental Allergy* 1992;**22 Suppl 3**:1-44.

3. Position paper on allergenic immunotherapy: report of a BSACI Working Party January-October 1992. *Clinical & Experimental Allergy* 1993;**23 (Supplement 3, August)**:1-44.

4. Allergy UK. Non-conventional allergy tests. http://www.allergyfoundation.com/allergy_othertests.html . Accessed July 2003.

5. Reisman RE. American Academy of Allergy: position statement -controversial techniques. *Journal of Allergy & Clinical Immunology* 1981;**67**:333-8.

6. Report from National Centre for Health Care Technology, Summary of Assessments. *Journal of the American Medical Association* 1981;**246**:1499.

7. Clinical ecology. American College of Physicians. [see comments]. [Review] [91 refs]. *Annals of Internal Medicine* 1989;**111**:168-78.

8. The Burghwood Clinic. http://www.burghwoodclinic.co.uk. Accessed February 2002.

9. The Breakspear Clinic. http://www.breakspearmedical.com.Accessed February 2002.

10. Airedale Allergy Centre. http://www.airedaleallergy.fsbusiness.co.uk/AAtreatment.htm . Accessed July 2002.

11. Gerdes KA. Provocation/Neutralization testing: A look at the controversy. *Clinical Ecology* 1989; **6**:21-3.

12. Aggressive Research Intelligence Facility (ARIF). http://www.bham.ac.uk/arif/. Accessed July 2003.

13. Royal College of Physicians Working Party on the provision of allergy services in the UK. Allergy – the unmet need. A blueprint for better patient care. London: Royal College of Physicians, June 2003

14. Sicherer, S. H. Food allergy. *Lancet* 2002; **360**: 701-710

15. Poulsen, L. K. *In vitro* tests in the diagnosis of food allergy. *Clinical & Experimental Allergy* 1998; **28**:1457-1459

16. Bindslev-Jensen C. Food allergy. [Review] [0 refs]. *BMJ* 1998;**316**:1299-302.

17. Ferguson A. Definitions and diagnosis of food intolerance and food allergy: consensus and controversy. [Review] [7 refs]. *Journal of Pediatrics* 1992;**121**:S7-11.

18. Egger J, Carter CM, Wilson J, Turner MW, Soothill JF. Is migraine food allergy? A double-blind controlled trial of oligoantigenic diet treatment. *Lancet* 1983;**2**:865-9.

19. Jones VA, McLaughlan P, Shorthouse M, Workman E, Hunter JO. Food intolerance: a major factor in the pathogenesis of irritable bowel syndrome. *Lancet* 1982;**2**:1115-7.

20. Darlington LG, Ramsey NW, Mansfield JR. Placebo-controlled, blind study of dietary manipulation therapy in rheumatoid arthritis. *Lancet* 1986;**1**:236-8.

21. Vatn, M. H. Symptoms and manifestations of food intolerance. *Environmental Toxicology and Pharmacology* 1997; **4**:51-53.

22. Young E, Stoneham MD, Petruckevitch A, Barton J, Rona R. A population study of food intolerance. [see comments]. *Lancet* 1994;**343**:1127-30.

23. British Nutrition Foundation. http://www.nutrition.org.uk. Accessed February 2002.

24. Jansen JJN, Kardinaal AFM, Huijbers G, Vlieg-Boerstra BJ, Martens BPM, Ockhuizen T. Prevalence of food allergy and intolerance in the adult Dutch population. *Journal of Allergy & Clinical Immunology* 1994;**93**:446-56.

25. Altman DR,.Chiaramonte LT. Public perception of food allergy. *Journal of Allergy & Clinical Immunology* 1996;**97**:1247-51.

26. Anthony HM, Birtwistle S, Brostoff J, Eaton KK, Hearn G, Maberly DJ *et al*. Food intolerance (letter). *Lancet* 1994;**344**:136-7.

27. Finn R. Food intolerance (letter). *Lancet* 1994;**344**:137.

28. Moneret-Vautrin DA,.Kanny G. Food intolerance (letter). *Lancet* 1994;**344**:137.

29. VanArsdel PP, Jr.,.Larson EB. Diagnostic tests for patients with suspected allergic disease. Utility and limitations. [see comments]. [Review] [45 refs]. *Annals of Internal Medicine* 1989;**110**:304-12.

30. Bahna, S. L. Diagnosis of food allergy. *Annals of Allergy, Asthma, & Immunology* 2003; **90** (Suppl 3):77-80

31. King HC,.King WP. Alternatives in the diagnosis and treatment of food allergies. [Review] [21 refs]. *Otolaryngologic Clinics of North America* 1998;**31**:141-56.

32. Anthony, H. M. Adverse reactions to foods. *Journal of Nutritional and Environmental Medicine* 2001;**11**: 127-135.

33. Caffarelli, C. and Petroccione, T. False-negative food challenges in children with suspected food allergy. *Lancet* 2001; **358**: 1871-1872.

34. Rusznak, C. and Davies , R. J. Diagnosing allergies. *BMJ* 1998; **316**:686-689.

35. Rinkel HJ. Food allergy II, the technique and clinical application of individual food tests. *Annals of Allergy* 1944;**2**:504-14.

36. Lee, C. H., Williams, R. I., and Binkley, E. L. Provocative testing and treatment for foods. *Arch Otolaryngol* 1969; **90**: 113.

37. Miller JB. Food Allergy, Provocative Testing and Injection Therapy. Springfield, IL: Charles Thomas Co., 1972.

38. Miller, J. B. The Case for Neutralizing (Optimal Dose) Immunotherapy. http://www.allergycentre.com/neutralizing_dose.htm . 2003. 15-7-0003. Accessed July 2003.

39. Miller, J. B. The Maximum Intradermally Tolerated Dose (MITD) Method of Food Allergy testing and Immunotherapy: New Concepts. http://www.allergycentre.com/mitd.htm. Accessed July 2003.

40. Rapp DJ. Food allergy treatment for hyperkinesis. *Journal of Learning Disabilities* 1979;**12**:608-16.

41. Miller JB. A double-blind study of food extract injection therapy: a preliminary report. *Annals of Allergy* 1977;**38**:185-91.

42. Grieco, M. H. Controversial practices in allergy. *JAMA* 1982; **247**(22): 3106-3111.

43. Rinkel H.J., Lee, C. H., and Brown, D. W. Jr et al. The diagnosis of food allergy. *Arch Otolaryngol* 1964; **79**: 71-76.

44. Goldberg BJ,.Kaplan MS. Controversial concepts and techniques in the diagnosis and management of food allergies. *Immunology & Allergy Clinics of North America* 1991;**11**:863-84.

45. Caplin, I. Report of the Committee on Provocative Food Testing. *Ann Allergy* 1973; **31**: 375.

46. King WP, Rubin WA, Fadal RG, Ward WA, Trevino RJ, Pierce WB *et al*. Provocation-neutralization: a two-part study. Part I. The intracutaneous provocative food test: a multi-center comparison study. *Otolaryngology - Head & Neck Surgery* 1988;**99**:263-71.

47. Podell RN. Food extract injection for food sensitivity. Valid technique or 'black magic'? *Postgraduate Medicine* 1984;**76**:59-62.

48. Breneman, J. C. Crook W. C. Deamer W. et al. Report of the Food Allergy Committee on the sublingual method of provocative testing for food allergy. *Ann Allergy* 1973; **31**: 382-385.

49. Hills, M. and Armitage, P. The two-period cross-over clinical trial. *British Journal of Clinical Pharmacology* 1979; **8**: 7-20.

50. Deeks, J. J. Systematic review in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001; **323**(7305): 157-162.

51. Undertaking Systematic Reviews of Reasearch on Effectiveness: CRD's Guidance for those carrying out or commissioning reviews. 2nd edition. 2001. NHS Centre for Reviews and dissemination, University of York.

52. Final report of the Food Allergy Committee of the American College of Allergists on the clinical evaluation of sublingual provocation testing method for diagnosis of food allergy. *Ann Allergy* 1974; **33**:164-166.

53. King WP, Fadal RG, Ward WA, Trevino RJ, Pierce WB, Stewart JA *et al*. Provocation-neutralization: a two-part study. Part II. Subcutaneous neutralization therapy: a multi-center study. *Otolaryngology - Head & Neck Surgery* 1988;**99**:272-7.

54. Jewett DL, Fein G, Greenberg MH. A double-blind study of symptom provocation to determine food sensitivity. [see comments]. *New England Journal of Medicine* 1990;**323** :429-33.

55. Mandell, M. and Conte, A. A. The Role of Allergy in Arthritis, Rheumatism, Polysymptomatic Cerebral Visceral and Somatic Disorders. *Journal of IAPM* 1982; **7**:5-16.

56. Fox RA, Sabo BM, Williams TP, Joffres MR. Intradermal testing for food and chemical sensitivities: a double-blind controlled study. *Journal of Allergy & Clinical Immunology* 1999;**103**:907-11.

57. Lehman CW. A double-blind study of sublingual provocative food testing: a study of its efficacy. *Annals of Allergy* 1980;**45**:144-9.

58. King DS. Can allergic exposure provoke psychological symptoms? A double-blind test. *Biological Psychiatry* 1981;**16**:3-19.

59. O'Shea JA,.Porter SF. Double-blind study of children with hyperkinetic syndrome treated with multi-allergen extract sublingually. *Journal of Learning Disabilities* 1981;**14**:189-91.

60. Rea WJ, Podell RN, Williams ML. Elimination of oral food challenge reaction by injection of food extracts. A double-blind evaluation. *Archives of Otolaryngology* 1984;**110**:248-52.

61. Surrey Allergy Clinic. Food Allergy Tests. http://www.allergy-clinic.co.uk . Accessed July 2003.

62. National Asthma Campaign. All about asthma: food for thought. http://www.asthma.org.uk/about/an026.php. Accessed July 2003.

63. YORKTEST. Laboratory Tests. http://homeinonhealth.com. Accessed July 2003.

64. King DS. The reliability and validity of provocative food testing: a critical review. [Review] [56 refs]. *Medical Hypotheses* 1988;**25**:7-16.

65. Smith, R. Beyond conflict of interest. *BMJ* 1998; **317**:291-292.

66. Teuber SS,.Vogt PJ. An unproven technique with potentially fatal outcome: provocation/neutralization in a patient with systemic mastocytosis. *Annals of Allergy, Asthma, & Immunology* 1999;**82**:61-5.

67. Bronsky E.A., Burley, D. P., and Ellis, E. F. Evaluation of the provocative skin test technique, abstracted. *J Allergy* 1971; **47**:104.

68. Crawford, L. V., Lieberman, P., Harfi, H. A., Hale, R., Nelson, H., Selner, J., Wittig, H., Postman, M., and Zietz, H. A double-blind study of subcutaneous food testing sponsored by the food committee of the American Academy of Allergy. *J Allergy Clin Immunol* 1976; **57**:236.

69. Kailin, E. W. and Collier, R. 'Relieving' therapy for antigen exposure. *JAMA* 1971; **217**: 78-82.

70. Provocation testing and food sensitivity. [letter; comment.]. *New England Journal of Medicine* 1991;**325**:1171-4.

71. Dixon HS. Dysphonia and delayed food allergy: a provocation/neutralization study with strobovideolaryngoscopy. *Otolaryngology - Head & Neck Surgery* 1999;**121**:418-29.

72. Draper, L. W. Food testing in allergy: Intradermal, provocative or deliberate feeding. *Arch Otolaryngol* 1972; **95**:169-174.

73. Duncan RB,.Duncan DD. Otolaryngeal allergy in Wellington. 1971-1975. *New Zealand Medical Journal* 1977;**85**:45-52.

74. Endicott, J. N. and Stucker, F. J. Allergy in Menieres disease related fluctuating hearing loss preliminary findings in a double-blind crossover clinical study. *The Laryngoscope* 1977; **87**:1650-1657.

75. Forman R. A critique of evaluation studies of sublingual and intracutaneous provocative tests for food allergy. *Medical Hypotheses* 1981;**7**:1019-27.

76. Green, M. Sublingual provocative testing for food and FD and C dyes. *Ann Allergy* 1974; **33**:274-281.

77. Jewett, D. L. and Greenberg, M. R. Placebo responses in intradermal provocation testing with food extracts. *J Allergy Clin Immunol* 1985; **75**:205.

78. King WP. Testing for food allergy: a statistical comparison of cytotoxic and intracutaneous tests. *Laryngoscope* 1978;**88**:1649-59.

79. Mandell, M. and Conte, A. The role of allergy in arthritis, rheumatism and associated polysymptomatic cerebro-viscero-somatic disorders: a double-blind provocation test study. *Annals of Allergy* 1980; **44**: 51.

80. McGovern JJ. The provocative sublingual method. *Annals of Allergy* 1981;**46**:44-7.

81. Podell RN. Provocative neutralization. *JAMA* 1983;**249**:2021.

82. Rapp DJ. Double-blind confirmation and treatment of milk sensitivity. *Medical Journal of Australia* 1978;**1**:571-2.

83. Rapp DJ. Sublingual provocative food testing. *Annals of Allergy* 1981;**46**:44.

84. Rapp DJ. Diagnostic testing and immunotherapy for allergy. *JAMA* 1988;**260**:341-2.

85. Shaver EF, Jr. Allergic management of Meniere disease. *Archives of Otolaryngology* 1975;**101**:96-9.

86. Rapp, D. J. Weeping eyes in wheat allergy. *Trans Am Soc Ophthalmol Otolaryngol Allergy* 1982; **18**(149):150.

87. Heyse, H. P. Intracutaneous and subcutaneous provocation and neutralization testing and neutralization therapy for food allergies. National Centre for Health Care Technology, US Department of Health and Human Services Assessment Report Series 1981.