National Institute for Health and Care Excellence

Final

Myalgic encephalomyelitis (or encephalopathy) / chronic fatigue syndrome: diagnosis and management

[F] Pharmacological interventions

NICE guideline NG206

Evidence reviews underpinning recommendations and research recommendations in the NICE guideline

October 2021

Final

These evidence reviews were developed by the National Guideline Centre

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-4221-3

Contents

1. Pharmacological interventions

1.1. Review question

What is the clinical and cost effectiveness of pharmacological interventions for people with ME/CFS? What are the experiences of people who have had interventions for ME/CFS?

1.1.1. Introduction

No drug treatment has been found to be a safe and effective cure for ME/CFS. Pharmacological interventions are however commonly used for symptomatic relief in people with ME/CFS, for example for pain and sleep, even though evidence from clinical trials in ME/CFS may be lacking. Approaches can also be used for co-morbid conditions such as irritable bowel syndrome, migraine-type headaches, postural orthostatic tachycardia syndrome or vitamin D deficiency. Many people report self-medicating with vitamins and supplements.

The committee evaluated evidence from clinical effectiveness studies and patient experience from a wide range of pharmacological management strategies to inform the recommendation in these areas.

The clinical and cost effectiveness methods and evidence found are outlined Evidence review G: Non pharmacological management as well as the methods and evidence found for the review on the experiences of people who have had interventions for ME/CFS.

1.1.2. Summary of the protocol

For full details see the review protocol in Appendix A.

Table 1: PICO characteristics of review question

Population	Adults, children and young people who are diagnosed as having ME/CFS.

Intervention(s)	These can include (but are not restricted to):
	• Antidepressants
	  ○ Include all SSRIs / SNRIs and tricyclics
	• Immunomodulatory drugs. For example:
	  ○ Rintatolimod (Ampligen)
	  ○ Rituximab
	• Pro-inflammatory cytokines. For example:
	  ○ Anakinra
	• Sleep medication. For example:
	  ○ Melatonin
	• Pain relief. For example:
	  ○ Pregabalin
	  ○ Gabapentin
	  ○ cannabinoids
	• Antiviral drugs
	• Oral corticosteroids
	  ○ fludrocortisone / hydrocortisone / other steroids
	• Modafinil
	• Sodium Valproate
	• Low dose Naltrexone

	Combinations of treatments (including combinations with non-pharmacological treatments) are allowed.
Comparison(s)	• No treatment • Each other (both within and between drug classes) • Placebo/control/usual care • Non-pharmacological interventions

Outcomes	Longest follow up available:
	
	CRITICAL OUTCOMES:
	• Mortality
	• Quality of life (any validated scales). For example:
	  ○ SF36
	  ○ EQ5D
	• General symptom scales (any validated scales). For example:
	  ○ De Paul Symptom Questionnaire
	  ○ Self-Rated Clinical Global Impression Change Score
	• Fatigue/fatiguability (any validated scales). For example:
	  ○ Chalder fatigue Scale
	  ○ Fatigue Severity Scale
	  ○ Fatigue Impact scale
	• Physical functioning (any validated scales). For example:
	  ○ SF36 physical function
	  ○ SF36 PCS
	• Cognitive function (any validated scales). For example:
	  ○ MMSE
	• Psychological status (any validated scales). For example:
	  ○ Hospital Anxiety and Depression Scale
	  ○ Becks Depression Inventory
	• Pain (VAS/NRS)
	• Sleep quality (any validated scales). For example:
	  ○ Pittsburgh Sleep quality Index
	  ○ Epworth Sleepiness Scale
	  ○ Leeds Sleep Evaluation Questionnaire VAS
	• Treatment-related adverse effects
	• Activity levels – step counts
	• Return to school / work
	• Exercise performance measures. For example:
	  ○ Hand grip
	  ○ Maximal Cycle Exercise Capacity
	  ○ 6 min walk
	  ○ Timed Up and Go
	  ○ 5 repetition sit to stand
	  ○ 40m walk speed
	  ○ Step test
	IMPORTANT OUTCOMES:
	• Care needs
	• Impact on families and carers

Study design	• Randomised controlled trials

	• Systematic review of randomised controlled trials. For a systematic review to be included it must be conducted to the same methodological standard as NICE guideline reviews. If sufficient details are not provided to include a relevant systematic review, the review will be used for citation searching. Cross-over RCTs will be considered provided the wash-out period is considered adequate. Non RCTs will not be considered.

1.1.3. Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A and the methods document.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

1.1.4. Effectiveness evidence

1.1.4.1. Included studies

A search was conducted for randomised trials comparing the effectiveness of pharmacological interventions for adults, children and young people who are diagnosed with ME/CFS. A variety of pharmacological interventions were identified; immunomodulatory drugs, antidepressants, corticosteroids, antihypertensive drugs, central nervous system stimulants, antiviral drugs, 5-HT3 antagonists, Galantamine hydrobromide, antihistamines, pro-inflammatory cytokine antagonists and Staphylococcus vaccine. The majority of the interventions are compared to placebo. The study populations were adults with mixed or unclear severity categories.

Thirty studies were included in the review;[4, 8, 9, 22, 23, 34, 37, 44, 47, 52, 53, 56, 61, 63-65, 68, 70, 77, 79, 84, 86, 88, 91, 93, 97, 99, 103, 107, 109] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3 - Table 19).

See also the study selection flow chart, study evidence tables, forest plots and GRADE tables in the appendices.

There was a small amount of limited evidence for a number of different drugs. Network meta-analysis was considered for the comparison of drugs but was not pursued because of insufficient data available for the relevant outcomes. In addition there were substantial differences between the study interventions, comparators, populations and outcomes. There was a general lack of evidence of clinically important differences for any pairwise comparisons.

1.1.4.2. Excluded studies

See the excluded studies list in Appendix J.

1.1.5. Summary of studies included in the effectiveness evidence

It should be noted that post exertional malaise (PEM) is also referred to as post exertional symptom exacerbation (PESE). PESE is the committee's preferred term.

Table 2: Summary of studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Immunomodulatory drugs				
Fluge 2011[22]	Immunomodulatory drug – Rituximab Rituximab 500 mg/m² (maximum 1000 mg), diluted in saline to a concentration of 2 mg/ml, given twice two weeks apart, with nurse surveillance and according to local guidelines used for treating B-cell lymphomas. Infusion bags had double plastic covers to avoid content identification by nurse or patient. Duration: 2 weeks Versus Placebo An equal volume of saline, given twice two weeks apart, with nurse surveillance and according to local guidelines used for treating B-cell lymphomas. Infusion bags had double plastic covers to avoid content identification by nurse or patient. Duration: 2 weeks No additional infusions, or other interventions, were given during follow-up. All	N=30 people with CFS, diagnosed by a neurologist, according to the CDC criteria 1994 (Fukuda 1994). Pre-treatment evaluation included thorough clinical examination, standard laboratory tests and further diagnostic tests if pre-treatment evaluation revealed any relevant abnormality that could explain the severe fatigue. Strata details: adults (18-65 years); severity mixed or unclear.	Quality of life (SF-36; physical and mental composite scores)	Conducted in Norway Fatigue, cognitive, pain and other symptoms scores calculated as the mean of different self-reported symptom scales (0-6) – measures not validated. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; Post exertional Malaise (PEM) is

Study	Intervention and comparison	Population	Outcomes	Comments
	patients were given oral cetirizine 10 mg, paracetamol 1 g, and dexamethasone 8 mg prior to infusion.			not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Fluge 2019[23]	Immunomodulatory drug – Rituximab Induction treatment with 2 infusions, 2 weeks apart, of rituximab (MabThera, Roche), 500 mg/m2 of body surface area (maximum of 1000 mg). In the maintenance phase, patients received a 500-mg fixed dose of rituximab at 3, 6, 9, and 12 months. Duration: 12 months Versus Placebo Induction treatment with 2 infusions, 2 weeks apart, of 500 mg/m2 of body surface area (maximum of 1000 mg) saline with added human albumin (Flexbumin [Baxalta] or Albunorm [Octapharma]), 0.4 mg/mL, to ensure no visible difference from the active comparator. In the maintenance phase, patients received a 500-mg fixed dose of saline with human albumin at 3, 6, 9, and 12 months. Duration: 12 months	N=152 people with ME/CFS according to Canadian consensus criteria (Carruthers 2003). Patients where the workup uncovered other pathology as a possible cause of symptoms were excluded. Strata details: adults (18-65 years); severity mixed or unclear (mild or mild/moderate 40%, moderate 30%, moderate/severe and severe 30%; patients with very severe ME/CFS (WHO function class IV), who were totally bedridden and in need of care were excluded).	Fatigue/fatigability (Fatigue severity scale; fatigue numeric rating scale) Physical function (SF36 physical function; function level percentage) Adverse events (any adverse events and any serious adverse events with possible/probable relation to intervention; suspected unexpected serious adverse reactions) Activity levels (mean steps per 24 hours)	Conducted in Norway PEM reporting: all participants met the Canadian criteria which has PEM as a compulsory feature.

Study	Intervention and comparison	Population	Outcomes	Comments
	One hour before infusions, all patients received premedication with1g of oral acetaminophen, 10 mg of cetirizine, and 8mg of dexamethasone.			
Lloyd 1990[44]	Immunomodulatory drug – IV immunoglobulin G High-dose intravenous IgG was administered intravenously by continuous infusion in a dosage of 2g/kg. Three infusions lasting 24 hours were administered at monthly intervals. Duration: 3 months (3 infusions) Versus Placebo A placebo solution of 10% w/v maltose was administered intravenously for 24 hours at an equivalent volume to the IgG infusion. Duration: 3 months (3 infusions)	N=49 people with CFS, diagnosed according to Lloyd 1988 criteria. A physical examination and standardized investigation protocol excluded other chronic infectious or immunodeficiency-related disorders. Strata details: adults (18-65 years); severity mixed or unclear.	Psychological status (Hamilton Depression Scale & Zung Self-Rating Depression Scale) Adverse events (phlebitis & constitutional symptoms)	Conducted in Australia 'Quality of life visual analogue scales modified to include 10 aspects of physical and neuropsychiatric symptomatology typical of CFS (fatigue, headaches, myalgia, concentration impairment and functional activity)' were completed, but a single value for QoL was reported for each group. This outcome was not extracted due to lack of information on how the overall score was derived, the range, or the direction of scales

Study	Intervention and comparison	Population	Outcomes	Comments
				PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – Lloyd 1988 criteria were excluded from the diagnostic criteria review on the basis there was unclear methodology for the development of the criteria and have therefore been downgraded here for indirectness. The study states that the criteria emphasize the same features as the criteria published subsequently by the Centers for Disease Control [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
Peterson 1990[65]	Immunomodulatory drug – IV immunoglobulin G Patients were scheduled to receive a total of six infusions of IV IgG. Infusions were initiated at a rate of 0.5 mL/kg/hour and increased as per the IV IgG package insert to a maximum of 4 mL/kg/hour. Duration: once per month for 6 months Versus Placebo Participants in the placebo group received the same course of IV but IgG was replaced with an exactly correlating volume of a 1% albumin solution as placebo. Duration: once per month for 6 months	N=30 people with CFS, diagnosed according to the diagnostic criteria of CDC /Holmes 1988 after thorough medical, psychometric, and psychiatric evaluations did not establish another explanation for chronic fatigue. Strata details: adults (18-65 years); severity mixed or unclear.	Physical functioning (on the Medical Outcome Study Short Form, a precursor to SF36) Psychological status (mental health on the Medical Outcome Study Short Form) Adverse events (major adverse events)	Conducted in USA PEM reporting: 96.4% of participants had “post-exertional fatigue (prolonged)” but as there was no additional information the committee was not satisfied that this symptom was PEM)Serious population indirectness – Holmes 1988 criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear; unclear if symptom described in study is PEM [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
Strayer 2012[88]	Immunomodulatory drug – rintatolimod (Ampligen) Patients initially received a 200 mg IV dose of rintatolimod twice weekly for two weeks. Following this, a 400 mg dose of rintatolimod was administered twice weekly for 40 weeks. Total duration: 42 weeks. Versus Placebo Patients initially received a 200 mg IV dose of placebo (physiological saline) twice weekly for two weeks. Following this, a 400 mg dose of saline placebo was administered twice weekly for 40 weeks. Total duration: 42 weeks.	N=234 people with CFS, diagnosed according to the CDC criteria 1988 (Holmes 1988) & 1994 CDC criteria. A complete medical history to confirm diagnosis was performed at baseline. A complete physical examination and laboratory tests were also performed. Only subjects with Karnofsky Performance Score values ranging from 20 to 60 were eligible. Strata details: adults (18-65 years); severity mixed or unclear. Participants were stratified according to treadmill duration (≥9 minutes vs >9 minutes) then randomised.	Quality of life (Vitality Score subscale) Adverse events (Serious Adverse Events with possible/probable relation to intervention) Physical functioning (Karnofsky Performance Score & Activities of Daily Living) Exercise performance measure (treadmill exercise duration)	Conducted in USA. SD or CIs were not reported for quality of life and physical functioning outcomes and therefore are not analysed. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – Holmes 1988 criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Vollmer-Conna 1997[99]	Immunomodulatory drug – IV immunoglobulin G Participants received 3 infusions, each lasting 24 hours at monthly intervals. Three dose arms:	N=99 people with CFS, diagnosed according to Schluederberg criteria.	Adverse events (constitutional symptoms) [pooled]	Conducted in Australia Study reported quality of life

Study	Intervention and comparison	Population	Outcomes	Comments
	1. IV IgG (Intragram) at 0.5 g/kg (n=22) 2. IV IgG (Intragram) at 1.0 g/kg (n=28) 3. IV IgG (Intragram) at 2.0 g/kg (n=23) Versus <u>Placebo</u> Identical placebo solution given via IV (n=26). Participants received 3 infusions each lasting 24 hours at monthly intervals. IV IgG dose arms are pooled for analysis.	Strata details: adults (18-65 years); severity mixed or unclear.		outcomes (QAL, POMS depression, confusion, fatigue & energy) and physical function (Karnofsky scale but no analysable data were reported for these outcomes. Results are reported narratively in the clinical evidence table. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – unclear criteria used. Schluederberg 1992 publication was not included in the diagnostic criteria review as it presented a review of the CDC 1988 criteria rather than an original set of criteria [original

Study	Intervention and comparison	Population	Outcomes	Comments
				analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Antidepressants				
Arnold 2015[4]	<u>Antidepressants – serotonin-norepinephrine reuptake inhibitors – Duloxetine hydrochloride</u> Duloxetine hydrochloride at 30 mg once a day for 1 week, then 60 mg one a day for a following 3 weeks. The dose then increased to 90mg per day for next 4 weeks (as tolerated). If highest doses not tolerated doses could be as low as 60mg per day. Duration: 12 weeks. Versus <u>Placebo</u> Identical placebo given in same way as study drug. Duration: 12 weeks.	N=60 people with CFS, diagnosed according to the CDC criteria 1994 (Fukuda 1994) Strata details: adults (18-65 years); severity mixed or unclear.	Quality of life (eight SF36 subscales) Fatigue (MFI-20, five subscales) Psychological status (HADS anxiety & depression) Pain (Brief Pain Inventory for severity & interference) Symptom scales (Clinical Global Impression of Severity & Improvement; CDC symptom inventory) Adverse events	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Hickie 2000[34]	<u>Antidepressants – MAOIs – Moclobemide</u> Moclobemide is a reversible inhibitor of monoamine oxidase (RIMA). Treatment was initially given as 150 mg tablets to be taken	N=90 people with CFS, diagnosed according to Lloyd 1988 criteria.	Physical functioning (Karnofsky Performance Index)	Conducted in Australia.

Study	Intervention and comparison	Population	Outcomes	Comments
	twice daily after meals. After 1 week the dose was increased to 2 tablets in morning and 1 tablet at night for a total dose of 450mg/day. This was increased to 600mg/day if tolerated. Duration: 6 weeks. Versus <u>Placebo</u> increased to 2 tablets in morning and 1 tablet at night for a total dose of 450mg/day. This was increased to 600mg/day if tolerated. Duration: 6 weeks. Concurrent care: Intermittent night dosages of benzodiazepines were allowed for insomnia.	Strata details: adults (18-65 years); severity mixed or unclear.	Psychological status (Profile Of Mood States, POMS fatigue, vigor and depression) Symptom scales (Globally improved cases (patient-rated))	Results reported are standard units of improvement (pre-treatment score-posttreatment score/SD of mean pre-treatment score) PEM reporting: the percentage of participants with PEM was not reported.Serious population indirectness - Lloyd 1988 criteria were excluded from the diagnostic criteria review on the basis there was unclear methodology for the development of the criteria and have therefore been downgraded here for indirectness [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
Pardini 2011[63]	Antidepressants – selective serotonin reuptake inhibitors – Fluoxetine Fluoxetine 20 mg u.i.d. Duration: 12 weeks. Versus Other - Amisulpride Amisulpride (a substituted benzamide) is an atypical antipsychotic. Given at 25 mg b.i.d. Duration: 12 weeks.	N=40 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994) Strata details: adults (18-65 years); severity mixed or unclear.	Quality of life (SF-12) Fatigue (Fatigue Severity Scale) Psychological status (HADS anxiety & depression) Pain (on VAS) Adverse events (FIBSER – global burden) Symptom scales (Clinical Global Impression of severity, CGI-S)	Conducted in Italy. PEM reporting: the percentage of participants with PEM was not reported.Serious population indirectness – 1994 CDC criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis.
Vercoulen 1996[97]	Antidepressants – selective serotonin reuptake inhibitors – Fluoxetine Fluoxetine 20 mg once daily. Duration: 8 weeks. Versus Placebo Placebo given once daily. Duration: 8 weeks.	N=107 people with CFS, diagnosed according to Oxford Criteria (Sharpe 1991) criteria. Strata details: adults (18-65 years); severity mixed or unclear.	Fatigue (Checklist Individual Strength (CIS) fatigue) Psychological status (Beck Depression Inventory) Adverse events (tremor & perspiration)	Conducted in Netherlands. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness - Oxford criteria used; PEM is not a compulsory feature

Study	Intervention and comparison	Population	Outcomes	Comments
			Symptom scales (self-reported global improvement)	[original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Wearden 1998[103]	This four-arm study compared an antidepressant, graded exercise and placebos of both: 1. Fluoxetine & exercise control 2. Graded exercise & placebo 3. Fluoxetine & graded exercise 4. Placebo & exercise control Fluoxetine (antidepressant – selective serotonin reuptake inhibitor) Fluoxetine at a fixed daily dose of 20 mg. Duration: 6 months. Versus Graded exercise Subjects were instructed to carry out their preferred aerobic activity (usually walking/ jogging, swimming or cycling), for 20 minutes, at least three times per week. The intensity of the activity was initially set at a level which utilised oxygen at approximately 75% of the subject's tested functional maximum. Exercise intensity was increased when there was a consistent recorded reduction of 10 beats per minute in post-exercise heart rate for one week and two points on the perceived exertion scale.	N=136 people with CFS, diagnosed according to Oxford Criteria (Sharpe 1991). Strata details: adults (18-65 years); severity mixed or unclear.	Fatigue (14-item Chalder fatigue scale) Psychological status (depression on the Hospital Anxiety and Depression Scale) Exercise performance measure (functional work capacity/VO2 peak)	Conducted in United Kingdom. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness - Oxford criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
	This group also received a placebo fluoxetine capsule of similar taste and appearance, taken daily. Duration: 6 months. Versus <u>Placebo or exercise control</u> Fluoxetine placebo: a capsule of similar taste and appearance, taken daily for 6 months. Exercise control consisted of a placebo exercise programme in which participant activity diaries were reviewed by a physiotherapist. Subjects were not offered any specific advice on how much exercise they should be taking but were told to do what they could when they felt capable and to rest when they felt they needed to.			
Corticosteroids				
Kakumanu 2003[37]	<u>Nasal corticosteroids – Flunisolide</u> Nasal (not oral) corticosteroid (Flunisolide) self-administered with two sprays twice daily. Duration: 4 weeks - 8 weeks (see comments). Versus <u>Placebo</u> Saline spray, two sprays daily. Duration: 4 weeks - 8 weeks (see comments).	N=28 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). All participants also had rhinitis Strata details: adults (18-65 years); severity mixed or unclear.	Sleep quality (Epworth Sleepiness Scale & Functional Outcomes of Sleep Questionnaire) Fatigue (Chronic Fatigue Syndrome Severity Rating)	Conducted in USA. This was a hybrid parallel/cross-over trial design. There were 4 groups of 7 who were treated as follows: active throughout; placebo throughout; active then placebo; placebo then active. Thus 21 had the active treatment at

Study	Intervention and comparison	Population	Outcomes	Comments
				one point. In the analysis the results from these 21 people were aggregated without any apparent adjustments for some having had the other treatment (with the possibility of carryover effects). PEM reporting: the percentage of participants with PEM was not reported. Very serious population indirectness – 1994 CDC criteria used; PEM is not a compulsory feature and all participants had rhinitis [original analysis]; percentage of participants with PEM unclear & all participants had rhinitis [PEM reanalysis]

Study	Intervention and comparison	Population	Outcomes	Comments
Mckenzie 1998[47]	Oral corticosteroids – Hydrocortisone Hydrocortisone pills, dose of 16 mg per square metre of body surface per day (20-30mg every morning at 8am and 5mg every day at 2pm). Duration: 12 weeks. Versus Placebo Identical placebo at same doses as hydrocortisone group. Duration: 12 weeks.	N=70 people with CFS, diagnosed according to the CDC criteria 1988 (Holmes 1988). CDC criteria 1994 (Fukuda 1994) were also met. Strata details: adults (18-65 years); severity mixed or unclear.	Psychological status (Beck Depression Inventory; Profile of Mood States seven subscales; Symptom checklist 90-R three subscales; Hamilton Depression Scale) Adverse events (adverse reaction) Activity levels (activity scale) Symptom scales (Wellness Scale & Sickness Impact Profile)	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported.Serious population indirectness – Holmes 1988 and 1994 CDC criteria used; PEM is not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Peterson 1998[64]	Oral corticosteroids – Fludrocortisone Initial dose of fludrocortisone acetate was 0.1mg via 1 tablet orally. Dose doubled if no AEs reported after 2 weeks of treatment. Duration: 6 weeks. Versus Placebo Identical tablets taken at same dosing regimen as study drug. Duration: 6 weeks. Patients told not to make any dietary changes (including salt intake) during study.	N=25 people with CFS, diagnosed according to the CDC criteria 1988 (Holmes 1988) and CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-65 years); severity mixed or unclear.	Quality of life (SF36) Cognitive function (inability to concentrate, forgetfulness and confusion all on VAS; reaction time) Psychological status (positive and negative effects scale, PANAS)	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – Holmes 1988 and 1994 CDC criteria used; PEM is not a compulsory feature

Study	Intervention and comparison	Population	Outcomes	Comments
			Pain (muscle pain and joint pain on VAS) Sleep quality (unrefreshing sleep on VAS) Adverse events (adverse events, adverse events causing withdrawal from the study) Activity levels (distance until exhausted) Exercise performance measures (time on treadmill) Symptom scales (headaches, painful lymph nodes and sore throat on VAS)	[original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Rowe 2001[77]	<u>Oral corticosteroids – Fludrocortisone</u> Fludrocortisone starting at a dose of 0.025 mg/day (1 capsule) for a week, then 0.05 mg/day (2 capsules) for the following week, and eventually increased to 0.1 mg/day (4	N=70 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-50 years); severity mixed or unclear.	Fatigue (Wood Mental Fatigue Inventory; POMS vigour and fatigue subscales)	Conducted in USA. PEM reporting: the percentage of participants with PEM was not

Study	Intervention and comparison	Population	Outcomes	Comments
	capsules) for remaining 7 weeks. Total duration: 9 weeks. Versus <u>Placebo</u> Identical capsules containing methylcellulose only were given exactly as the study drug in the same dose increments. Total duration: 9 weeks. Patients advised to drink at least 2L of fluid per day and to keep normal NaCl intake to their usual levels. Both groups also had KCl tablets 10mEq/day for duration of treatment. If AEs emerged, doses were reduced to previously tolerated levels.	All participants also had neurally-mediated hypotension.	Physical functioning (SF-36 physical function subscale) Psychological status (Beck Depression Inventory & SF-36 mental health subscale) Adverse events (adverse effects) Activity levels (Duke Activity Status) Symptom scales (Wellness Score)	reported.Very serious population indirectness –1994 CDC criteria used; PEM is not a compulsory feature and all participants also had neurally-mediated hypotension [original analysis]; percentage of participants with PEM unclear and all participants had neurally-mediated hypotension [PEM reanalysis]
Antihypertensive drugs				
Morriss 2002[56]	<u>Sympathomimetic/central antihypertensive drugs – Clonidine</u> Clonidine IV infusion 2.5 micro-g/kg in 10ml normal saline over 5 minutes. One-off treatment. Versus <u>Placebo</u> IV infusion of 10ml normal saline over 5 mins. One-off treatment. Heparinised cannula used for infusion.	N=10 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994) Strata details: adults (18-60 years); severity mixed or unclear.	Cognitive function (13 tests/performance tasks)	Conducted in United Kingdom Crossover: randomised order for clonidine/placebo with washout of 2 weeks. PEM reporting: the percentage of participants with

Study	Intervention and comparison	Population	Outcomes	Comments
				PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Sulheim 2014[91]	<u>Sympathomimetic/central antihypertensive drugs – Clonidine</u> Tablets containing 25µg of clonidine hydrochloride were enclosed in orange opaque, demolition-resistant lactose capsules. Clonidine dosages were chosen to yield plasma concentrations within the lower range of what is considered clinically effective. One-half of the dose was given during the first 3 days to minimize introductory adverse effects. After 8 weeks of the full dose, the dose was halved for 1 additional week to avoid rebound effects, after which treatment was discontinued. Versus	N=120 CFS patients (excluding 68 healthy controls), diagnosed according to the following: 3 months of unexplained, disabling chronic/relapsing fatigue of new onset, and no accompanying symptoms. Strata details: young people (12-18 years); severity mixed or unclear.	Fatigue (Chalder Fatigue Questionnaire) Physical functioning (functional disability inventory) Cognitive function (digit span backward test total) Pain (BPI average pain score) Sleep quality (KSQ insomnia score)	Conducted in Norway PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – PEM not compulsory feature of diagnosis [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
	Placebo Empty capsules were used as placebo comparators. Duration: 8 weeks.		Adverse events (self-reported) Activity levels (steps per day measured by accelerometer) Symptom scales (CFS symptom inventory hypersensitivity score)	
Central nervous system stimulants				
Blockmans 2006[9]	Sympathomimetic/central antihypertensive drugs – Methylphenidate 10 mg twice daily (8am and 2pm). Taken for 1 month. Vs Placebo Crossover: The same 60 patients took both drugs, but in a random order. Thus about half would have had the study drug in the first period, whilst the other half would have had the placebo first. A washout period of 1 week was used before each patient took the alternative compound in the second period of 4 weeks. Patients who stopped the treatment during the first period but who returned after	N=60 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-50 years); severity mixed or unclear.	Quality of life (SF-36 physical and mental subscales) Fatigue (CIS fatigue total score) Psychological status (HADS depression and HADS anxiety) Adverse events (six categories)	Conducted in Belgium. Crossover: 1 week (half- life of drug = 2 hours, so likely to be appropriate). PEM reporting: the percentage of participants with PEM was not reported.Serious population indirectness – 1994 CDC criteria used; PEM not a

Study	Intervention and comparison	Population	Outcomes	Comments
	4 weeks were allowed to start therapy with the second compound			compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Montoya 2018[52]	<u>Sympathomimetic/central antihypertensive drugs – Methylphenidate</u> 5mg methylphenidate daily for week 1 and 10mg twice daily for weeks 2 to 12. Mitochondrial modulator (nutritional supplement) given as 4 tablets twice daily. The combination of these two agents is called KPAX002. KPAX002 is comprised of a low dosage of methylphenidate hydrochloride, combined with nutrients believed to modulate mitochondrial function. Duration: 12 weeks Versus <u>Placebo</u> Placebo version of KPAX002 treatment. Duration: 12 weeks	N=135 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults; severity mixed or unclear.	Fatigue (CIS fatigue total score; fatigue on VAS) Cognitive function (concentration disturbance on VAS) Adverse events (AEs leading to discontinuation; serious AEs - pyelonephritis)	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Olson 2003[61]	<u>Amphetamines – Dexamphetamine</u> Dexamphetamine 5mg twice daily for first week. Dose increased to 10mg twice daily if indicated at start of 2nd week. Increment repeated if appropriate at start of 3rd week.	N=20 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (17-72); severity mixed or unclear.	Quality of life (SF36 physical and mental) Fatigue (Severity Scale)	Conducted in Australia. PEM reporting: the percentage of participants with

Study	Intervention and comparison	Population	Outcomes	Comments
	This dose continued for a further 4 weeks. Duration 6 weeks. Versus Placebo Identical doses and strategies for increase as study drug. Duration 6 weeks		Sleep quality (sleep latency) Adverse events (anorexia)	PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Young 2013[107]	Amphetamines - Lisdexamphetamine. Lisdexamfetamine given as a flexible morning dose (progressing from 30, through 50, and then to 70 mg/day) provided no serious AEs occurred. Duration 6 weeks Versus Placebo Identical placebo given in same incremental doses. Duration 6 weeks.	N=30 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-60); severity mixed or unclear. Note: most participants have executive functioning impairment. Not downgraded for indirectness.	Fatigue (Fatigue Severity Scale) Cognitive function (Behaviour Rating Inventory of Executive Function, BRIEF) Psychological status (Hamilton Anxiety Scale) Pain (McGill Pain Questionnaire) Symptom scales (Clinical Global	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].

Study	Intervention and comparison	Population	Outcomes	Comments
			Improvement, severity) Adverse events (headache, dry mouth, insomnia; discontinuation due to adverse events)	
Randall 2005[68]	<u>Modafinil – two dose arms</u> Modafinil (200mg) Modafinil (400mg) dose increased slowly at 3 day intervals starting at 100mg until required dose reached. Duration: 20 days. The two dose arms were pooled for analysis. Versus <u>Placebo</u> Identical doses of placebo. Duration 20 days	N=14 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-70 years); severity mixed or unclear.	Quality of life (SF36) Fatigue (Chalder physical and mental fatigue scales) Adverse events	Conducted in United Kingdom. Two intervention arms at different dose – pooled for analysis. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Antiviral drugs				

Study	Intervention and comparison	Population	Outcomes	Comments
Montoya 2013[53]	Antiviral drug – Acyclovir Valganciclovir 900 mg (two 450 mg tablets) twice daily for 21 days followed by 900 mg once daily to complete 6 months. Versus Placebo Identical appearing placebo 900 mg (two 450 mg tablets) twice daily for 21 days followed by 900 mg once daily to complete 6 months	N=30 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18+ years); severity mixed or unclear. Inclusion criteria included suspected viral onset and elevated antibody titers. Antibody titers were required to fit one of the following schema: (i) HHV-6 IgG ≥ 1:640, EBV VCA IgG ≥ 1:640, and EBV EA IgG ≥ 1:160 or (ii) HHV-6 IgG ≥ 1:320, EBV VCA IgG ≥ 1:1,280 and EBV EA IgG ≥ 1:160.	Fatigue (multidimensional fatigue inventory, MFI-20) Adverse events (treatment-related adverse events)	Conducted in USA. Other outcomes reported but insufficient information for analysis: general symptom scores, sleep, psychological status, cognitive function PEM reporting: 96.7% of participants had PEM. Very serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature and requirement for participants to have suspected viral onset and elevated antibody titers [original analysis]; serious population indirectness – requirement for participants to have

Study	Intervention and comparison	Population	Outcomes	Comments
				suspected viral onset and elevated antibody titers [PEM reanalysis].
Straus 1988[86]	Antiviral drugs – IV Acyclovir IV acyclovir (500mg per square metre) infused over a period of 60 minutes in 150ml of saline every 8 hrs for 7 days of hospitalisation Versus Placebo IV placebo (500mg per square metre) infused over a period of 60 minutes in 150ml of saline every 8 hrs for 7 days of hospitalisation.	N=27 people with CFS, diagnosed according to CDC criteria 1988 (Holmes 1988) Strata details: adults (mean age 34.1 years); severity mixed or unclear.	Psychological status (Profile of Mood States – 6 subscales) Adverse events (reversible renal failure) Activity levels (rest in hours/day) Symptom scales (Wellness score)	Conducted in USA. PEM reporting: the percentage of participants with PEM was not reported.Serious population indirectness – Holmes 1988 criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
5-HT3 antagonists				
The 2010[93]	5-HT3 antagonists – Ondansetron Ondansetron (8 mg tablets). 16mg per day in 2 doses. Duration 10 weeks. Versus	N=67 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (range of mean age – 34.7 to 35.8 years); severity mixed or unclear.	Fatigue (Checklist Individual Strength (CIS) fatigue) Adverse events (constipation & malaise)	Conducted in Netherlands. PEM reporting: the percentage of participants with PEM was not reported.Serious

Study	Intervention and comparison	Population	Outcomes	Comments
	Placebo Identical placebo - 2 tablets taken per day. Duration 10 weeks		Activity levels (Actometer) Adverse events Symptom scales (Sickness Impact Profile)	population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Galantamine hydrobromide				
Blacker 2004[8]	Galantamine hydrobromide Galantamine hydrobromide, 3 x 2.5mg per day or 3 x 5mg per day or 3 x 7.5mg per day or 3 x 10mg per day. Duration: 16 weeks. Versus Placebo Placebo 3 x daily. Titration details not clear. Duration 16 weeks	N=434 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata criteria: adults (18-65 years); severity mixed or unclear.	Symptom scales (Clinical Global Impression Scores – no change or worse)	Conducted in USA. Other outcomes reported but insufficient information for analysis: fatigue, cognitive function, and sleep quality – results reported narratively in clinical evidence table. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994

Study	Intervention and comparison	Population	Outcomes	Comments
				CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Snorrason 1996[79]	Galantamine hydrobromide Galantamine hydrobromide 10mg 3x daily, reached by a schedule of escalating dosage. Duration 8 weeks. Versus Placebo Placebo 3 x daily. Duration 8 weeks.	N=49 people with CFS, not diagnosed according to a consensus-based set of criteria. Strata criteria: adults (range of mean ages 43.44 to 44.46 years); severity mixed or unclear	Fatigue (on VAS) Cognitive function (memory on VAS) Pain (myalgia on VAS) Sleep quality (sleep disturbance on VAS) Return to school/work (work capacity/ satisfaction on VAS) Adverse events	Conducted in Iceland. In placebo group an optional cross-over design was added to parallel group RCT design - patients could cross-over after 2 weeks if failed to improve or had symptoms worsening.PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – unclear criteria for diagnosis [original analysis]; percentage of

Study	Intervention and comparison	Population	Outcomes	Comments
				participants with PEM unclear [PEM reanalysis].
Antihistamines				
Steinberg 1996[84]	Antihistamines - Terfenadine. 60mg terfenadine twice daily. Duration 2 months Versus Placebo Placebo twice daily. Duration 2 months. Participants were permitted to take oral contraceptives, antibiotics, vitamins, aspirin, NSAIDS, beta adrenergic agonists; not permitted to take antihistamines, decongestants, tricyclic antidepressants or ENT anti-inflammatory agents.	N=30 people with CFS, diagnosed according to CDC criteria 1988 (Holmes 1988). Strata details: adults (mean age 36.2 years); severity mixed or unclear.	Physical functioning (modified Medical Outcome Study Short Form – physical functioning) Psychological status (modified Medical Outcome Study Short Form – mental health)	Conducted in USA. PEM reporting: 82.1% of participants had “post-exertional fatigue (prolonged)” but as there was no additional information the committee was not satisfied that this symptom was PEM. Serious population indirectness – Holmes 1988 criteria used; PEM not a compulsory feature [original analysis]; unclear if symptom described in study is PEM and <95% had this symptom [PEM reanalysis].
Pro-inflammatory cytokine antagonists				

Study	Intervention and comparison	Population	Outcomes	Comments
Roerink 2017[70]	Pro-inflammatory cytokine antagonists - Anakinra. Anakinra (Interleukin-1 receptor antagonist) 100mg subcutaneously per day. Each participant provided with a box containing 28 syringes and supplies of drug. Patients instructed by physician on how to administer. Daily alarm used to assist compliance, along with adherence monitoring. Duration: 4 weeks Versus Placebo Identical placebo given in identical doses intramuscularly. Duration: 4 weeks	N=50 people with CFS, diagnosed according to CDC criteria 1994 (Fukuda 1994). Strata details: adults (18-59 years); severity mixed or unclear.	Mortality Fatigue (CIS fatigue) Physical functioning (SF36 physical function) Psychological status (Symptom Checklist 90) Pain (VAS maximum pain score) Adverse events (AEs & withdrawal due to AEs) Symptom scales (Sickness Impact Profile)	Conducted in Netherlands. PEM reporting: the percentage of participants with PEM was not reported. Serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis].
Staphylococcus vaccine				
Zachrisson 2002[109]	Staphylococcus toxoid preparation, Staphypan Berna (SB). Composed of undefined extracts of 2 strains of staphylococci (S. aureus and S. epidermidis), and a preservative compound thiomersal. Injection given subcutaneously in gluteal region by nurse. Drug administered in increasing doses of 0.1ml, 0.2ml, 0.3ml, 0.4ml, 0.6ml, 0.8ml, 0.9ml, and 1.0ml weekly, followed by booster doses of 1.0ml every 4	N=100 people with CFS (and fibromyalgia), diagnosed according to CDC criteria 1994 (Fukuda 1994) (and ACR criteria for FM). Investigations prior to study entry included physical examination, vital signs and blood parameters.	Pain (Visual analogue of pain scale) General symptom scales (Clinical global assessment of change – observer rated; clinical global	Conducted in Sweden. Comprehensive Psychopathological Rating Scale (CPRS-15) - authors selected 15 items relevant to FM/CFS from original 65-item

Study	Intervention and comparison	Population	Outcomes	Comments
	weeks with final dose given week 24. Drug kept in 1ml ampoules and packed in boxes marked with patient numbers. Versus Sterile water. Injection given subcutaneously in gluteal region by nurse. Administered in increasing doses of 0.1ml, 0.2ml, 0.3ml, 0.4ml, 0.6ml, 0.8ml, 0.9ml, and 1.0ml weekly, followed by booster doses of 1.0ml every 4 weeks with final dose given week 24. Drug kept in 1ml ampoules and packed in boxes marked with patient numbers. Similar in colour to active treatment. Both active substance and placebo caused slight local pain and reaction after injection.	Strata details: adults (age 18-65); severity mixed or unclear (according to global assessment of illness severity measured at baseline 17% were moderately ill, 70% markedly ill, 12% severely ill, 1% most extremely ill)	assessment of severity – observer rated) Adverse events (most frequent side effects; clinical global assessment of side effects)	scale which covers a broad range of psychiatric illnesses. Does not seem to be validated subscale. PEM reporting: the percentage of participants with PEM was not reported. Very serious population indirectness – 1994 CDC criteria used; PEM not a compulsory feature and all participants also had fibromyalgia [original analysis]; percentage of participants with PEM unclear and all participants also had fibromyalgia [PEM reanalysis].

1.1.6. Quality assessment of clinical studies included in the evidence review

Table 3: Clinical evidence summary: Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
Quality of Life: SF36 physical composite (max % change from baseline) Scale from: 0 to 100.	28 (1 study) 10 months	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		The mean quality of life: sf36 physical composite (max % change from baseline) at 10 months in the control groups was 26	The mean quality of life: sf36 physical composite (max % change from baseline) at 10 months in the intervention group (rituximab) was 28 higher (1.56 to 54.44 higher)
Quality of Life: SF36 mental composite (max % change from baseline) Scale from: 0 to 100.	28 (1 study) 10 months	⨁⨀⨀⨀ VERY LOW1,2 due to indirectness, imprecision		The mean quality of life: sf36 mental composite (max % change from baseline) at 10 months in the control groups was 5	The mean quality of life: sf36 mental composite (max % change from baseline) at 10 months in the intervention group (rituximab) was 4 higher (29.52 lower to 37.52 higher)
Fatigue/fatigability: Fatigue severity scale Scale from: 9 to 63.	151 (1 study) 18 months	⨁⨁⨁⨁ HIGH		The mean fatigue/fatigability: fatigue severity scale in the control groups was 56.05	The mean fatigue/fatigability: fatigue severity scale in the intervention group (rituximab) was 0.07 lower (3.21 lower to 3.07 higher)
Fatigue/fatigability: numeric rating scale Scale from: 0 to 10.	151 (1 study) 16-20 months	⨁⨁⨁⨁ HIGH		The mean fatigue/fatigability: numeric rating scale in the control groups was 3.18	The mean fatigue/fatigability: numeric rating scale in the intervention group (rituximab) was 0.06 lower (0.5 lower to 0.39 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
Psychological status: Hamilton Depression Scale Scale from: 0 to 52.	49 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		The mean psychological status: hamilton depression scale at 6 months in the control groups was 10	The mean psychological status: hamilton depression scale at 6 months in the intervention groups (IV immunoglobulin G) was 1 lower (3.35 lower to 1.35 higher)
Psychological status: Zung Self-Rating Depression Scale Scale from: 0 to 80.	49 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		The mean psychological status: zung self-rating depression scale at 6 months in the control groups was 40	The mean psychological status: zung self-rating depression scale at 6 months in the intervention group (IV immunoglobulin G) was 1 higher (5.44 lower to 7.44 higher)
Psychological status: mental health on the Medical Outcome Study Short Form Scale from: 0 to 100.	28 (1 study) 150 days	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		The mean psychological status: mental health on the medical outcome study short form at 150 days in the control groups was 62.9	The mean psychological status: mental health on the medical outcome study short form at 150 days in the intervention group (IV immunoglobulin G) was 4.6 lower (16.07 lower to 6.87 higher)
Physical functioning: physical functioning on the Medical Outcome Study Short Form/SF36 Scale from: 0 to 100.	28 (1 study) 150 days	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		The mean physical functioning: physical functioning on the medical outcome study short form/sf36 in the control groups was 51.8	The mean physical functioning: physical functioning on the medical outcome study short form/sf36 in the intervention groups (IV immunoglobulin G) was 4.2 higher (12.62 lower to 21.02 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
Physical functioning: physical functioning on the Medical Outcome Study Short Form/SF36 Scale from: 0 to 100.	151 (1 study) 24 months	⨁⨁⨁⨁ HIGH			The mean physical functioning: physical functioning on the medical outcome study short form/sf36 in the intervention groups (rituximab) was 1.24 higher (7.38 lower to 9.86 higher)
Physical functioning: functional level percentage	151 (1 study) 16-20 months	⨁⨁⨁⊖ MODERATE2 due to imprecision		The mean physical functioning: functional level percentage in the control groups was 25.93	The mean physical functioning: functional level percentage in the intervention group (rituximab) was 0.68 lower (5.9 lower to 4.54 higher)
Adverse events: Serious Adverse Events with possible/probable relation to intervention	234 (1 study) 42 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 0.5 (0.05 to 5.44)	17 per 1000	9 fewer per 1000 (from 16 fewer to 76 more) (with rintatolimod)
Adverse events: major adverse events	30 (1 study) 21 weeks	⨁⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, indirectness, imprecision	RR 1 (0.24 to 4.18)	200 per 1000	0 fewer per 1000 (from 152 fewer to 636 more) (with IV immunoglobulin G)
Adverse events: constitutional symptoms	99 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3	RR 0.87 (0.72	885 per 1000	115 fewer per 1000 (from 248 fewer to 44 more) (with IV immunoglobulin G)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
		due to risk of bias, indirectness, imprecision	to 1.05)		
Adverse events: any serious adverse events (hospitalisations) with possible/probable relation to intervention	151 (1 study) 24 months	⊕⊕⊕⊕ HIGH	Peto OR 7.82 (1.89 to 32.35)	0 per 1000	100 more per 1000 (from 30 more to 180 more) (with rituximab)
Adverse events: any adverse events of at least moderate severity (CTCAE grade ≥2) with possible/probable relation to intervention	151 (1 study) 24 months	⊕⊕⊕⊖ MODERATE2 due to imprecision	RR 2.08 (1.14 to 3.81)	162 per 1000	175 more per 1000 (from 23 more to 456 more) (with rituximab)
Adverse events: suspected unexpected adverse reactions	151 (1 study) 24 months	⊕⊕⊖⊖ LOW2 due to imprecision	RR 1.92 (0.18 to 20.75)	14 per 1000	12 more per 1000 (from 11 fewer to 267 more) (with rituximab)
Activity levels: mean number of steps per 24 hours	151 (1 study) 17-21 months	⊕⊕⊕⊕ HIGH		The mean activity levels: mean number of steps per 24 hours in the control groups was 3904	The mean activity levels: mean number of steps per 24 hours in the intervention group (rituximab) was 127 lower (1004 lower to 750 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
Exercise performance measure: Treadmill exercise duration in seconds	208 (1 study) 42 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean exercise performance measure: treadmill exercise duration in seconds at 42 weeks in the control groups was 616	The mean exercise performance measure: treadmill exercise duration in seconds at 42 weeks in the intervention group (rintatolimod) was 56 higher (25.94 lower to 137.94 higher)
Return to school or work: Resumption of pre-morbid employment status (full-time)	49 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness	Peto OR 10.79 (1.98 to 58.68)	0 per 1000	260 more per 1000 (from 80 more to 450 more) (with IV immunoglobulin G)
Symptom scales: Marked reduction in symptoms and improvement in functional capacity	49 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 3.77 (1.18 to 12.04)	115 per 1000	320 more per 1000 (from 21 more to 1000 more) (with IV immunoglobulin G)

1 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; no clear definition of "post-exertional fatigue (prolonged) to confirm this is PEM, so unclear how many participants had PEM [PEM reanalysis – see Appendix G for additional details]

2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

3 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

4 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. Further downgraded for outcome indirectness (unclear if major adverse events were treatment-related) (Peterson 1990)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo (95% CI)
* Studies included for immunomodulatory drugs: Rituximab – Fluge 2011, Fluge 2019; IVIG – Lloyd 1990, Peterson 1990, Vollmer-Conna 1997; Rintatolimod – Strayer 2012					

Table 4: Clinical evidence summary: Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Quality of Life: SF36 vitality Scale from: 0 to 100.	46 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 vitality at 12 weeks in the control group was 11.9	The mean quality of life: sf36 vitality at 12 weeks in the intervention group (duloxetine) was 3.3 higher (10.3 lower to 16.9 higher)
Quality of Life: SF-36 physical functioning Scale from: 0 to 100.	46 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 physical functioning at 12 weeks in the control group was 7.5	The mean quality of life: sf-36 physical functioning at 12 weeks in the intervention group (duloxetine) was 6.8 higher (8.5 lower to 22.1 higher)
Quality of Life: SF-36 role physical Scale from: 0 to 100.	46 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of		The mean change score in quality of life: SF36 role physical at 12 weeks in the control group was	The mean quality of life: sf-36 role physical at 12 weeks in the intervention group (duloxetine) was

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
		bias, indirectness, imprecision		11.5	11 higher (9 lower to 31 higher)
Quality of Life: SF36 mental health Scale from: 0 to 100.	46 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 mental health at 12 weeks in the control group was 7.5	The mean quality of life: sf36 mental health at 12 weeks in the intervention group (duloxetine) was 1.1 lower (11.8 lower to 9.6 higher)
Quality of Life: SF36 role emotional Scale from: 0 to 100.	46 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 role emotional at 12 weeks in the control group was 9	The mean quality of life: sf36 role emotional at 12 weeks in the intervention group (duloxetine) was 4.4 higher (24.2 lower to 33 higher)
Quality of Life: SF36 bodily pain Scale from: 0 to 100.	46 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 bodily pain at 12 weeks in the control group was 7.5	The mean quality of life: sf36 bodily pain at 12 weeks in the intervention group (duloxetine) was 11.4 higher (0.5 lower to 23.3 higher)
Quality of Life: SF36 general health Scale from: 0 to 100.	46 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean change score in quality of life: SF36 general health at 12 weeks in the control group was 2.7	The mean quality of life: sf36 general health at 12 weeks in the intervention group (duloxetine) was 0 higher (10.8 lower to 10.8 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
		indirectness, imprecision			
Quality of Life: SF36 social functioning Scale from: 0 to 100.	46 (1 study) 12 weeks	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in quality of life: SF36 social functioning at 12 weeks in the control group was 10.6	The mean quality of life: sf36 social functioning at 12 weeks in the intervention group (duloxetine) was 0.7 higher (14.7 lower to 16.1 higher)
Fatigue: 14-item Chalder fatigue scale Scale: not reported.	69 (1 study) 12 weeks	⨁⨀⨀⨀ VERY LOW2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the control groups was -2.7	The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the intervention group (fluoxetine) was 0.3 lower (4.06 lower to 3.46 higher)
Fatigue: MFI-20 general fatigue Scale: not reported.	57 (1 study) 12 weeks	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in fatigue: MFI-20 general fatigue at 12 weeks in the control group was -1.8	The mean fatigue: mfi-20 general fatigue at 12 weeks in the intervention group (duloxetine) was 1 lower (2.8 lower to 0.8 higher)
Fatigue: MFI-20 physical fatigue Scale: not reported.	57 (1 study) 12 weeks	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in fatigue: MFI-20 physical fatigue at 12 weeks in the control group was -1	The mean fatigue: mfi-20 physical fatigue at 12 weeks in the intervention group (duloxetine) was 0.9 lower (2.7 lower to 0.9 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Fatigue: MFI-20 reduced activity Scale: not reported.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in fatigue: MFI-20 reduced activity at 12 weeks in the control group was -1.5	The mean fatigue: mfi-20 reduced activity at 12 weeks in the intervention group (duloxetine) was 0 higher (1.8 lower to 1.8 higher)
Fatigue: MFI-20 reduced motivation Scale: not reported.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in fatigue: MFI-20 reduced motivation at 12 weeks in the control group was -1.6	The mean fatigue: mfi-20 reduced motivation at 12 weeks in the intervention group (duloxetine) was 0.8 lower (2.6 lower to 1 higher)
Fatigue: MFI-20 mental fatigue Scale: not reported.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in fatigue: MFI-20 mental fatigue at 12 weeks in the control group was -1.4	The mean fatigue: mfi-20 mental fatigue at 12 weeks in the intervention group (duloxetine) was 2.5 lower (4.4 to 0.6 lower)
Fatigue: Checklist Individual Strength (CIS) fatigue Scale from: 8 to 56.	97 (1 study) 16 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: checklist individual strength (cis) fatigue at 16 weeks in the control group was not reported (between-group difference only)	The mean fatigue: checklist individual strength (CIS) fatigue at 16 weeks in the intervention group (fluoxetine) was 0.16 lower (0.64 lower to 0.31 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Physical functioning: Karnofsky Performance Index (measured in units of standard deviation at baseline)	77 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean physical functioning: karnofsky performance index at 6 weeks in the control groups was 0.58	The mean physical functioning: karnofsky performance index at 6 weeks in the intervention group (moclobemide) was 0.28 higher (0.28 lower to 0.84 higher)
Psychological status: Profile of mood states (POMS) fatigue Scale from: 0 to 28.	77 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: profile of mood states (poms) fatigue at 6 weeks in the control groups was -0.01	The mean psychological status: profile of mood states (poms) fatigue at 6 weeks in the intervention group (moclobemide) was 0.04 lower (0.2 lower to 0.12 higher)
Psychological status: Profile of mood states (POMS) vigour Scale from: 0 to 32.	77 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: profile of mood states (poms) vigour at 6 weeks in the control groups was 0	The mean psychological status: profile of mood states (poms) vigour at 6 weeks in the intervention group (moclobemide) was 0.51 higher (0 to 1.02 higher)
Psychological status: Profile of mood states (POMS) depression Scale from: 0 to 60.	77 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: profile of mood states (poms) depression at 6 weeks in the control groups was -0.08	The mean psychological status: profile of mood states (poms) depression at 6 weeks in the intervention group (moclobemide) was 0.02 higher (0.36 lower to 0.4 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Psychological status: HADS depression Scale from: 0 to 21.	126 (2 studies) 12-26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, indirectness, imprecision, inconsistency		The mean change in psychological status: hads depression at 12-26 weeks in the control groups was -1.6	The mean change in psychological status: hads depression at 12-26 weeks in the intervention groups (fluoxetine or duloxetine) was 0.51 higher (0.72 lower to 1.74 higher)
Psychological status: HADS anxiety Scale from: 0 to 21.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in psychological status: HADS anxiety at 12 weeks in the control group was -2	The mean psychological status: hads anxiety at 12 weeks in the intervention group (duloxetine) was 0.9 lower (2.4 lower to 0.6 higher)
Psychological status: Beck Depression Inventory Scale from: 0 to 63.	97 (1 study) 16 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: beck depression inventory at 16 weeks in the control group was not reported (between-group difference only)	The mean psychological status: beck depression inventory at 16 weeks in the intervention group (fluoxetine) was 0.19 lower (0.35 to 0.02 lower)
Pain: Brief Pain Inventory severity Scale from: 0 to 10.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in pain: Brief Pain Inventory severity at 12 weeks in the control group was -0.8 not reported (between-group difference only)	The mean pain: brief pain inventory severity at 12 weeks in the intervention group (duloxetine) was 0.73 lower (1 to 0.46 lower)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Pain: Brief Pain Inventory interference Scale from: 0 to 10.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, indirectness		The mean change score in pain: Brief Pain Inventory interference at 12 weeks in the control group was -1.1	The mean pain: brief pain inventory interference at 12 weeks in the intervention group (duloxetine) was 0.7 lower (0.96 to 0.44 lower)
Adverse events: tremor	96 (1 study) 16 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RR 1.57 (0.87 to 2.83)	255 per 1000	145 more per 1000 (from 33 fewer to 466 more) (with fluoxetine)
Adverse events: perspiration	96 (1 study) 16 weeks	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, indirectness	RR 1.7 (1.14 to 2.53)	392 per 1000	275 more per 1000 (from 55 more to 600 more) (with fluoxetine)
Exercise performance measure: VO2 max (mL O2/kg/min)	69 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean change score in exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the control groups was -0.1	The mean exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the intervention group (fluoxetine) was 1.1 higher (1.43 lower to 3.63 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
Symptom scales: Clinical Global Impression of Severity Scale from: 1 to 7.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in symptom scale: Clinical Global Impression of Severity at 12 weeks in the control group was -0.4	The mean symptom scales: clinical global impression of severity at 12 weeks in the intervention group (duloxetine) was 0.1 lower (0.3 lower to 0.1 higher)
Symptom scales: Clinical Global Impression of Improvement Scale from: 1 to 7.	57 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW2,3 due to risk of bias, indirectness, imprecision		The mean change score in symptom scale: Clinical Global Impression of Improvement at 12 weeks in the control group was 3.3	The mean symptom scales: clinical global impression of improvement at 12 weeks in the intervention group (duloxetine) was 0.8 lower (1.7 lower to 0.1 higher)
Symptom scales: CDC symptom inventory Scale from: not reported.	46 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW2,3 due to risk of bias, indirectness, imprecision		The mean change score in symptom scale: CDC symptom inventory at 12 weeks in the control group was -13	The mean symptom scales: cdc symptom inventory at 12 weeks in the intervention group (duloxetine) was 2.7 lower (15.5 lower to 10.1 higher)
Symptom scales: Improvement of symptoms (patient-reported)	186 (2 studies) 6-16 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.63 (1.02 to 2.59)	202 per 1000	127 more per 1000 (from 4 more to 321 more) (with fluoxetine or moclobemide)
Symptom scales: Worsening of symptoms (patient-reported)	96 (1 study) 16 weeks	⨁⊖⊖⊖ VERY LOW1,2,3	RR 0.66 (0.28	235 per 1000	80 fewer per 1000 (from 169 fewer to 125 more) (with fluoxetine)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo (95% CI)
		due to risk of bias, indirectness, imprecision	to 1.53)		

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
4 Downgraded for inconsistency. I^2=63%

* Studies included for antidepressants: Duloxetine – Arnold 2015; Moclobemide – Hickie 2000; Fluoxetine – Vercoulen 1996, Wearden 1998

Table 5: Clinical evidence summary: Antidepressants (fluoxetine) versus graded exercise for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (fluoxetine) versus graded exercise (95% CI)
Fatigue: 14-item Chalder fatigue scale Scale from: not reported.	69 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the control groups was -5.7	The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the intervention group was 2.7 higher (1.85 lower to 7.25 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (fluoxetine) versus graded exercise (95% CI)
Psychological status: HADS depression Scale from: 0 to 21.	69 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads depression at 26 weeks in the control groups was -1.2	The mean psychological status: hads depression at 26 weeks in the intervention group was 0.5 lower (2.27 lower to 1.27 higher)
Exercise performance measure: VO2 max (mL O2/kg/min)	69 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the control groups was 2.8	The mean exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the intervention group was 1.8 lower (4.53 lower to 0.93 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
Studies included: Wearden 1998					

Table 6: Clinical evidence summary: Antidepressants (fluoxetine) versus combined antidepressants (fluoxetine) & graded exercise for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (fluoxetine) versus combined antidepressants (fluoxetine) & graded exercise (95% CI)
Fatigue: 14-item Chalder fatigue scale Scale from: not reported.	68 (1 study) 26 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the control groups was -6	The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the intervention group was 3 higher (1.47 lower to 7.47 higher)
Psychological status: HADS depression Scale from: 0 to 21.	69 (1 study) 26 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads depression at 26 weeks in the control groups was -2	The mean psychological status: hads depression at 26 weeks in the intervention group was 0.3 higher (1.51 lower to 2.11 higher)
Exercise performance measure: VO2 max (mL O2/kg/min)	68 (1 study) 26 weeks	⊕⊖⊖⊖ VERY LOW2,3 due to risk of bias, indirectness, imprecision		The mean change score in exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the control groups was 2	The mean exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the intervention group was 1 lower (3.41 lower to 1.41 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
Studies included: Wearden 1998					

Table 7: Clinical evidence summary: Combined antidepressants (fluoxetine) & graded exercise versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Combined antidepressants (fluoxetine) & graded exercise versus placebo (95% CI)
Fatigue: 14-item Chalder fatigue scale Scale from: not reported.	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the control groups was -2.7	The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the intervention group was 3.3 lower (7.71 lower to 1.11 higher)
Psychological status: HADS depression Scale from: 0 to 21.	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads depression at 26 weeks in the control groups was -1.3	The mean psychological status: hads depression at 26 weeks in the intervention group was 0.7 lower (2.28 lower to 0.88 higher)
Exercise performance measure: VO2 max (mL O2/kg/min)	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the control groups was -0.1	The mean exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the intervention group was 2.1 higher (0.08 lower to 4.28 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Studies included: Wearden 1998

Table 8: Clinical evidence summary: Combined antidepressants (fluoxetine) & graded exercise versus graded exercise for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Combined antidepressants (fluoxetine) & graded exercise versus graded exercise (95% CI)
Fatigue: 14-item Chalder fatigue scale Scale from: not reported.	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness		The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the control groups was -5.7	The mean fatigue: 14-item chalder fatigue scale at 26 weeks in the intervention group was 0.3 lower (5.41 lower to 4.81 higher)
Psychological status: HADS depression Scale from: 0 to 21.	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads depression at 26 weeks in the control groups was -1.2	The mean psychological status: hads depression at 26 weeks in the intervention group was 0.8 lower (2.52 lower to 0.92 higher)
Exercise performance measure: VO2 max (mL O2/kg/min)	67 (1 study) 26 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean change score in exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the control groups was 2.8	The mean exercise performance measure: vo2 max (ml o2/kg/min) at 26 weeks in the intervention group was 0.8 lower (3.21 lower to 1.61 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Studies included: Wearden 1998

Table 9: Clinical evidence summary: Antidepressants (fluoxetine) versus antipsychotics (amisulpride) for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (fluoxetine) versus antipsychotics (amisulpride) (other) (95% CI)
Quality of Life: SF12 Scale from: 0 to 100.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness		The mean quality of life: sf12 at 12 weeks in the control groups was 53.2	The mean quality of life: sf12 at 12 weeks in the intervention group was 15.6 lower (18.61 to 12.59 lower)
Fatigue: Fatigue Severity Scale Scale from: 9 to 63.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness		The mean fatigue: fatigue severity scale at 12 weeks in the control groups was 36.3	The mean fatigue: fatigue severity scale at 12 weeks in the intervention group was 12.6 higher (8.26 to 16.94 higher)
Psychological status: HADS anxiety Scale from: 0 to 21.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads anxiety at 12 weeks in the control groups was 4.5	The mean psychological status: hads anxiety at 12 weeks in the intervention group was 0.4 higher (0.22 lower to 1.02 higher)
Psychological status: HADS depression Scale from: 0 to 21.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hads depression at 12 weeks in the control groups was 4.3	The mean psychological status: hads depression at 12 weeks in the intervention group was 0.1 lower (0.69 lower to 0.49 higher)
Pain: pain on VAS Scale from: 0 to 100.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to risk of		The mean pain: pain on vas at 12 weeks in the control groups was 40.5	The mean pain: pain on vas at 12 weeks in the intervention group was 12.6 higher (5.8 to 19.4 to higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressants (fluoxetine) versus antipsychotics (amisulpride) (other) (95% CI)
		bias, indirectness			
Adverse events: FIBSER global burden Scale from: not reported.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean adverse events: fibser global burden at 12 weeks in the control groups was 0.8	The mean adverse events: fibser global burden at 12 weeks in the intervention group was 0.2 lower (0.67 lower to 0.27 higher)
Symptom scales: Clinical Global Impression Severity (CGI-S) Scale from: 1 to 7.	40 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness		The mean symptom scales: clinical global impression severity (cgi-s) at 12 weeks in the control groups was 2.9	The mean symptom scales: clinical global impression severity (cgi-s) at 12 weeks in the intervention group was 1.3 higher (0.75 to 1.85 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
Studies included: Pardini 2011					

Table 10: Clinical evidence summary: Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Quality of Life: SF36 physical total Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 physical total at 6-11 weeks in the control groups was 46.75	The mean quality of life: sf36 physical total at 6-11 weeks in the intervention group (fludrocortisone) was 7.54 higher (0.71 lower to 15.79 higher)
Quality of Life: SF36 energy or fatigue Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 energy or fatigue at 6 weeks in the control groups was 18.2	The mean quality of life: sf36 energy or fatigue at 6 weeks in the intervention group (fludrocortisone) was 2.1 higher (7.43 lower to 11.63 higher)
Quality of Life: SF36 emotional wellbeing Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 emotional wellbeing at 6 weeks in the control groups was 68.8	The mean quality of life: sf36 emotional wellbeing at 6 weeks in the intervention group (fludrocortisone) was 3.8 higher (5.29 lower to 12.89 higher)
Quality of Life: SF36 role emotional Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 role emotional at 6 weeks in the control groups was 87.8	The mean quality of life: sf36 role emotional at 6 weeks in the intervention group (fludrocortisone) was 0 higher (14.96 lower to 14.96 higher)
Quality of Life: SF36 role physical Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean quality of life: sf36 role physical at 6 weeks in the control	The mean quality of life: sf36 role physical at 6 weeks in the intervention group (fludrocortisone)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
		indirectness, imprecision		groups was 25	was 11.8 lower (29.09 lower to 5.49 higher)
Quality of Life: SF36 pain Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 pain at 6 weeks in the control groups was 50.5	The mean quality of life: sf36 pain at 6 weeks in the intervention group (fludrocortisone) was 0.6 lower (15.29 lower to 14.09 higher)
Quality of life: SF36 social Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 social at 6 weeks in the control groups was 38.2	The mean quality of life: sf36 social at 6 weeks in the intervention group (fludrocortisone) was 1.9 higher (11.06 lower to 14.86 higher)
Quality of life: SF36 general well-being Scale from: 0 to 100.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 general wellbeing at 6 weeks in the control groups was 35.8	The mean quality of life: sf36 general wellbeing at 6 weeks in the intervention groups (fludrocortisone) was 3.7 lower (12.54 lower to 5.14 higher)
Fatigue: fatigue on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: fatigue on vas at 6 weeks in the control groups was 7.5	The mean fatigue: fatigue on vas at 6 weeks in the intervention group (fludrocortisone) was 0 higher (1.1 lower to 1.1 higher)
Fatigue: Chronic Fatigue Syndrome Severity Rating	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias,		The mean [nasal] symptom scales: rhinitis severity rating at	The mean symptom scales: rhinitis severity rating at 4-8 weeks in the intervention group (nasal

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Scale from: not reported.	4-8 weeks	indirectness, imprecision		4-8 weeks in the control group was 18.13	flunisolide) was 3.17 lower (7.48 lower to 1.14 higher)
Fatigue: Profile of Mood States – fatigue Scale from: 0 to 28.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, indirectness		The mean fatigue: profile of mood states - fatigue at 11 weeks in the control groups was 16.4	The mean fatigue: profile of mood states - fatigue at 11 weeks in the intervention groups (fludrocortisone) was 0.20 lower (3.47 lower to 3.07 higher)
Fatigue: Profile of Mood States – fatigue Scale from: 0 to 28.	68 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean fatigue: profile of mood states - fatigue at 12 weeks in the control groups was -1.8	The mean fatigue: profile of mood states - fatigue at 12 weeks in the intervention groups (hydrocortisone) was 1.8 lower (4.14 lower to 0.54 higher)
Fatigue: Profile of Mood States – vigour Scale from: 0 to 32.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, indirectness		The mean fatigue: profile of mood states - vigour at 11 weeks in the control groups was 8.6	The mean fatigue: profile of mood states - vigour at 11 weeks in the intervention groups (fludrocortisone) was 0.2 higher (2.56 lower to 2.96 higher)
Fatigue: Profile of Mood States – vigour Scale from: 0 to 32.	68 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean fatigue: profile of mood states - vigour at 12 weeks in the control groups was 3.3	The mean fatigue: profile of mood states - vigour at 12 weeks in the intervention groups (hydrocortisone) was 0.5 higher (1.07 lower to 2.07 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Fatigue: Wood Mental Fatigue Inventory Scale from: 0 to 36.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean fatigue: wood mental fatigue inventory at 11 weeks in the control groups was 13.3	The mean fatigue: wood mental fatigue inventory at 11 weeks in the intervention group (fludrocortisone) was 0.8 higher (3.66 lower to 5.26 higher)
Physical function: SF36 physical function Scale from: 0 to 100.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean physical function: sf36 physical function at 11 weeks in the control groups was 51.4	The mean physical function: sf36 physical function at 11 weeks in the intervention group (fludrocortisone) was 7.5 higher (3.2 lower to 18.2 higher)
Psychological status: SF36 mental health Scale from: 0 to 100.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean psychological status: sf36 mental health at 11 weeks in the control groups was 69.8	The mean psychological status: sf36 mental health at 11 weeks in the intervention group (fludrocortisone) was 1.2 lower (8.92 lower to 6.52 higher)
Adverse events: adverse events leading to study withdrawal	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	Peto OR 0.13 (0.01 to 2.13)	100 per 1000	100 fewer per 1000 (from 250 fewer to 50 more) (with fludrocortisone)
Adverse events: adverse effects/adverse events	123 (2 studies)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, indirectness	RR 0.86 (0.63	554 per 1000	78 fewer per 1000 (from 205 fewer to 94 more) (with fludrocortisone)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
	6-11 weeks		to 1.17)		
Adverse events: any adverse reaction	70 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.15 (0.93 to 1.43)	771 per 1000	116 more per 1000 (from 54 fewer to 332 more) (with hydrocortisone)
Psychological status: Beck Depression Inventory Scale from: 0 to 63.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean psychological status: beck depression inventory at 11 weeks in the control groups was 10.8	The mean psychological status: beck depression inventory at 11 weeks in the intervention groups (fludrocortisone) was 0.4 lower (3.43 lower to 2.63 higher)
Psychological status: Beck Depression Inventory Scale from: 0 to 63.	68 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: beck depression inventory at 12 weeks in the control groups was -0.4	The mean psychological status: beck depression inventory at 12 weeks in the intervention groups (hydrocortisone) was 1.7 lower (3.90 lower to 0.5 higher)
Psychological status: Profile of Mood States – anger Scale from: 0 to 48.	68 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean psychological status: profile of mood states - anger, at 12 weeks in the control groups was -0.8	The mean psychological status: profile of mood states - anger, at 12 weeks in the intervention group (hydrocortisone) was 0.8 lower (2.63 lower to 1.03 higher)
Psychological status: Profile of Mood States – anxiety Scale from: 0 to 36.	68 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean psychological status: profile of mood states - anxiety, at 12 weeks in the control groups	The mean psychological status: profile of mood states - anxiety, at 12 weeks in the intervention group (hydrocortisone) was

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
		indirectness, imprecision		was -2.1	1.3 higher (0.17 lower to 2.77 higher)
Psychological status: Profile of Mood States – confusion Scale from: 0 to 28.	68 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean psychological status: profile of mood states - confusion, at 12 weeks in the control groups was -1.4	The mean psychological status: profile of mood states - confusion, at 12 weeks in the intervention group (hydrocortisone) was 0.3 higher (1.18 lower to 1.78 higher)
Psychological status: Profile of Mood States – depression Scale from: 0 to 60.	68 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		The mean psychological status: profile of mood states - depression, at 12 weeks in the control groups was 0	The mean psychological status: profile of mood states - depression, at 12 weeks in the intervention group (hydrocortisone) was 1.6 lower (3.61 lower to 0.41 higher)
Psychological status: Symptom checklist-90-R general sensitivity index Scale from: not reported.	68 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean psychological status: symptom checklist-90-r general sensitivity index at 12 weeks in the control groups was -0.1	The mean psychological status: symptom checklist-90-r general sensitivity index at 12 weeks in the intervention group (hydrocortisone) was 0 higher (0.1 lower to 0.1 higher)
Psychological status: Symptom checklist-90-R positive symptom distress index Scale from: not reported.	68 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: symptom checklist-90-r positive symptom distress index at 12 weeks in the control groups was -0.1	The mean psychological status: symptom checklist-90-r positive symptom distress index at 12 weeks in the intervention group (hydrocortisone) was 0.1 higher (0.04 lower to 0.24 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Psychological status: Symptom checklist-90-R positive symptom total Scale from: not reported.	68 (1 study) 12 weeks	⨁⨁⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean psychological status: symptom checklist-90-r positive symptom total at 12 weeks in the control groups was -2.4	The mean psychological status: symptom checklist-90-r positive symptom total at 12 weeks in the intervention group (hydrocortisone) was 0.2 lower (5.5 lower to 5.1 higher)
Psychological status: Hamilton Depression Rating Scale Scale from: not reported.	65 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: hamilton depression rating scale at 12 weeks in the control groups was 0.1	The mean psychological status: hamilton depression rating scale at 12 weeks in the intervention group (hydrocortisone) was 0.9 lower (2.55 lower to 0.75 higher)
Psychological status: Positive and negative effect scale (PANAS) positive affect Scale from: 10 to 50.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: positive and negative effect scale (panas) positive affect at 6 weeks in the control groups was 21.7	The mean psychological status: positive and negative effect scale (panas) positive affect at 6 weeks in the intervention group (fludrocortisone) was 1 higher (3.67 lower to 5.67 higher)
Activity levels: activity scale Scale from: not reported.	68 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean activity levels: activity scale at 12 weeks in the control groups was 0.7	The mean activity levels: activity scale at 12 weeks in the intervention group (hydrocortisone) was 0.4 lower (1 lower to 0.2 higher)
Activity levels: distance before exhausted (ordinal scale) Scale from: 1 to 5.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean activity levels: distance before exhausted (ordinal scale) at 6 weeks in the control groups	The mean activity levels: distance before exhausted (ordinal scale) at 6 weeks in the intervention group

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
		indirectness, imprecision		was 2.7	(fludrocortisone) was 0 higher (0.72 lower to 0.72 higher)
Activity levels: Duke Activity Status Index Scale from: 0 to 58.2.	83 (1 study) 11 weeks	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean activity levels: duke activity status index at 11 weeks in the control groups was 6.7	The mean activity levels: duke activity status index at 11 weeks in the intervention group (fludrocortisone) was 2.5 higher (1.49 lower to 6.49 higher)
Cognitive function: Reaction time (secs)	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: reaction time (secs) at 6 weeks in the control groups was 0.36	The mean cognitive function: reaction time (secs) at 6 weeks in the intervention group (fludrocortisone) was 0.01 lower (0.06 lower to 0.04 higher)
Cognitive function: inability to concentrate on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: inability to concentrate on vas at 6 weeks in the control groups was 5.8	The mean cognitive function: inability to concentrate on vas at 6 weeks in the intervention group (fludrocortisone) was 0.6 lower (2.18 lower to 0.98 higher)
Cognitive function: forgetfulness on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: forgetfulness on vas at 6 weeks in the control groups was 5.6	The mean cognitive function: forgetfulness on vas at 6 weeks in the intervention group (fludrocortisone) was 0.9 lower (2.45 lower to 0.65 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Cognitive function: confusion on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: confusion on vas at 6 weeks in the control groups was 4.4	The mean cognitive function: confusion on vas at 6 weeks in the intervention group (fludrocortisone) was 0.1 lower (1.68 lower to 1.48 higher)
Pain: muscle pain on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean pain: muscle pain on vas at 6 weeks in the control groups was 5.9	The mean pain: muscle pain on vas at 6 weeks in the intervention group (fludrocortisone) was 0.1 lower (1.82 lower to 1.62 higher)
Pain: joint pain on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean pain: joint pain on vas at 6 weeks in the control groups was 5.1	The mean pain: joint pain on vas at 6 weeks in the intervention group (fludrocortisone) was 0.3 lower (2.39 lower to 1.79 higher)
Sleep quality: unrefreshing sleep on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean sleep quality: unrefreshing sleep on vas at 6 weeks in the control groups was 8.2	The mean sleep quality: unrefreshing sleep on vas at 6 weeks in the intervention group (fludrocortisone) was 0.5 lower (1.68 lower to 0.68 higher)
Sleep quality: Functional Outcomes of Sleep Questionnaire Scale from: not reported.	28 (1 study) 4-8 weeks	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		The mean [nasal] sleep quality: functional outcomes of sleep questionnaire at 4-8 weeks in the control group was 12.4	The mean [nasal] sleep quality: functional outcomes of sleep questionnaire at 4-8 weeks in the intervention group (nasal flunisolide) was 0.89 higher (0.99 lower to 2.77 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
Sleep quality: Epworth Sleepiness Scale Scale from: 0 to 24.	28 (1 study) 4-8 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean [nasal] sleep quality: epworth sleepiness scale at 4-8 weeks in the control group was 11.66	The mean [nasal] sleep quality: epworth sleepiness scale at 4-8 weeks in the intervention group (nasal flunisolide) was 3.18 lower (6.57 lower to 0.21 higher)
Exercise performance measure: Treadmill time (mins)	40 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean exercise performance measure: treadmill time (mins) at 6 weeks in the control groups was 20.2	The mean exercise performance measure: treadmill time (mins) at 6 weeks in the intervention group (fludrocortisone) was 2.6 higher (3.85 lower to 9.05 higher)
Symptom scales: Wellness scale Scale from: 0 to 100.	83 (1 study) 11 weeks	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean symptom scales: wellness scale at 11 weeks in the control groups was 2.7	The mean symptom scales: wellness scale at 11 weeks in the intervention groups (fludrocortisone) was 1.1 higher (3.58 lower to 5.78 higher)
Symptom scales: Wellness scale Scale from: 0 to 100.	65 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: wellness scale at 112 weeks in the control groups was 1.7	The mean symptom scales: wellness scale at 12 weeks in the intervention groups (hydrocortisone) was 4.6 higher (0.5 lower to 9.70 higher)
Symptom scales: Sickness Impact Profile Scale from: 0 to 68.	67 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean symptom scales: sickness impact profile at 12 weeks in the control groups was -2.2	The mean symptom scales: sickness impact profile at 12 weeks in the intervention group (hydrocortisone) was

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
					0.3 lower (3.46 lower to 2.86 higher)
Symptom scales: headaches on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: headaches on vas at 6 weeks in the control groups was 6	The mean symptom scales: headaches on vas at 6 weeks in the intervention group (fludrocortisone) was 0 higher (1.55 lower to 1.55 higher)
Symptom scales: painful lymph nodes on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: painful lymph nodes on vas at 6 weeks in the control groups was 3.7	The mean symptom scales: painful lymph nodes on vas at 6 weeks in the intervention group (fludrocortisone) was 0.2 lower (2.31 lower to 1.91 higher)
Symptom scales: sore throat on VAS Scale from: 0 to 10.	40 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: sore throat on vas at 6 weeks in the control groups was 3.3	The mean symptom scales: sore throat on vas at 6 weeks in the intervention group (fludrocortisone) was 0.2 lower (1.8 lower to 1.4 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment) : downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
4 The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. 2) Additionally downgraded due to all participants having rhinitis

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo (95% CI)
(Kakumanu 2003) 5 The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. 2) Additionally downgraded due the majority of evidence coming from a study where all participants had neurally-mediated hypotension (Rowe 2001)					
Studies included for corticosteroids: hydrocortisone – Mckenzie 1998; fludrocortisone – Peterson 1998, Rowe 2001; nasal flunisolide – Kakumanu 2003					

Table 11: Clinical evidence summary: Central antihypertensive drugs (clonidine) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Cognitive function: Stockings of Cambridge - minimum moves	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: stockings of cambridge - minimum moves at 30 minutes in the control groups was 10.22	The mean cognitive function: stockings of cambridge - minimum moves at 30 minutes in the intervention group was 1.22 lower (3.33 lower to 0.89 higher)
Cognitive function: Stockings of Cambridge - initial think time (secs)	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: stockings of cambridge - initial think time (secs) at 30 minutes in the control groups was 9.27	The mean cognitive function: stockings of cambridge - initial think time (secs) at 30 minutes in the intervention group (clonidine) was 1.28 lower (5.19 lower to 2.63 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Cognitive function: Stockings of Cambridge - subsequent thinking time (secs)	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: stockings of cambridge - subsequent thinking time (secs) at 30 minutes in the control groups was 1.89	The mean cognitive function: stockings of cambridge - subsequent thinking time (secs) at 30 minutes in the intervention group was 0.51 lower (3.08 lower to 2.06 higher)
Cognitive function: Rapid Visual Information Processing - reaction time (secs)	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: rapid visual information processing - reaction time (secs) at 30 minutes in the control groups was 5.15	The mean cognitive function: rapid visual information processing - reaction time (secs) at 30 minutes in the intervention group (clonidine) was 0.15 lower (1.42 lower to 1.12 higher)
Cognitive function: Intradimensional (IDS) set sift errors	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: intradimensional (ids) set sift errors at 30 minutes in the control groups was 0.22	The mean cognitive function: intradimensional (ids) set sift errors at 30 minutes in the intervention group was 0.22 higher (0.34 lower to 0.78 higher)
Cognitive function: Extradimensional (EDS) set shift errors	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: extradimensional (eds) set shift errors at 30 minutes in the control groups was 4.44	The mean cognitive function: extradimensional (eds) set shift errors at 30 minutes in the intervention group was 2.66 lower (7.12 lower to 1.8 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Cognitive function: Spatial working memory: between-search errors	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: spatial working memory: between-search errors at 30 minutes in the control groups was 9.26	The mean cognitive function: spatial working memory: between-search errors at 30 minutes in the intervention group was 2.17 lower (7.41 lower to 3.07 higher)
Cognitive function: Spatial working memory: strategy score	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: spatial working memory: strategy score at 30 minutes in the control groups was 31.78	The mean cognitive function: spatial working memory: strategy score at 30 minutes in the intervention group (clonidine) was 0.22 lower (5.92 lower to 5.48 higher)
Cognitive function: pattern recognition - number correct	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: pattern recognition - number correct at 30 minutes in the control groups was 21.4	The mean cognitive function: pattern recognition - number correct at 30 minutes in the intervention group was 0.9 higher (0.77 lower to 2.57 higher)
Cognitive function: spatial recognition - number correct	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: spatial recognition - number correct at 30 minutes in the control groups was 15.3	The mean cognitive function: spatial recognition - number correct at 30 minutes in the intervention group was 0.1 lower (2.44 lower to 2.24 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Cognitive function: spatial span - length	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: spatial span - length at 30 minutes in the control groups was 6.1	The mean cognitive function: spatial span - length at 30 minutes in the intervention group was 0.3 higher (0.84 lower to 1.44 higher)
Cognitive function: delayed matching to sample 2 sec delay	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: delayed matching to sample 2 sec delay at 30 minutes in the control groups was 7.78	The mean cognitive function: delayed matching to sample 2 sec delay at 30 minutes in the intervention group was 1.22 lower (2.65 lower to 0.21 higher)
Cognitive function: paired associate learning - sets completed	18 (1 study) 30 minutes	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: paired associate learning - sets completed at 30 minutes in the control groups was 8.89	The mean cognitive function: paired associate learning - sets completed at 30 minutes in the intervention group was 0 higher (0.3 lower to 0.3 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
Studies included: Morriss 2002					

Table 12: Clinical evidence summary: Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo for ME/CFS

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
Quality of Life: SF36 physical total Scale from: 0 to 100.	140 (2 studies) 4-6 weeks	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		The mean quality of life: sf36 physical total at 4-6 weeks in the control groups was 51.2	The mean quality of life: sf36 physical total at 4-6 weeks in the intervention groups (methylphenidate or dexamphetamine) was 1.63 higher (4.11 lower to 7.37 higher)
Quality of Life: SF36 mental total Scale from: 0 to 100.	140 (2 studies) 4-6 weeks	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		The mean quality of life: sf36 mental total at 4-6 weeks in the control groups was 47.3	The mean quality of life: sf36 mental total at 4-6 weeks in the intervention groups (methylphenidate or dexamphetamine) was 3.51 higher (1.67 lower to 8.69 higher)
Quality of Life: SF36 vitality Scale from: 0 to 100.	42 (1 study) 20 days	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 vitality at 20 days in the control groups was 26.1	The mean quality of life: sf36 vitality at 20 days in the intervention group (modafinil) was 0.6 lower (15.95 lower to 14.75 higher)
Quality of Life: SF36 physical role limitation Scale from: 0 to 100.	42 (1 study) 20 days	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 physical role limitation at 20 days in the control groups was 21.4	The mean quality of life: sf36 physical role limitation at 20 days in the intervention group (modafinil) was 6.45 lower (26.66 lower to 13.76 higher)

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
Quality of Life: SF36 physical function Scale from: 0 to 100.	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 physical function at 20 days in the control groups was 53.6	The mean quality of life: sf36 physical function at 20 days in the intervention group (modafinil) was 1.6 lower (19.6 lower to 16.4 higher)
Quality of Life: SF36 mental health Scale from: 0 to 100.	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 mental health at 20 days in the control groups was 74.9	The mean quality of life: sf36 mental health at 20 days in the intervention group (modafinil) was 6.3 lower (16.26 lower to 3.66 higher)
Quality of Life: SF36 emotional role limitation Scale from: 0 to 100.	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 emotional role limitation at 20 days in the control groups was 95.2	The mean quality of life: sf36 emotional role limitation at 20 days in the intervention group (modafinil) was 19.3 lower (35.88 to 2.72 lower)
Quality of Life: SF36 pain Scale from: 0 to 100.	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 pain at 20 days in the control groups was 57.2	The mean quality of life: sf36 pain at 20 days in the intervention group (modafinil) was 2.45 lower (22.61 lower to 17.71 higher)
Quality of Life: SF36 social Scale from: 0 to 100.	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean quality of life: sf36 social in the control groups was 43.7	The mean quality of life: sf36 social in the intervention group (modafinil) was 2.4 lower (21.85 lower to 17.05 higher

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
		indirectness, imprecision			
Quality of Life: general health Scale from: 0 to 100.	42 (1 study) 20 days	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: sf36 general health in the control groups was 49.2	The mean quality of life: sf36 general health in the intervention group (modafinil) was 0.4 lower (14.35 lower to 13.55 higher)
Fatigue: Checklist Individual Strength (CIS) total score Scale from: 20 to 140.	248 (2 studies) 4-12 weeks	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		The mean fatigue: checklist individual strength (cis) total score at 4-12 weeks in the control groups was 112.5 final score (Blockmans)	The mean fatigue: checklist individual strength (CIS) total score at 4-12 weeks in the intervention groups (methylphenidate) was 7.12 lower (12.07 to 2.16 lower)
Fatigue: Fatigue Severity Scale Scale from: 9 to 63.	44 (2 studies) 6 weeks	⨁⨀⨀⨀ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		The mean fatigue: fatigue severity scale at 6 weeks in the control groups was -2.5	The mean fatigue: fatigue severity scale at 6 weeks in the intervention groups (dexamphetamine or lisdexamphetamine) was 7.67 lower (21.75 lower to 6.4 higher)
Fatigue: Chalder Physical Fatigue scale Scale from: 0 to 21.	42 (1 study) 20 days	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: chalder physical fatigue scale at 20 days in the control groups was 13.6	The mean fatigue: chalder physical fatigue scale at 20 days in the intervention group (modafinil) was 0.25 lower (4.92 lower to 4.42 higher)

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
Fatigue: Chalder Mental Fatigue scale Scale from: 0 to 12.	42 (1 study) 20 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: chalder mental fatigue scale at 20 days in the control groups was 7.4	The mean fatigue: chalder mental fatigue scale at 20 days in the intervention group (modafinil) was 0.4 higher (1.55 lower to 2.35 higher)
Sleep quality: sleep latency (time taken to fall asleep in mins)	20 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to indirectness, imprecision		The mean sleep quality: sleep latency (time taken to fall asleep in mins) at 6 weeks in the control groups was 11.8	The mean sleep quality: sleep latency (time taken to fall asleep in mins) at 6 weeks in the intervention group (dexamphetamine) was 1.2 higher (2.91 lower to 5.31 higher)
Psychological status: HADS anxiety Scale from: 0 to 21.	120 (1 study) 4 weeks	⨁⨁⨁⊖ MODERATE1 due to indirectness		The mean psychological status: hads anxiety at 4 weeks in the control groups was 7.7	The mean psychological status: hads anxiety at 4 weeks in the intervention group (methylphenidate) was 0.4 lower (1.74 lower to 0.94 higher)
Psychological status: HADS depression Scale from: 0 to 21.	120 (1 study) 4 weeks	⨁⨁⨁⊖ MODERATE1 due to indirectness		The mean psychological status: hads depression at 4 weeks in the control groups was 8.7	The mean psychological status: hads depression at 4 weeks in the intervention group (methylphenidate) was 0.4 lower (1.93 lower to 1.13 higher)
Psychological status: Hamilton Anxiety Scale Scale from: 0 to 56.	24 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean psychological status: hamilton anxiety scale at 6 weeks in the control groups was 6.18 improvement	The mean psychological status: hamilton anxiety scale improvement at 6 weeks in the intervention group

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
		indirectness, imprecision			(lisdexamphetamine) was 5.13 higher (2.08 lower to 12.34 higher)
Adverse events: AEs leading to discontinuation	154 (2 studies) 6-12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 2.91 (0.9 to 9.43)	39 per 1000	75 more per 1000 (from 4 fewer to 333 more) (with methylphenidate or lisdexamphetamine)
Adverse events: Serious AEs (pyelonephritis)	128 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	Peto OR 7.63 (0.15 to 384.58)	0 per 1000	20 more per 1000 (from 30 fewer to 60 more) (with methylphenidate)
Adverse events: sleepiness	120 (1 study) 4 weeks	⊕⊖⊖⊖ VERY LOW1,2 due to indirectness, imprecision	RR 0.91 (0.57 to 1.46)	383 per 1000	34 fewer per 1000 (from 165 fewer to 176 more) (with methylphenidate)
Adverse events: dry mouth	146 (2 studies) 4-6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.9 (1.22 to 2.96)	254 per 1000	228 more per 1000 (from 56 more to 497 more) (with methylphenidate or lisdexamphetamine)

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
Adverse events: dizziness	120 (1 study) 4 weeks	⨁⨁⨀⨀ LOW[1,2] due to indirectness, imprecision	RR 0.79 (0.57 to 1.08)	633 per 1000	133 fewer per 1000 (from 272 fewer to 51 more) (with methylphenidate)
Adverse events: akathisia	120 (1 study) 4 weeks	⨁⨁⨀⨀ LOW[1,2] due to indirectness, imprecision	RR 0.85 (0.61 to 1.2)	567 per 1000	85 fewer per 1000 (from 221 fewer to 113 more) (with methylphenidate)
Adverse events: abdominal pain	120 (1 study) 4 weeks	⨁⨀⨀⨀ VERY LOW[1,2] due to indirectness, imprecision	RR 1.22 (0.8 to 1.85)	383 per 1000	84 more per 1000 (from 77 fewer to 326 more) (with methylphenidate)
Adverse events: chest pain	120 (1 study) 4 weeks	⨁⨁⨀⨀ LOW[1,2] due to indirectness, imprecision	RR 0.68 (0.41 to 1.12)	417 per 1000	133 fewer per 1000 (from 246 fewer to 50 more) (with methylphenidate)
Adverse events: anorexia	20 (1 study) 6 weeks	⨁⨁⨀⨀ LOW[1,2] due to indirectness, imprecision	RR 5 (0.7 to 35.5)	100 per 1000	400 more per 1000 (from 30 fewer to 1000 more) (with dexamphetamine)
Adverse events: headache	26 (1 study) 6 weeks	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias,	RR 1.47 (0.15 to	91 per 1000	43 more per 1000 (from 77 fewer to 1000 more) (with lisdexamphetamine)

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
		indirectness, imprecision	14.21)		
Adverse events: insomnia	26 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	Peto OR 5.66 (0.11 to 299.01)	0 per 1000	70 more per 1000 (from 120 fewer to 250 more) (with lisdexamphetamine)
Adverse events	42 (1 study) 20 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.31 (0.79 to 2.17)	571 per 1000	177 more per 1000 (from 120 fewer to 669 more) (with modafinil)
Cognitive function: Behaviour Rating Inventory of Executive Function (BRIEF), global executive composite Scale from: not reported.	24 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness		The mean improvement in cognitive function: behaviour rating inventory of executive function (brief), global executive composite at 6 weeks in the control groups was 3.36	The mean improvement in cognitive function: behaviour rating inventory of executive function (brief), global executive composite at 6 weeks in the intervention group (lisdexamphetamine) was 18.02 higher (8.39 to 27.65 higher)
Pain: McGill pain Questionnaire Scale from: 0 to 78.	24 (1 study) 6 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean pain: mcgill pain questionnaire at 6 weeks in the control groups was 2.54 improvement	The mean pain: mcgill pain questionnaire improvement at 6 weeks in the intervention group (lisdexamphetamine) was

Outcomes	No of Participants (studies *) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo (95% CI)
		indirectness, imprecision			7.84 higher (0.44 to 15.24 higher)
Symptom scales: Clinical Global Improvement - severity Scale from; 1 to 7.	24 (1 study) 6 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: clinical global improvement - severity, at 6 weeks in the control groups was 0.64 improvement	The mean symptom scales: clinical global improvement - severity, at 6 weeks in the intervention group (lisdexamphetamine) was 1.28 higher (0.3 to 2.26 higher)
1 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs 3 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 4 Heterogeneity, I2=86%, p=0.05, unexplained by subgroup analysis.					
Studies included for CNS stimulants: methylphenidate – Blockmans 2006, Montoya 2018; modafinil – Randall 2005; dexamphetamine – Olson 2003; lis-dexamphetamine – Young 2013					

Table 13: Clinical evidence summary: Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo (95% CI)
Fatigue: Multidimensional fatigue inventory (MFI-20) Scale from: 20 to 100.	30 (1 study) 9 months	Original analysis: ⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision PEM reanalysis: ⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		The mean fatigue: multidimensional fatigue inventory (mfi-20) at 9 months in the control groups was -1.1	The mean fatigue: multidimensional fatigue inventory (mfi-20) at 9 months in the intervention group (oral valganciclovir) was 5.05 lower (11.48 lower to 1.38 higher)
Fatigue: POMS fatigue Scale from: 0 to 28.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: poms fatigue at 37 days in the control group was not reported (between-group difference only)	The mean fatigue: poms fatigue at 37 days in the intervention group (IV acyclovir) was 1.26 higher (1.01 lower to 3.53 higher)
Fatigue: POMS vigour Scale from: 0 to 32.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of		The mean fatigue: poms vigour at 37 days in the control group was not reported (between-group difference only)	The mean fatigue: poms vigour at 37 days in the intervention group (IV acyclovir) was

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo (95% CI)
		bias, indirectness, imprecision			2.05 lower (4.65 lower to 0.55 higher)
Psychological status: POMS anxiety Scale from: 0 to 36.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: poms anxiety at 37 days in the control group was not reported (between-group difference only)	The mean psychological status: poms anxiety at 37 days in the intervention group (IV acyclovir) was 2.92 higher (0.63 to 5.21 higher)
Psychological status: POMS depression Scale from: 0 to 60.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: poms depression at 37 days in the intervention groups was not reported (between-group difference only)	The mean psychological status: poms depression at 37 days in the intervention group (IV acyclovir) was 3.97 higher (0.69 to 7.25 higher)
Psychological status: POMS anger Scale from: 0 to 48.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: poms anger at 37 days in the control group was not reported (between-group difference only)	The mean psychological status: poms anger at 37 days in the intervention group (IV acyclovir) was 2.3 higher (0.13 lower to 4.73 higher)
Psychological status: POMS confusion Scale from: 0 to 28.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean psychological status: poms confusion at 37 days in the control group was not reported (between-group difference only)	The mean psychological status: poms confusion at 37 days in the intervention group (IV acyclovir) was 1.83 higher (0.57 to 3.09 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo (95% CI)
		indirectness, imprecision			
Adverse events: treatment-related adverse events	30 (1 study) 9 months	Original analysis: ⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, indirectness PEM reanalysis: ⨁⊖⊖⊖ VERY LOW1,5 due to risk of bias, indirectness	RD 0.00 (-0.14 to 0.14)	0 per 1000	0 more per 1000 (from 140 fewer to 140 more) (with oral valganciclovir)
Adverse events: reversible renal failure	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	Peto OR 7.99 (0.8 to 80.28)	0 per 1000	11 more per 1000 (from 20 fewer to 240 more) (with IV acyclovir)
Activity levels: rest (hours/day)	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean activity levels: rest (hours/day) at 37 days in the control group was not reported (between-group difference only)	The mean activity levels: rest (hours/day) at 37 days in the intervention group (IV acyclovir) was 0.05 lower (0.83 lower to 0.73 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo (95% CI)
Symptom scales: Wellness score Scale from: not reported.	54 (1 study) 37 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: wellness score at 37 days in the control group was not reported (between-group difference only)	The mean symptom scales: wellness score at 37 days in the intervention group (IV acyclovir) was 1.08 lower (7.28 lower to 5.12 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. 2) Montoya 2013 was additionally downgraded due to population having suspected viral onset and elevated antibody tiers. 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs 4 The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature; 2) additionally downgraded due to population having suspected viral onset and elevated antibody tiers (Montoya 2013) [original analysis] 5 The majority of the evidence included an indirect population (downgraded by one increment): requirement for suspected viral onset and elevated viral antibody tiers (Montoya 2013). [PEM reanalysis]					
Studies included: Montoya 2013 (oral valganciclovir); Strauss 1988 (IV acyclovir)					

Table 14: Clinical evidence summary: 5-HT3 antagonists (ondansetron) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with 5-HT3 antagonists (ondansetron) versus placebo (95% CI)
Fatigue: Checklist Individual Strength fatigue Scale from: 8 to 56.	67 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: cis fatigue at 12 weeks in the control groups was 45.4	The mean fatigue: cis fatigue at 12 weeks in the intervention groups was 1.4 lower (6.81 lower to 4.01 higher)
Activity levels: Actometer (objective accelerometer-based method of measuring activity)	67 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean activity levels: actometer (objective accelerometer-based method of measuring activity) at 12 weeks in the control groups was 60.6	The mean activity levels: actometer (objective accelerometer-based method of measuring activity) at 12 weeks in the intervention groups was 5.6 lower (13.61 lower to 2.41 higher)
Adverse events: constipation	67 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	Peto OR 7.86 (0.48 to 128.37)	0 per 1000	60 more per 1000 (from 40 fewer to 160 more)
Adverse events: malaise	67 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 3.09 (0.34 to 28.23)	29 per 1000	61 more per 1000 (from 19 fewer to 801 more)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with 5-HT3 antagonists (ondansetron) versus placebo (95% CI)
Symptom scales: Sickness Impact Profile (SIP) 8 Scale from: 0 to 5799.	67 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean symptom scales: sickness impact profile (sip) 8 at 12 weeks in the control groups was 1172	The mean symptom scales: sickness impact profile (sip) 8 at 12 weeks in the intervention groups was 109 lower (403.38 lower to 185.38 higher)
1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias 2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
* Studies included: The 2010					

Table 15: Clinical evidence summary: Galantamine hydrobromide versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Galantamine hydrobromide versus placebo (95% CI)
Fatigue: fatigue on VAS Scale from: 0 to 10.	49 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		The mean fatigue: fatigue on vas at 2 weeks in the control groups was 7.11	The mean fatigue: fatigue on vas at 2 weeks in the intervention groups was 0.14 higher (0.84 lower to 1.12 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Galantamine hydrobromide versus placebo (95% CI)
		indirectness, imprecision			
Cognitive function: memory on VAS Scale from: 0 to 10.	49 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: memory on vas at 2 weeks in the control groups was 4.72	The mean cognitive function: memory on vas at 2 weeks in the intervention groups was 0.91 higher (0.67 lower to 2.49 higher)
Pain: myalgia on VAS Scale from: 0 to 10.	49 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean pain: myalgia on vas at 2 weeks in the control groups was 7.99	The mean pain: myalgia on vas at 2 weeks in the intervention groups was 0.47 lower (1.39 lower to 0.45 higher)
Sleep quality: sleep disturbance on VAS Scale from: 0 to 10.	49 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean sleep quality: sleep disturbance on vas at 2 weeks in the control groups was 6.66	The mean sleep quality: sleep disturbance on vas at 2 weeks in the intervention groups was 0.34 higher (1.02 lower to 1.7 higher)
Adverse events: AEs dizziness on VAS Scale from: 0 to 10.	49 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean adverse events: aes dizziness on vas at 2 weeks in the control groups was 3.54	The mean adverse events: aes dizziness on vas at 2 weeks in the intervention groups was 0.72 higher (0.93 lower to 2.37 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Galantamine hydrobromide versus placebo (95% CI)
Return to school/work: work capacity/satisfaction on VAS Scale from: 0 to 10.	39 (1 study) 2 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean quality of life: work capacity/satisfaction on vas at 2 weeks in the control groups was 5.09	The mean quality of life: work capacity/satisfaction on vas at 2 weeks in the intervention groups was 0.17 lower (1.38 lower to 1.04 higher)
Symptom scales: clinical global impression score, no change or worse	347 (1 study) 20 weeks	⨁⊖⊖⊖ VERY LOW2,3 due to risk of bias, indirectness, imprecision	RR 0.86 (0.72 to 1.03)	701 per 1000	98 fewer per 1000 (from 196 fewer to 21 more)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Studies included: Blacker 2004, Snorrason 1996

Table 16: Clinical evidence summary: Antihistamines (terfenadine) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antihistamines (terfenadine) versus placebo (95% CI)
Physical functioning: modified Medical Outcome Study Short Form - physical functioning Scale from: 0 to 100.	28 (1 study) 2 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean physical functioning: modified medical outcome study short form - physical functioning at 2 months in the control groups was 69.66	The mean physical functioning: modified medical outcome study short form - physical functioning at 2 months in the intervention groups was 6.56 lower (19.75 lower to 6.63 higher)
Psychological status: modified Medical Outcome Study Short Form - mental health Scale from: 0 to 100.	28 (1 study) 2 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean psychological status: modified medical outcome study short form - mental health at 2 months in the control group was 74.62	The mean psychological status: modified medical outcome study short form - mental health at 2 months in the intervention groups was 10.73 lower (24.5 lower to 3.04 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; the percentage of participants with PEM is <95%; additionally it was unclear if this actually was PEM (described as “post-exertional fatigue (prolonged)”) [PEM reanalysis – see Appendix G for additional details]
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Studies included: Steinberg 1996

Table 17: Clinical evidence summary: Pro-inflammatory cytokine antagonists (anakinra) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Pro-inflammatory cytokine antagonists (anakinra) versus placebo (95% CI)
Mortality	50 (1 study) 24 weeks	⨁⨁⨁⨀ MODERATE1 due to indirectness	RD 0.00 (-0.07 to 0.07)	0 per 1000	0 more per 1000 (from 70 fewer to 70 more)
Fatigue: Checklist Individual Strength fatigue Scale from: 8 to 56.	50 (1 study) 24 weeks	⨁⨀⨀⨀ VERY LOW1,2 due to indirectness, imprecision		The mean fatigue: cis fatigue at 24 weeks in the control groups was 0	The mean fatigue: cis fatigue at 24 weeks in the intervention groups was 1.3 higher (5.3 lower to 7.9 higher)
Physical functioning: SF36 physical function Scale from: 0 to 100.	50 (1 study) 24 weeks	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		The mean physical functioning: sf36 physical function at 24 weeks in the control group was 64.8	The mean physical functioning: sf36 physical function at 24 weeks in the intervention groups was 4 lower (15.1 lower to 7.1 higher)
Psychological status: Symptom Checklist 90 Scale from: 90 to 450.	50 (1 study) 24 weeks	⨁⨁⨁⨀ MODERATE1, 2 due to indirectness		The mean psychological status: symptom checklist 90 at 24 weeks in the control group was 140.5	The mean psychological status: symptom checklist 90 at 24 weeks in the intervention groups was 3 higher (8.6 lower to 14.6 higher)
Pain: VAS maximum pain score Scale from: 0 to 10.	50 (1 study) 24 weeks	⨁⨀⨀⨀ VERY LOW1,2 due to indirectness, imprecision		The mean pain: vas maximum pain score at 24 weeks in the control group was 6.6	The mean pain: vas maximum pain score at 24 weeks in the intervention groups was 0.34 higher (1.1 lower to 1.78 higher)
Adverse events	50 (1 study) 24 weeks	⨁⨁⨁⨀ MODERATE1 due to indirectness	RR 1.71 (1.2 to 2.45)	560 per 1000	398 more per 1000 (from 112 more to 812 more)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Pro-inflammatory cytokine antagonists (anakinra) versus placebo (95% CI)
Adverse events: withdrawal due to adverse events	50 (1 study) 24 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to indirectness, imprecision	Peto OR 7.39 (0.15 to 372.38)	0 per 1000	40 more per 1000 (from 60 fewer to 140 more)
Symptom scales: Sickness Impact Profile Scale from: 0 to 5799.	50 (1 study) 24 weeks	⨁⨁⊖⊖ LOW1,2 due to indirectness, imprecision		The mean symptom scales: sickness impact profile at 24 weeks in the control groups was 1260.4	The mean symptom scales: sickness impact profile at 24 weeks in the intervention groups was 91.2 higher (275.8 lower to 458.2 higher)
1 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details] 2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					
Studies included: Roerink 2017					

Table 18: Clinical evidence summary: Staphylococcus vaccine (Staphypan Berna) versus placebo for ME/CFS

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Staphylococcus vaccine (Staphypan Berna) versus placebo (95% CI)
Pain: pain on VAS Scale from: unclear	98 (1 study) 32 weeks	⨁⨁⊖⊖ LOW1 due to indirectness		The mean pain: pain on vas at 32 weeks in the control groups was 6.2	The mean pain: pain on vas at 32 weeks in the intervention groups was 0.3 lower (1.12 lower to 0.52 higher)
Adverse events	100 (1 study) 32 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to	RR 1.08 (0.75	520 per 1000	42 more per 1000 (from 130 fewer to 286 more)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Staphylococcus vaccine (Staphypan Berna) versus placebo (95% CI)
		indirectness, imprecision	to 1.55)		
Symptom scales: clinical global impression of change Scale from; 1 to 7.	98 (1 study) 32 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to indirectness, imprecision		The mean symptom scales: clinical global impression of change at 32 weeks in the control group was 4.4	The mean symptom scales: clinical global impression of change at 32 weeks in the intervention groups was 0.7 lower (1.22 to 0.18 lower)
Symptom scales: clinical global impression of severity Scale from: 1 to 7.	98 (1 study) 32 weeks	⨁⊖⊖⊖ VERY LOW1,2 due to indirectness, imprecision		The mean symptom scales: clinical global impression of severity at 32 weeks in the control group was 4.8	The mean symptom scales: clinical global impression of severity at 32 weeks in the intervention groups was 0.3 lower (0.53 to 0.07 lower)

1 The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments):1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]; 2) further downgraded as the population was also required to also meet diagnostic criteria for fibromyalgia.
2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Studies included: Zachrisson 2002

Table 19: Clinical evidence summary: Central antihypertensive drugs (clonidine) versus placebo for ME/CFS (children and young people)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Children and young people: Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Fatigue: Chalder Fatigue Questionnaire (CFQ) total sum score Scale from: not reported.	103 (1 study) 30 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean fatigue: chalder fatigue questionnaire (cfq) total sum score at 30 weeks in the control group was 13.5	The mean fatigue: chalder fatigue questionnaire (cfq) total sum score at 30 weeks in the intervention groups was 0.5 higher (14.7 lower to 15.7 higher)
Physical functioning: Fatigue Disability Index (FDI) total sum score Scale from: not reported.	103 (1 study) 30 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean physical functioning: fatigue disability index (fdi) total sum score at 30 weeks in the control group was 16.8	The mean physical functioning: fatigue disability index (fdi) total sum score at 30 weeks in the intervention groups was 0.2 higher (13.3 lower to 13.7 higher)
Pain: BPI average pain score Scale from: 0 to 10.	103 (1 study) 30 weeks	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		The mean pain: bpi average pain score at 30 weeks in the control group was 3.3	The mean pain: bpi average pain score at 30 weeks in the intervention groups was 0.4 higher (0.4 lower to 1.2 higher)
Sleep quality: KSQ insomnia score Scale from: not reported.	103 (1 study) 30 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean sleep quality: ksq insomnia score at 30 weeks in the control group was 3.6	The mean sleep quality: ksq insomnia score at 30 weeks in the intervention groups was 0.1 higher (0.3 lower to 0.5 higher)

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Children and young people: Central antihypertensive drugs (clonidine) versus placebo (95% CI)
Adverse effects: various self-reported	108 (1 study) 9 weeks	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, indirectness, imprecision	RR 1.17 (0.91 to 1.5)	647 per 1000	110 more per 1000 (from 58 fewer to 324 more)
Activity levels: steps per day (accelerometer)	103 (1 study) 30 weeks	⨁⨁⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean activity levels: steps per day (accelerometer) at 30 weeks in the control group was 4652	The mean activity levels: steps per day (accelerometer) at 30 weeks in the intervention groups was 119 higher (796 lower to 1034 higher)
Cognitive function: Digit span backward test total Scale from: not reported	103 (1 study) 30 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		The mean cognitive function: digit span backward test total at 30 weeks in the control group was 6.7	The mean cognitive function: digit span backward test total at 30 weeks in the intervention groups was 0.5 lower (1.2 lower to 0.2 higher)
Symptom scales: CFS symptom inventory hypersensitivity score Scale from: not reported	103 (1 study) 30 weeks	⨁⨁⊖⊖ LOW1,2 due to risk of bias, indirectness		The mean symptom scales: cfs symptom inventory hypersensitivity score at 30 weeks in the control group was 2.6	The mean symptom scales: cfs symptom inventory hypersensitivity score at 30 weeks in the intervention groups was 0.03 lower (0.4 lower to 0.34 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 The majority of the evidence included an indirect population (downgraded by one increment): downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional

Outcomes	No of Participants (studies*) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Children and young people: Central antihypertensive drugs (clonidine) versus placebo (95% CI)
details] 3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs 4 Outcome indirectness: Some adverse effects are poorly defined, e.g. "unwellness" and "other"					
Studies included: Sulheim 2014					

See Appendix F for full GRADE and/or GRADE-CERQual tables.

More information on the minimally important differences (MIDs) used and the interpretation can be found in Appendix K of this review and the Methods Chapter of this guideline.

1.1.7. Economic evidence

1.1.7.1. Included studies

No health economic studies were included.

1.1.7.2. Excluded studies

No relevant health economic studies were specifically excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in Appendix G.

1.2. The committee's discussion and interpretation of the evidence

The committee's discussion on the evidence reviews for the clinical and cost-effectiveness of pharmacological interventions and the experiences of people who have had interventions for ME/CFS are included here. See Evidence review G – Non-pharmacological management for the full methods and results sections of the review on the experiences of people who have had interventions (pharmacological and non-pharmacological) for ME/CFS.

The committee discussed this evidence with the findings from the review on diagnosis (report D) and the reports on Children and Young people (Appendix 1) and people with severe ME/CFS (Appendix 2). Where relevant this is noted.

1.1.8. The outcomes that matter most – review of the clinical and cost effectiveness

Mortality, quality of life, general symptom scales, fatigue/fatigability, physical function, cognitive function, psychological status, pain, sleep quality, treatment-related adverse events, activity levels, return to school/work and exercise performance measures were considered by the committee to be critical outcomes for decision making.

Fatigue/fatigability, unrefreshing sleep and physical and cognitive dysfunction are recognised as key symptoms of ME/CFS. The worsening or improvement of these symptoms reflect the impact of an intervention or strategy. The committee agreed that pain though not key to the diagnosis of ME/CFS, is a common symptom in people with ME/CFS and should be considered by the committee in their decision making. The committee agreed that any decisions on interventions and strategies should be informed by treatment related adverse events as a possible indicator of harm.

Care needs, impact on families and carers and ability to resume occupation, school or study were considered important outcomes for decision making reflecting the effectiveness of an intervention.

The committee acknowledged the lack of existing objective outcome measures of effectiveness of interventions for ME/CFS and the limitations of subjective measures (see Professor Edwards expert testimony – Appendix 3: Expert testimonies). Only validated outcome measurement scales were included in the evidence review.

No evidence was identified for care needs or impact on families and carers.

1.1.9. The outcomes that matter most – qualitative review of experiences of interventions review of the clinical and cost effectiveness

Themes emerging from qualitative data regarding the experiences of people that have had interventions for ME/CFS. Themes were derived from the evidence identified and were not pre-specified by the committee.

Only findings that were relevant to the review question were included; findings related to people's experiences of general ME/CFS services rather than specific interventions were not extracted.

1.1.10. The quality of the evidence – summary of quality in review of clinical and cost effectiveness

Evidence from 30 studies was identified for the following pharmacological interventions: immunomodulatory drugs (n=6), antidepressants (n=5), corticosteroids (n=4), antihypertensive drugs (n=2), central nervous system stimulants (n=5), antiviral drugs (n=2), 5-HT3 antagonists (n=1), galantamine hydrobromide (n=2), antihistamines (n=1), pro-inflammatory cytokine antagonists(n=1) and staphylococcus vaccine (n=1). No evidence was identified for sleep medication, pain relief, sodium valproate or low dose naltrexone.

The majority of the interventions were compared with placebo. The study populations were mostly adults all with mixed or unclear ME/CFS severity. One study comparing clonidine to placebo included young people (12-18 years).

Most of the evidence was of low and very low quality. The main reasons for downgrading were risk of bias, indirectness and imprecision. Several outcomes were at high risk of selection bias due to insufficient randomisation and allocation concealment methods reported in the studies. The majority of the studies were placebo-controlled and double blinded, but those that were not double blinded were at high risk of performance bias. This was particularly relevant for subjective outcomes and the committee considered this limitation when interpreting the evidence.

After considering the stakeholder comments the committee agreed to revisit the evidence for the intervention reviews, further scrutinising the information on PEM reported in the trials and the application of indirectness in the evidence. For outcomes that were reanalysed, this did not result in any changes to the overall quality rating of the evidence. Further information on this analysis is briefly summarised elsewhere in this section. Full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence are in Appendix G.

For most outcomes, meta-analysis was not appropriate due to important differences between the types of drugs or multiple relevant measures of the same outcome being reported within the same study. Most of the comparisons only included one study. Therefore, evidence for most outcomes was based on single studies, many of which included very small sample sizes. This resulted in imprecision around the point estimates.

Population indirectness

The committee discussed the CDC 1994 diagnostic criteria used in the studies to recruit eligible participants. The committee have identified PEM/PESE as an essential symptom that is central to the diagnosis of ME/CFS (see evidence report D: diagnosis) and the CDC 1994 criteria does not include this as a compulsory requirement. It should be noted that PESE is also referred to as post exertional malaise (PEM) in the criteria, PESE is the committee's preferred term. The committee agreed that a population diagnosed with such criteria may not accurately represent the ME/CFS population and that people experiencing PEM/PESE are likely to respond differently to treatment than those who do not experience PEM/PESE. This raised concerns over the generalisability of findings to the ME/CFS population. It was therefore agreed to downgrade the evidence for population indirectness.

After considering the stakeholder comments the committee agreed to revisit the evidence for the intervention reviews further scrutinising the information on PEM reported in the trials and the application of indirectness in the evidence. As part of this they agreed that any evidence with a population ≥ 95% with PEM would be considered direct. Studies where < 95% of participants had PEM, or where the percentage of participants with PEM was not reported would be considered indirect. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Evidence was not stratified by diagnostic criteria used, so theoretically, studies including potentially different populations could have been combined. In practice, for the majority of outcomes, meta-analysis was not appropriate due to important differences between the types of interventions, comparators, population strata, or multiple relevant measures of the same outcome being reported within the same study. Therefore, potentially different populations were rarely combined. Where they were combined, no serious heterogeneity was identified.

Evidence quality by intervention

Immunomodulatory drugs

Evidence from six randomised controlled trials were identified for immunomodulatory drugs compared to placebo; two rituximab, three IV immunoglobulin G and one rintatolimod. Most of the evidence was low and very low quality (apart from some moderate and high quality evidence for rituximab) and based on single small studies. No evidence was identified for mortality, cognitive function, pain, sleep quality, activity levels, care needs and impact on families and carers.

Antidepressants and antipsychotics

Evidence from five randomised controlled trials were identified for antidepressants. Three trials (single trials on the serotonin-norepinephrine reuptake inhibitor (SNRI) duloxetine hydrochloride, the monoamine oxidase inhibitor (MAOI) moclobemide, and the selective serotonin reuptake inhibitor (SSRI) fluoxetine) were compared only to placebo. One trial had four arms comparing fluoxetine and exercise control, graded exercise and placebo, fluoxetine and graded exercise, and placebo and exercise control. One trial compared fluoxetine to amisulpride (an atypical antipsychotic). All the evidence very low quality and the majority was based on single small studies. No evidence was identified for mortality, cognitive function, sleep quality, activity levels, return to school/work, exercise performance measures, care needs and impact on families and carers.

Corticosteroids

Evidence from four randomised controlled trials were identified for corticosteroids (one nasal flunisolide, two oral fludrocortisone and one oral hydrocortisone) compared to placebo. Most of the evidence was very low quality and based on single small studies. No evidence was identified for mortality, physical function, activity levels, return to school/work, care needs and impact on families and carers.

Central antihypertensive drugs

Evidence from two randomised controlled trials compared clonidine to placebo. Most of the evidence was very low to low quality and based on single small studies. No evidence was identified for mortality, quality of life, psychological status, return to school/work, exercise performance measures, care needs and impact on families and carers.

Central nervous system (CNS) stimulants

Evidence from five randomised controlled trials identified for CNS stimulants (two methylphenidate, and one each of dexamphetamine, lisdexamphetamine, and Modafinil) compared to placebo. Most of the evidence was very low to low quality based on single small studies. No evidence was identified for mortality, physical function, activity levels, return to school/work, exercise performance measures, care needs and impact on families and carers.

Antiviral drugs

Evidence from two randomised controlled trials compared acyclovir (IV and oral) to placebo. All the evidence was very low quality and based on single small studies. No evidence was identified for mortality, quality of life, physical function, cognitive function, pain, sleep quality,

return to school/work, exercise performance measures, care needs and impact on families and carers.

5-HT3 antagonists

Evidence from one randomised controlled trial compared ondansetron to placebo. All the evidence was very low quality. No evidence was identified for mortality, quality of life, physical function, psychological status, sleep quality, exercise performance measures, care needs and impact on families and carers were also considered to be important outcomes.

Galantamine hydrobromide

Evidence from two randomised controlled trials compared galantamine hydrobromide to placebo. All the evidence was very low quality and based on single small studies. No evidence was identified for mortality, quality of life, physical function, psychological status, activity levels, exercise performance measures, care needs and impact on families and carers were also considered to be important outcomes.

Antihistamines

Evidence from one randomised controlled trial compared terfenadine to placebo. All the evidence was very low quality. No evidence was identified for mortality, quality of life, general symptom scales, fatigue/fatigability, cognitive function, pain, sleep quality, treatment-related adverse events, activity levels, return to school/work, exercise performance measures, care needs and impact on families and carers.

Pro-inflammatory cytokine antagonists

Evidence from one randomised controlled trial compared anakinra to placebo. The evidence was very low to moderate quality. No evidence was identified for quality of life, cognitive function, sleep quality, activity levels, return to school/work, exercise performance measures, care needs and impact on families and carers.

Staphylococcus vaccine

Evidence from one randomised controlled trial compared staphylococcus vaccine to placebo. All the evidence was very low quality. No evidence was identified for mortality, quality of life, fatigue/fatigability, physical function, cognitive function, psychological status, sleep quality, activity levels, return to school/work, exercise performance measures, care needs and impact on families and carers.

1.1.11. The quality of the evidence - qualitative review of people's experience of interventions

The majority of the studies included in the qualitative review reported experiences of non-pharmacological interventions. One study in adults, using a survey with open-ended questions, reported experiences of antidepressants. Two studies in children and young people, using semi-structured interviews, reported experiences of sickness or stomach acid relief medication or pharmacological interventions in general.

Confidence in the review findings was very low. The main reasons for downgrading were methodological limitations, relevance and adequacy. Issues regarding recruitment strategy and data analysis were the main contributory factors to concerns over methodological limitations in the study in adults. The main methodological limitations of the studies in children/young people included the role of the researcher and concerns regarding the richness of the data. The study in adults reported limited information on participant characteristics, so it was unclear how applicable the findings were to the wider ME/CFS population. There were also concerns regarding applicability of the findings reported in the

studies on children/young people; the population in one study was limited to adolescents with ME/CFS who experienced eating difficulties and the population in the other study was limited to children/young people with comorbid depression. Findings were reported without elaboration or examples and were based on single studies, leading to concerns regarding adequacy.

1.1.12. Benefits and harms - Review of clinical and cost effectiveness

Immunomodulatory drugs

The evidence showed a clinical benefit of rituximab compared with placebo for the physical component of SF36 quality of life, however there was some uncertainty (imprecision) around the point estimate. The evidence showed no clinically important difference of rituximab for the mental component of SF36 quality of life, fatigue/fatigability, activity levels, and physical functioning. High and moderate quality evidence showed harm of rituximab for serious adverse events and adverse events of at least moderate severity, respectively. Serious adverse events included febrile neutropenia, infusion-related reactions, and other events also requiring hospitalisation, and were considered to be possibly or probably related to the study intervention.

The evidence showed a clinical benefit of intravenous immunoglobulin G (IV Ig) compared with placebo for symptom improvement and for return to work, however there was some uncertainty (imprecision) around the point estimate for symptom improvement. There was no clinically important difference of IV Ig for psychological status, or physical functioning. There was a clinically important difference in adverse events, with participants from one study receiving IVIG reporting fewer constitutional symptoms compared to those receiving placebo. In another study, there was no clinical difference in unspecified major adverse events between study arms.

There was no clinically important difference of rintatolimod for exercise performance (treadmill test). There was no clinical difference in serious adverse events with possible or probable relation to the intervention, between study arms.

The committee considered that the majority of the evidence for immunomodulatory drugs was of low and very low quality and based on single small studies and the committee was not confident about the effects for any of the treatments. The committee were aware from their clinical experience that immunomodulatory drugs can cause serious adverse events, and they acknowledged the high quality evidence of harm of rituximab. The committee were aware of anecdotal reports of some of these drugs working for some people with ME/CFS, however they decided that due to the limitations of the evidence, the lack of any clear benefit, and potential for serious harms, immunomodulatory drugs should not be used for the purposes of treating or curing ME/CFS.

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Antidepressants and antipsychotics

The evidence showed a clinical benefit of duloxetine (SNRI antidepressant) compared with placebo for the bodily pain sub scale of SF36 quality of life and the general fatigue sub scale of the MFI-20 fatigue scale, however there was some uncertainty (imprecision) around the point estimates. There was no clinical difference for the remainder of the SF36 or MFI-20 sub scales, the hospital anxiety and depression scale, the brief pain inventory, or general symptom scales (clinical global impression of severity and improvement) for duloxetine. Evidence showed a clinical benefit of fluoxetine (SSRI antidepressant) and moclobemide

(MAOI antidepressant) for general symptom scales, however there was considerable uncertainty around the point estimates. There was no clinical difference of fluoxetine for fatigue, beck depression inventory and exercise performance. For moclobemide there was no clinical difference for physical functioning or profile of mood states. There was a harm of fluoxetine for adverse events (tremor/perspiration).

The committee considered that the majority of the evidence was of low and very low quality and based on single studies, and they were not confident about the effects. The committee noted the evidence suggesting harm of fluoxetine in the form of side effects was also broadly reflected in the qualitative review of people's experiences of interventions, though this evidence was also of low quality (see Evidence review G for the full methods and results of this review, and section 1.1.13 below). The committee are also aware from their own experience that ME/CFS is commonly misdiagnosed as depression and that treatment with antidepressants is often given on the basis of these incorrect beliefs. The committee decided based on the lack of any clear benefit from the evidence and their own clinical experience that antidepressants should not be used for the purpose of ME/CFS. However, they acknowledged that people with ME/CFS can experience comorbid depression, and that antidepressants may be useful in some of these people as a treatment for depression (as for any other person with depression regardless of whether or not they have ME/CFS). The committee cross referred to the NICE guideline on depression.

The committee also reviewed the evidence for fluoxetine compared with amisulpride (atypical anti-psychotic) and graded exercise therapy. Evidence from one study showed a clinical benefit of amisulpride over fluoxetine for quality of life, general symptom scales and fatigue, but no clinically important difference for psychological status, or pain, There was no clinically important difference in adverse events (FIBSER global burden) between study arms. The committee considered the lack of robust evidence identified for anti-psychotics and their own experience of potential harms and decided that anti-psychotics should not be used for the purposes of treating or curing ME/CFS.

Very low quality evidence from one four armed study showed no clinically important difference in fatigue, psychological status or exercise performance between fluoxetine, graded exercise therapy, placebo and exercise control. The committee considered that there was insufficient evidence to conclude whether SSRIs were more effective than graded exercise therapy. The evidence for graded exercise therapy is discussed further in Evidence review G - non pharmacological management.

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Corticosteroids

Evidence for corticosteroids was mainly of very low quality. Evidence for hydrocortisone showed no clinical difference for any of the outcomes assessed for fatigue, psychological status, general symptom scales, and activity levels.. There was a clinically important harm of hydrocortisone in terms of adverse events, with participants in the hydrocortisone study arm experience more adverse reactions compared to those in the placebo arm.

Similarly, for fludrocortisone there was no clinical difference for any SF36 quality of life subscales, fatigue, physical functioning, psychological status, cognitive functioning, pain, sleep, activity levels, exercise performance, or general symptom scales. There was a clinically important difference in any adverse events and adverse events leading to study withdrawal, with participants receiving fludrocortisone experiencing fewer events, the latter of which was from one small study with considerable uncertainty (imprecision) around the point estimate.

There was no clinical difference for symptom severity and sleep for nasal flunisolide.

The committee raised concerns about the long-term safety of these drugs for people with ME/CFS, specifically disruption to the hypothalamic-pituitary-adrenal axis and weakening of muscle and bone. Taking into account the very low quality of the evidence and lack of any clear benefit, as well as their own clinical experience of the potential harms, the committee decided that corticosteroids should not be used as a curative treatment of ME/CFS. The committee was aware that fludrocortisone is sometimes given for orthostatic intolerance syndromes, such as postural hypotension or Postural Tachycardia Syndrome (POTS). They agreed that this recommendation would not prevent people with ME/CFS being offered fludrocortisone treatment for relevant comorbidities, but that it should not be offered for the purpose of treating or curing ME/CFS. See Evidence review G - non pharmacological management report for further recommendations and discussion on the management of orthostatic intolerance.

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Central antihypertensive drugs

Evidence from one small study showed a clinical benefit of clonidine compared with placebo for some measures of cognitive function, but no clinically important difference for others. There was considerable uncertainty (imprecision) around most of the effect estimates. The committee noted that the evidence of benefit for cognitive function was based on a small study whereby a single dose of the drug was administered and follow up was at thirty minutes and the committee was not confident in the effect.

Low to very low quality evidence from one study in young people showed no clinically importance difference in general symptom scales, fatigue, physical function, sleep quality or activity levels, and harm of clonidine for cognitive function, and pain. There was a harm of clonidine for adverse events, with participants receiving clonidine experiencing more self-reported adverse events. This was from one small study and there was uncertainty (imprecision) around the effect estimate.

The committee considered the limitations of the evidence, the evidence of potential harm as well as their own clinical knowledge regarding evidence for other relevant conditions and decided that that clonidine should not be used for the curative treatment of ME/CFS.

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Central nervous system (CNS) stimulants

Evidence showed a clinical benefit of amphetamines (dexamphetamine and lisdexamphetamine) compared with placebo for reducing fatigue on the fatigue severity scale, anxiety measured by the Hamilton anxiety scale, general symptom scales, pain and cognitive function, however there was uncertainty (imprecision) around the point estimates for most of these outcomes. There was no clinical difference for SF36 quality of life and sleep scales. The evidence showed harm of amphetamines for adverse events leading to discontinuation and other adverse events, including anorexia, dry mouth, headache and insomnia, however there was considerable uncertainty (imprecision) around the point estimates.

For methylphenidate there was no clinical difference for fatigue, psychological status, and SF36 quality of life. There was harm for serious adverse events, however the adverse event that occurred was not considered to be treatment-related (pyelonephritis), dry mouth, and abdominal pain. The remaining evidence on adverse events was mixed depending on the adverse event. There was a clinical difference in some adverse events (dizziness, chest pain and akathisia), with participants receiving methylphenidate experiencing fewer events, however there was uncertainty (imprecision) around these point estimates. There was no clinical difference between study arms for the adverse event, sleepiness.

Finally, short-term evidence from one study (20-day follow-up) showed a harm of modafinil for adverse events (none were serious and included headache and nausea) and some sub scales of SF36 quality of life. There was considerable uncertainty (imprecision) around all of these point estimates. There was no clinical difference for other SF36 sub scales and the Chalder fatigue scale.

The committee noted the very low quality of the evidence and was not confident in the effects for CNS stimulants. The committee discussed their experience of CNS stimulants and were concerned about possible harms. They noted that CNS stimulants could cause people with ME/CFS to push themselves outside of their energy limits which could have damaging effects. They also discussed side effects which could be particularly detrimental to people with ME/CFS, some of which were noted in the evidence (such as anorexia and insomnia). The committee considered the low to very low quality of the evidence, as well as their own clinical knowledge regarding evidence for other chronic conditions, and possible harms, and decided that CNS stimulants should not be used for the curative treatment of ME/CFS.

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Antiviral drugs

The evidence showed a clinical benefit of oral valganciclovir compared with placebo for fatigue (MFI-20), however there was some uncertainty (imprecision) around the effect estimate. There was no clinical difference between study arms in terms of treatment-related adverse events.

The evidence showed harm of intravenous (IV) acyclovir for profile of mood states and adverse events (reversible renal failure), although there was some uncertainty (imprecision) around the effect estimates. There was no clinically important difference of IV acyclovir for general symptom scales or activity levels.

The committee noted that evidence for antiviral drugs came from two small studies and was of very low quality, and they could not be confident of the effects. Evidence of harm came from a single study on IV acyclovir with a short follow up of 37 days. The committee discussed that antiviral drugs are used by some specialists and they were aware of anecdotal evidence of benefit in some people, but they recognised the absence of convincing clinical evidence and possible harms. Therefore, the committee recommended that antiviral drugs should not be used for purposes of treating or curing ME/CFS, however they acknowledged this recommendation should not stop antiviral drugs being used where a genuine indication exists, for example for the treatment of some viral infections.

After further scrutinising the information on PEM reported in the trials, the study on oral valganciclovir (Montoya 2013) that was previously downgraded for indirectness for using a diagnostic criteria that did not have PEM as a compulsory feature (1994 CDC criteria) reported that ≥95% of participants had PEM. As a result, the indirectness rating was changed from very serious to serious indirectness (the study remained downgraded by one increment for other population concerns relating to a requirement for viral onset of disease and elevated

viral titres). Only two outcomes were reported in this trial: fatigue and adverse events. The overall quality rating remained very low for both outcomes and the committee's interpretation of the evidence did not change.

For the other trial of antiviral drugs, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

Other drugs

The committee also reviewed evidence for antidepressants combined with graded exercise, 5HT3 antagonists, galantamine, antihistamines, proinflammatory cytokine antagonists and staphylococcus vaccine. Evidence for these comparisons was mostly low and very low quality and based on individual studies. Due to the significant limitations of the evidence the committee agreed that none of these drug treatments should be offered for the purpose of treating or curing ME/CFS, but they noted there may be other indications for the use of some of these medications (for example for management of specific symptoms or comorbidities).

PEM re-analysis

After further scrutinising the information on PEM reported in the trials, no new information on PEM was identified that required re-analysis. See Appendix G for full details of the approach taken, the analysis, and the impact on the results and interpretation of the evidence.

1.1.13. Benefits and harms - qualitative review of people's experience of interventions

Evidence from one study showed that in people who did not attend specialist ME services, antidepressants were prescribed for ME/CFS symptoms by health care professionals and people experienced negative side effects, although these side effects were not described. There was very low confidence in this finding. See section 1.1.12 above for full discussion regarding antidepressants.

Evidence from one study in children/young people showed that some took prescribed sickness or stomach acid relief medication, which they found to be helpful. However, it was not common to have been offered medication to relieve their symptoms which frustrated some. There was very low confidence in this finding. Evidence from one study in children/young people showed that they generally did not mind taking medication providing they found it helpful. There was very low confidence in this finding. The committee considered that this qualitative evidence was too limited to support any recommendations.

1.1.14. Overall summary for pharmacological interventions for ME/CFS

Overall, the evidence for pharmacological interventions is limited and most was of low to very low quality and the committee was not confident in the effects. There is little evidence for most of the interventions identified and little evidence of clinical benefit and some evidence of harm. After discussing the clinical effectiveness of pharmacological interventions and people's experiences and considering the reports from the young people (see Appendix 1: Children and Young people) and people with severe ME/CFS (see Appendix 2: People with severe ME/CFS) the committee agreed there is no current pharmacological cure for ME/CFS. The committee discussed the claims that have been made about cures for people with ME/CFS and lack of evidence for this. The committee were aware of interventions that are promoted as cures and there is often a financial cost to people with ME/CFS when these are pursued. To address this the committee made a recommendation to raise awareness

that there is no current pharmacological cure for people with ME/CFS. In addition, the committee made a clear recommendation not to offer any medicines or supplements to cure ME/CFS.

The committee acknowledged that while there are not any current pharmacological cures for ME/CFS, people with ME/CFS have found some drugs when used appropriately with advice and support from health care professionals can be helpful in managing the symptoms of ME/CFS and they could be discussed on an individual basis.

1.1.15. Cost effectiveness and resource use

There were no published economic evaluations of pharmacological treatment of ME/CFS.

The annual cost of the drugs per patient that have been trialled range from only a few pounds to thousands of pounds a year.

Since there was no good quality evidence of clinical effectiveness for any of the drugs trialled, their cost effectiveness remains unproven.

Therefore, the committee did not recommend any drugs, other than those for the treatment of symptoms as recommended in other guidelines.

1.1.16. Other factors the committee took into account

The committee noted that no clinical or cost effectiveness evidence was identified for interventions evaluating some of the drugs that have been commonly used in people with ME/CFS, for example thyroxine. The committee was aware of people with ME/CFS who have been given thyroxine and other thyroid supplements as a treatment for ME/CFS fatigue and noted there is no evidence for it use in people with ME/CFS.

Medicines management

The committee highlighted that in their clinical experience people with ME/CFS may be more intolerant of drug treatment than people who do not have ME/CFS. The committee agreed it was important to raise awareness to clinicians of possible greater intolerance in this group in order to allow consideration when medications are being prescribed and taken, especially as people with ME/CFS may not initially know they are sensitive to medicines. Therefore, the committee made a recommendation to be aware that people with ME/CFS may be more intolerant of drug treatment. The committee discussed using a cautious approach to medicines prescribing, which includes starting the medicine at a lower dose than in usual clinical practice and monitoring how the person responds before adjusting the dose. The committee agreed that this type of approach would reduce the risk of harm and recommended that it be considered. The committee agreed it was important that medicines management was tailored to the person with ME/CFS and they could not provide detailed advice on how to manage intolerance.

The committee discussed medicines management for children and young people. Committee members who had experience of general paediatric services expressed that ME/CFS specialists were better placed to deliver care in this context than paediatricians. It was considered by the committee that prescribing should be initiated under the supervision of a paediatrician with expertise in ME/CFS and made a consensus based recommendation. It was acknowledged that the current availability of paediatric specialist care is limited. The committee considered whether a lack of access by GPs to specialist ME/CFS paediatricians may result in children and young people with ME/CFS being prevented from accessing medicines. However, it was agreed that telephone supervision/consultation and shared care protocols would help to overcome this. It was also agreed that continuation of prescribing by a specialist ME/CFS paediatrician may not be necessary and the committee noted that

prescribing may be continued in primary care, depending on the preferences of the patient and their carers and local circumstances.

Appendices

Appendix A Review protocols

Review protocol for pharmacological interventions

ID	Field	Content
	Scope	Management of ME/CFS
	Draft review question	3.1 What is the clinical and cost effectiveness of pharmacological interventions for people with ME/CFS?
0.	PROSPERO registration number	Not registered.
1.	Review title	What is the clinical and cost-effectiveness of pharmacological interventions for people with ME/CFS?
2.	Review question	What is the clinical effectiveness, cost-effectiveness and acceptability (including patient experiences) of pharmacological interventions for people with ME/CFS.
3.	Objective	Intervention review • To identify the most clinically and cost-effective pharmacological methods to improve outcomes in adults and children with a diagnosis of ME/CFS Qualitative review • To identify the experiences of people who have had pharmacological interventions for ME/CFS.
4.	Searches	The following databases will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR)

		• Embase • MEDLINE • Cinahl • PsychInfo Searches will be restricted by: • English language • Human studies • Letters and comments are excluded. Other searches: • Inclusion lists of relevant systematic reviews will be checked by the reviewer. The searches may be re-run 6 weeks before the final committee meeting and further studies retrieved for inclusion if relevant. The full search strategies will be published in the final review
5.	Condition or domain being studied	ME/CFS

6.	Population	Adults, children and young people who are diagnosed as having ME/CFS.

7.	Intervention/Exposure/Test
	
	(intervention review)
		Mode of delivery, dose and duration of drug treatment are not pre-specified in this protocol. This is partly because as there are no known drug liscenced fo use in ME/CFS we are interested in evaluating different drug parameters. Furthermore, because this question is intended to cover any pharmaceutical treatments evaluated by RCTs in this population, we cannot possibly list treatment parameters for all drugs we might encounter.
		
		These can include (but are not restricted to):
		
		• Antidepressants
		  ○ Include all SSRIs / SNRIs and tricyclics
		• Immunomodulatory drugs. For example:
		  ○ Rintatolimod (Ampligen)
		  ○ Rituximab
		• Pro-inflammatory cytokines. For example:
		  ○ Anakinra
		• Sleep medication. For example:
		  ○ Melatonin
		• Pain relief. For example:
		  ○ Pregabalin
		  ○ Gabapentin
		  ○ cannabinoids
		• Antiviral drugs
		• Oral corticosteroids
		  ○ fludrocortisone / hydrocortisone / other steroids
		• Modafinil
		• Sodium Valproate

		• Low dose Naltrexone
8.	Comparator/Reference standard/Confounding factors (intervention review)	• No treatment • Each other (both within and between classes) • Placebo/control/usual care
9.	Phenomena of interest (qualitative review)	The perceptions of people that have had pharmacological interventions for ME/CFS and about the benefits and harms they experienced.
11.	Types of study to be included	<u>Intervention review</u> • Randomised controlled trials • Systematic review of randomised controlled trials. For a systematic review to be included it must be conducted to the same methodological standard as NICE guideline reviews. If sufficient details are not provided to include a relevant systematic review, the review will be used for citation searching. Cross-over RCTs will be considered provided wash-out period is considered adequate. Non RCTs will not be considered as they will yield data that is at too high a risk of bias for decision-making <u>Qualitative review</u> Qualitative studies (e.g. transcript data collected from focus groups / semi structured interviews) and surveys
11.	Other exclusion criteria	Non-English language studies.

		Abstracts will be excluded as it is expected there will be sufficient full text published studies available.
12.	Context	N/A

13.	Primary outcomes (critical outcomes)	<u>Intervention review</u>
		Longest follow up available:
		
		CRITICAL OUTCOMES:
		• Mortality
		• Quality of life (any validated scales). For example:
		  o SF36
		  o EQ5D
		• General symptom scales (any validated scales). For example:
		  o De Paul Symptom Questionnaire
		  o Self Rated Clinical Global Impression Change Score
		• Fatigue/fatiguability (any validated scales). For example:
		  o Chalder fatigue Scale
		  o Fatigue Severity Scale
		  o Fatigue Impact scale
		• Physical functioning (any validated scales). For example:
		  o SF36 physical function
		  o SF36 PCS
		• Cognitive function (any validated scales). For example:
		  o MMSE
		• Psychological status (any validated scales). For example:
		  o Hospital Anxiety and Depression Scale
		  o Becks Depression Inventory
		• Pain (VAS/NRS)

		• Sleep quality (any validated scales). For example:
		  ○ Pittsburgh Sleep quality Index
		  ○ Epworth Sleepiness Scale
		  ○ Leeds Sleep Evaluation Questionnaire VAS
		• Treatment-related adverse effects
		• Activity levels – step counts
		• Return to school / work
		• Exercise performance measures. For example:
		  ○ Hand grip
		  ○ Maximal Cycle Exercise Capacity
		  ○ 6 min walk
		  ○ Timed Up and Go
		  ○ 5 repetition sit to stand
		  ○ 40m walk speed
		  ○ Step test
		
		<u>Qualitative review</u>
		Themes emerging from qualitative data

14.	Secondary outcomes (important outcomes)	<u>Intervention review</u> • Care needs • Impact on families and carers
15.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. Titles and/or abstracts of studies retrieved using the search strategy and those from additional sources will be screened for inclusion. The full text of potentially eligible studies will be retrieved and will be assessed for eligibility in line with the criteria outlined above. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer.

		Intervention review An in-house developed database; EviBase, will be used for data extraction. A standardised form is followed to extract data from studies (see Developing NICE guidelines: the manual section 6.4) and for undertaking assessment of study quality. Summary evidence tables will be produced including information on: study setting; study population and participant demographics and baseline characteristics; details of the intervention and control interventions; study methodology' recruitment and missing data rates; outcomes and times of measurement; critical appraisal ratings. Qualitative review A standardised form will be used to extract data from studies (see Developing NICE guidelines: the manual section 6.4). A second reviewer will quality-assure the extracted data. Discrepancies will be identified and resolved through discussion (with a third reviewer where necessary).
16.	Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. For the intervention review the following checklist will be used according to study design being assessed: • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) For the qualitative review the CASP qualitative checklist will be used to assess risk of bias of individual studies. 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking:

		• papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
17.	Strategy for data synthesis	<u>Intervention review</u> Where possible, data will be meta-analysed. Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5) to combine the data given in all studies for each of the outcomes stated above. A fixed effect meta-analysis, with weighted mean differences for continuous outcomes and risk ratios for binary outcomes will be used, and 95% confidence intervals will be calculated for each outcome. Heterogeneity between the studies in effect measures will be assessed using the I^2 statistic and visually inspected. We will consider an I^2 value greater than 50% indicative of substantial heterogeneity. Sensitivity analyses will be conducted based on pre-specified subgroups using stratified meta-analysis to explore the heterogeneity in effect estimates. If this does not explain the heterogeneity, the results will be presented using random-effects. GRADE pro will be used to assess the quality of each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) will be appraised for each outcome. Indirectness: If the population included in an individual study includes children aged under 12, it will be included if the majority of the population is aged over 12, and downgraded for indirectness if the overlap into those aged less than 12 is greater than 20%.

		The criteria used to diagnose people with CFS/ME should include post exertional malaise (PEM) as a compulsory feature. If the criteria does not include PEM the population will be downgraded for indirectness. Publication bias is tested for when there are more than 5 studies for an outcome. Other bias will only be taken into consideration in the quality assessment if it is apparent. Where meta-analysis is not possible, data will be presented and quality assessed individually per outcome. If sufficient data is available to make a network of treatments, WinBUGS will be used for network meta-analysis. <u>Qualitative review</u> The synthesis of qualitative data will follow a thematic analysis approach. Information will be synthesised into main review findings. Results will be presented in a detailed narrative and in table format with summary statements of main review findings. GRADE CERQual will be used to synthesise the qualitative data and assess the certainty of evidence for each review finding.
18.	Analysis of sub-groups	<u>Stratification:</u> Age: children and young people vs adults Severity: severe vs moderate as defined by the studies Where populations are mixed/unclear, these will be analysed in mixed/unclear population strata.

		Subgroups to investigate if heterogeneity is present None		
19.	Type and method of review	☒ Intervention ☐ Diagnostic ☐ Prognostic ☒ Qualitative ☐ Epidemiologic ☐ Service Delivery ☐ Other (please specify)		
20.	Language	English		
20.	Country	England		
21.	Anticipated or actual start date	01/01/20		
22.	Anticipated completion date	01/01/21		
23.	Stage of review at time of this submission	**Review stage**	**Started**	**Completed**
		Preliminary searches	☑	☐

		Piloting of the study selection process	☑	☐
		Formal screening of search results against eligibility criteria	☑	☐
		Data extraction	☑	☐
		Risk of bias (quality) assessment	☑	☐
		Data analysis	☐	☐
24.	Named contact	**5a. Named contact** National Guideline Centre **5b Named contact e-mail** **5e Organisational affiliation of the review** National Institute for Health and Care Excellence (NICE) and the National Guideline Centre		
25.	Review team members	From the National Guideline Centre: • Dr Kate Kelley [Guideline lead] • Ms Maria Smyth [Senior systematic reviewer] • Ms Melina Vasileiou [Systematic reviewer]		

		• Dr Richard Clubbe [Systematic reviewer] • Dr Karin van Bart [Systematic reviewer] • Mr David Wonderling [Health economist] • Ms Agnes Cuyas [Information specialist] • Ms Kate Ashmore [Project manager]
26.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10091
29.	Other registration details	N/A
30.	Reference/URL for published protocol	N/A
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication

		• publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
32.	Keywords	
33.	Details of existing review of same topic by same authors	N/A
34.	Current review status	☒ Ongoing ☐ Completed but not published ☐ Completed and published ☐ Completed, published and being updated ☐ Discontinued
35..	Additional information	N/A
36.	Details of final publication	www.nice.org.uk

Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.

Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2004, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[59] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies.

	Setting: • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2004 or later but that depend on unit costs and resource data entirely or predominantly from before 2004 will be rated as 'Not applicable'. • Studies published before 2004 will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B Literature search strategies

This literature search strategy was used for the following review questions:

- What is the clinical effectiveness, cost-effectiveness and acceptability (including patient experiences) of pharmacological interventions for people with ME/CFS?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[59]

For more information, please see the Methodology review published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve.

Searches for patient views were run in Medline (OVID), Embase (OVID), CINAHL, and PsycINFO (ProQuest).

Table 20: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 23 June 2020	Exclusions
Embase (OVID)	1974 – 23 June 2020	Exclusions
The Cochrane Library (Wiley)	Cochrane Reviews to 2020 Issue 6 of 12 CENTRAL to 2020 Issue 6 of 12	None
CINAHL, Current Nursing and Allied Health Literature (EBSCO)	Inception – 23 June 2020	None
PsycINFO (ProQuest)	Inception – 23 June 2020	Exclusions
Epistemonikos (The Epistemonikos Foundation)	Inception - 23 June 2020	None

Medline (Ovid) search terms

1.	Fatigue Syndrome, Chronic/
2.	chronic* fatigue*.ti,ab.
3.	(fatigue* adj2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*)).ti,ab.
4.	((myalgic or post infection* or postinfection*) adj (encephalomyelitis or encephalopathy)).ti,ab.
5.	((ME adj CFS) or (CFS adj ME) or CFIDS or PVFS).ti,ab.
6.	(Systemic Exertion Intolerance Disease or SEID).ti,ab.
7.	((CFS adj SEID) or (SEID adj CFS) or (ME adj CFS adj SEID) or (ME adj SEID) or (SEID adj ME)).ti,ab.
8.	((Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome or POTS) adj6 (CFS or chronic* fatigue* or ME or myalgic or SEID or systemic exertion)).ti,ab.

9.	((Post-exertional or postexertional) adj2 malaise).ti,ab.
10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia or neurasthenia).ti,ab.
11.	((atypical or simulating or resembling) adj poliomyelitis).ti,ab.
12.	((chronic adj2 epstein Barr virus) or CEBV or CAEBV or chronic mononucleosis).ti,ab.
13.	xenotropic murine leukemia virus-related virus.ti,ab.
14.	effort syndrome*.ti,ab.
15.	(((akureyri or iceland or tapanui or royal free or royal free hospital) adj disease*) or ((yuppie or yuppy or tapanui) adj flu)).ti,ab.
16.	or/1-15
17.	letter/
18.	editorial/
19.	news/
20.	exp historical article/
21.	Anecdotes as Topic/
22.	comment/
23.	case report/
24.	(letter or comment*).ti.
25.	or/17-24
26.	randomized controlled trial/ or random*.ti,ab.
27.	25 not 26
28.	animals/ not humans/
29.	exp Animals, Laboratory/
30.	exp Animal Experimentation/
31.	exp Models, Animal/
32.	exp Rodentia/
33.	(rat or rats or mouse or mice).ti.
34.	or/27-33
35.	16 not 34
36.	limit 35 to English language

Embase (Ovid) search terms

1.	chronic fatigue syndrome/
2.	chronic* fatigue*.ti,ab.
3.	(fatigue* adj2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*)).ti,ab.
4.	((myalgic or post infection* or postinfection*) adj (encephalomyelitis or encephalopathy)).ti,ab.
5.	((ME adj CFS) or (CFS adj ME) or CFIDS or PVFS).ti,ab.
6.	(Systemic Exertion Intolerance Disease or SEID).ti,ab.
7.	((CFS adj SEID) or (SEID adj CFS) or (ME adj CFS adj SEID) or (ME adj SEID) or (SEID adj ME)).ti,ab.
8.	((Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome or POTS) adj6 (CFS or chronic* fatigue* or ME or myalgic or SEID or systemic exertion)).ti,ab.
9.	((Post-exertional or postexertional) adj2 malaise).ti,ab.

10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia or neurasthenia).ti,ab.
11.	((atypical or simulating or resembling) adj poliomyelitis).ti,ab.
12.	((chronic adj2 epstein Barr virus) or CEBV or CAEBV or chronic mononucleosis).ti,ab.
13.	xenotropic murine leukemia virus-related virus.ti,ab.
14.	effort syndrome*.ti,ab.
15.	(((akureyri or iceland or tapanui or royal free or royal free hospital) adj disease*) or ((yuppie or yuppy or tapanui) adj flu)).ti,ab.
16.	or/1-15
17.	letter.pt. or letter/
18.	note.pt.
19.	editorial.pt.
20.	case report/ or case study/
21.	(letter or comment*).ti.
22.	or/17-21
23.	randomized controlled trial/ or random*.ti,ab.
24.	22 not 23
25.	animal/ not human/
26.	nonhuman/
27.	exp Animal Experiment/
28.	exp Experimental Animal/
29.	animal model/
30.	exp Rodent/
31.	(rat or rats or mouse or mice).ti.
32.	or/24-31
33.	16 not 32
34.	limit 33 to English language

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Fatigue Syndrome, Chronic] this term only
#2.	chronic* fatigue*:ti,ab
#3.	(fatigue* near/2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*)):ti,ab
#4.	((myalgic or post infection* or postinfection*) near/1 (encephalomyelitis or encephalopathy)):ti,ab
#5.	((ME near/1 CFS) or (CFS near/1 ME) or CFIDS or PVFS):ti,ab
#6.	(Systemic Exertion Intolerance Disease or SEID):ti,ab
#7.	((CFS near/1 SEID) or (SEID near/1 CFS) or (ME near/1 CFS near/1 SEID) or (ME near/1 SEID) or (SEID near/1 ME)):ti,ab
#8.	(Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome or POTS)
#9.	((Post-exertional or postexertional) near/2 malaise):ti,ab
#10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia or neurasthenia):ti,ab
#11.	((atypical or simulating or resembling) near/1 poliomyelitis):ti,ab
#12.	((chronic epstein Barr virus) or CEBV or CAEBV or chronic mononucleosis):ti,ab
#13.	xenotropic murine leukemia virus-related virus:ti,ab
#14.	effort syndrome*:ti,ab

#15.	((akureyri or iceland or tapanui or "royal free" or "royal free hospital") near/1 disease*):ti,ab
#16.	((yuppie or yuppy or tapanui) near flu):ti,ab
#17.	(or #1-#16)

CINAHL (EBSCO) search terms

S1.	(MH "Fatigue Syndrome, Chronic")
S2.	chronic* fatigue*
S3.	(fatigue* n2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*))
S4.	((myalgic or post infection* or postinfection*) and (encephalomyelitis or encephalopathy))
S5.	((ME and CFS) or (CFS and ME) or CFIDS or PVFS)
S6.	(Systemic Exertion Intolerance Disease or SEID)
S7.	((CFS and SEID) or (SEID and CFS) or (ME and CFS and SEID) or (CFS and ME and SEID) or (ME and SEID) or (SEID and ME))
S8.	((Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome) and (CFS or chronic* fatigue* or ME or myalgic or SEID or systemic exertion))
S9.	((Post-exertional or postexertional) n2 malaise)
S10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia)
S11.	((atypical or simulating or resembling) and poliomyelitis)
S12.	(chronic epstein Barr virus or chronic mononucleosis)
S13.	xenotropic murine leukemia virus-related virus
S14.	effort syndrome*
S15.	(((akureyri or iceland or tapanui or royal free or royal free hospital) and disease*) or ((yuppie or yuppy or tapanui) and flu))
S16.	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15

PsycINFO (ProQuest) search terms

1.	((((chronic* fatigue*) OR (fatigue* NEAR2 (disorder* OR syndrome* OR post viral OR postviral OR immune dysfunction* OR post infection* OR postinfection*)) OR ((myalgic OR post infection* OR postinfection*) NEAR1 (encephalomyelitis OR encephalopathy)) OR ((ME NEAR1 CFS) OR (CFS NEAR1 ME) OR CFIDS OR PVFS) OR (Systemic Exertion Intolerance Disease OR SEID) OR ((CFS NEAR1 SEID) OR (SEID NEAR1 CFS)) OR ((ME NEAR1 CFS NEAR1 SEID) OR (ME NEAR1 SEID) OR (SEID NEAR1 ME)) OR ((Orthostatic intolerance OR postural orthostatic tachycardia syndrome OR postural tachycardia syndrome OR POTS) NEAR6 (CFS OR chronic* fatigue* OR ME OR myalgic OR SEID OR systemic exertion)) OR (neurasthenic neuroses OR epidemic neuromyasthenia OR neurataxia OR neuroasthenia OR neurasthenia) OR ((atypical OR simulating OR resembling) NEAR1 poliomyelitis)) OR (((chronic NEAR2 epstein Barr virus) OR CEBV OR CAEBV OR chronic mononucleosis) OR (xenotropic murine leukemia virus-related virus) OR (effort syndrome*) OR ((akureyri OR iceland OR tapanui OR royal free OR royal free hospital) NEAR1 disease*) OR ((yuppie OR yuppy OR tapanui) NEAR1 flu) OR MAINSUBJECT.EXACT.EXPLODE("Chronic Fatigue Syndrome"))) AND (stype.exact("Scholarly Journals") AND la.exact("ENG") AND po.exact("Human") NOT (me.exact("Empirical Study" OR "Quantitative Study" OR "Longitudinal Study" OR "Clinical Trial" OR "Qualitative Study" OR "Prospective Study" OR "Followup Study" OR "Literature Review" OR "Retrospective Study" OR "Systematic Review" OR "Meta Analysis") AND po.exact("Human"))

Epistemonikos search terms

1.	(advanced_title_en:((advanced_title_en:((chronic* fatigue* syndrome*) OR (fatigue* syndrome* OR fatigue* disorder* OR postviral fatigue* OR post viral fatigue* OR

	fatigue* immune dysfunction OR post infection fatigue* OR postinfection fatigue*) OR (encephalomyelitis OR encephalopathy) OR ("ME/CFS" OR "CFS/ME" OR "CFIDS" OR "PVFS") OR (Systemic Exertion Intolerance Disease OR SEID) OR ((CFS AND SEID) OR (SEID AND CFS) OR (ME AND CFS AND SEID) OR (ME AND SEID) OR (SEID AND ME)) OR (Orthostatic intolerance OR postural orthostatic tachycardia syndrome OR postural tachycardia syndrome OR POTS) OR ((Post-exertional OR postexertional) AND malaise) OR (neurasthenic neuroses OR epidemic neuromyasthenia OR neurataxia OR neuroasthenia OR neurasthenia) OR (atypical poliomyelitis OR simulating poliomyelitis OR resembling poliomyelitis) OR (chronic epstein Barr virus OR CEBV OR CAEBV OR chronic mononucleosis) OR (xenotropic murine leukemia virus-related virus) OR (effort syndrome*) OR (akureyri OR iceland disease OR tapanui OR royal free disease) OR (yuppie flu OR yuppy flu OR tapanui flu)) OR advanced_abstract_en:((chronic* fatigue* syndrome*) OR (fatigue* syndrome* OR fatigue* disorder* OR postviral fatigue* OR post viral fatigue* OR fatigue* immune dysfunction OR post infection fatigue* OR postinfection fatigue*) OR (encephalomyelitis OR encephalopathy) OR ("ME/CFS" OR "CFS/ME" OR "CFIDS" OR "PVFS") OR (Systemic Exertion Intolerance Disease OR SEID) OR ((CFS AND SEID) OR (SEID AND CFS) OR (ME AND CFS AND SEID) OR (ME AND SEID) OR (SEID AND ME)) OR (Orthostatic intolerance OR postural orthostatic tachycardia syndrome OR postural tachycardia syndrome OR POTS) OR ((Post-exertional OR postexertional) AND malaise) OR (neurasthenic neuroses OR epidemic neuromyasthenia OR neurataxia OR neuroasthenia OR neurasthenia) OR (atypical poliomyelitis OR simulating poliomyelitis OR resembling poliomyelitis) OR (chronic epstein Barr virus OR CEBV OR CAEBV OR chronic mononucleosis) OR (xenotropic murine leukemia virus-related virus) OR (effort syndrome*) OR (akureyri OR iceland disease OR tapanui OR royal free disease) OR (yuppie flu OR yuppy flu OR tapanui flu)))) OR advanced_abstract_en:((advanced_title_en:((chronic* fatigue* syndrome*) OR (fatigue* syndrome* OR fatigue* disorder* OR postviral fatigue* OR post viral fatigue* OR fatigue* immune dysfunction OR post infection fatigue* OR postinfection fatigue*) OR (encephalomyelitis OR encephalopathy) OR ("ME/CFS" OR "CFS/ME" OR "CFIDS" OR "PVFS") OR (Systemic Exertion Intolerance Disease OR SEID) OR ((CFS AND SEID) OR (SEID AND CFS) OR (ME AND CFS AND SEID) OR (ME AND SEID) OR (SEID AND ME)) OR (Orthostatic intolerance OR postural orthostatic tachycardia syndrome OR postural tachycardia syndrome OR POTS) OR ((Post-exertional OR postexertional) AND malaise) OR (neurasthenic neuroses OR epidemic neuromyasthenia OR neurataxia OR neuroasthenia OR neurasthenia) OR (atypical poliomyelitis OR simulating poliomyelitis OR resembling poliomyelitis) OR (chronic epstein Barr virus OR CEBV OR CAEBV OR chronic mononucleosis) OR (xenotropic murine leukemia virus-related virus) OR (effort syndrome*) OR (akureyri OR iceland disease OR tapanui OR royal free disease) OR (yuppie flu OR yuppy flu OR tapanui flu)) OR advanced_abstract_en:((chronic* fatigue* syndrome*) OR (fatigue* syndrome* OR fatigue* disorder* OR postviral fatigue* OR post viral fatigue* OR fatigue* immune dysfunction OR post infection fatigue* OR postinfection fatigue*) OR (encephalomyelitis OR encephalopathy) OR ("ME/CFS" OR "CFS/ME" OR "CFIDS" OR "PVFS") OR (Systemic Exertion Intolerance Disease OR SEID) OR ((CFS AND SEID) OR (SEID AND CFS) OR (ME AND CFS AND SEID) OR (ME AND SEID) OR (SEID AND ME)) OR (Orthostatic intolerance OR postural orthostatic tachycardia syndrome OR postural tachycardia syndrome OR POTS) OR ((Post-exertional OR postexertional) AND malaise) OR (neurasthenic neuroses OR epidemic neuromyasthenia OR neurataxia OR neuroasthenia OR neurasthenia) OR (atypical poliomyelitis OR simulating poliomyelitis OR resembling poliomyelitis) OR (chronic epstein Barr virus OR CEBV OR CAEBV OR chronic mononucleosis) OR (xenotropic murine leukemia virus-related virus) OR (effort syndrome*) OR (akureyri OR iceland disease OR tapanui OR royal free disease) OR (yuppie flu OR yuppy flu OR tapanui flu)))))

B.2 Health economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to ME/CFS population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA – this ceased to be updated after March 2018), with no date restrictions. NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics.

Table 21: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2014 – 30 June 2020	Exclusions Health economics studies
Embase	2014 –30 June 2020	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	HTA - 2003 – 31 March 2018 NHSEED - 2003 to 31 March 2015	None

Medline (Ovid) search terms

1.	Fatigue Syndrome, Chronic/
2.	chronic* fatigue*.ti,ab.
3.	(fatigue* adj2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*)).ti,ab.
4.	((myalgic or post infection* or postinfection*) adj (encephalomyelitis or encephalopathy)).ti,ab.
5.	((ME adj CFS) or (CFS adj ME) or CFIDS or PVFS).ti,ab.
6.	(Systemic Exertion Intolerance Disease or SEID).ti,ab.
7.	((CFS adj SEID) or (SEID adj CFS) or (ME adj CFS adj SEID) or (ME adj SEID) or (SEID adj ME)).ti,ab.
8.	((Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome or POTS) adj6 (CFS or chronic* fatigue* or ME or myalgic or SEID or systemic exertion)).ti,ab.
9.	((Post-exertional or postexertional) adj2 malaise).ti,ab.
10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia or neurasthenia).ti,ab.
11.	((atypical or simulating or resembling) adj poliomyelitis).ti,ab.
12.	((chronic adj2 epstein Barr virus) or CEBV or CAEBV or chronic mononucleosis).ti,ab.
13.	xenotropic murine leukemia virus-related virus.ti,ab.
14.	effort syndrome*.ti,ab.
15.	(((akureyri or iceland or tapanui or royal free or royal free hospital) adj disease*) or ((yuppie or yuppy or tapanui) adj flu)).ti,ab.
16.	or/1-15
17.	letter/
18.	editorial/
19.	news/
20.	exp historical article/
21.	Anecdotes as Topic/

22.	comment/
23.	case report/
24.	(letter or comment*).ti.
25.	or/17-24
26.	randomized controlled trial/ or random*.ti,ab.
27.	25 not 26
28.	animals/ not humans/
29.	exp Animals, Laboratory/
30.	exp Animal Experimentation/
31.	exp Models, Animal/
32.	exp Rodentia/
33.	(rat or rats or mouse or mice).ti.
34.	or/27-33
35.	16 not 34
36.	limit 35 to English language
37.	Economics/
38.	Value of life/
39.	exp "Costs and Cost Analysis"/
40.	exp Economics, Hospital/
41.	exp Economics, Medical/
42.	Economics, Nursing/
43.	Economics, Pharmaceutical/
44.	exp "Fees and Charges"/
45.	exp Budgets/
46.	budget*.ti,ab.
47.	cost*.ti.
48.	(economic* or pharmaco?economic*).ti.
49.	(price* or pricing*).ti,ab.
50.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
51.	(financ* or fee or fees).ti,ab.
52.	(value adj2 (money or monetary)).ti,ab.
53.	or/37-52
54.	36 and 53

Embase (Ovid) search terms

1.	chronic fatigue syndrome/
2.	chronic* fatigue*.ti,ab.
3.	(fatigue* adj2 (disorder* or syndrome* or post viral or postviral or immune dysfunction* or post infection* or postinfection*)).ti,ab.
4.	((myalgic or post infection* or postinfection*) adj (encephalomyelitis or encephalopathy)).ti,ab.
5.	((ME adj CFS) or (CFS adj ME) or CFIDS or PVFS).ti,ab.
6.	(Systemic Exertion Intolerance Disease or SEID).ti,ab.
7.	((CFS adj SEID) or (SEID adj CFS) or (ME adj CFS adj SEID) or (ME adj SEID) or (SEID adj ME)).ti,ab.

8.	((Orthostatic intolerance or postural orthostatic tachycardia syndrome or postural tachycardia syndrome or POTS) adj6 (CFS or chronic* fatigue* or ME or myalgic or SEID or systemic exertion)).ti,ab.
9.	((Post-exertional or postexertional) adj2 malaise).ti,ab.
10.	(neurasthenic neuroses or epidemic neuromyasthenia or neurataxia or neuroasthenia or neurasthenia).ti,ab.
11.	((atypical or simulating or resembling) adj poliomyelitis).ti,ab.
12.	((chronic adj2 epstein Barr virus) or CEBV or CAEBV or chronic mononucleosis).ti,ab.
13.	xenotropic murine leukemia virus-related virus.ti,ab.
14.	effort syndrome*.ti,ab.
15.	(((akureyri or iceland or tapanui or royal free or royal free hospital) adj disease*) or ((yuppie or yuppy or tapanui) adj flu)).ti,ab.
16.	or/1-15
17.	letter.pt. or letter/
18.	note.pt.
19.	editorial.pt.
20.	case report/ or case study/
21.	(letter or comment*).ti.
22.	or/17-21
23.	randomized controlled trial/ or random*.ti,ab.
24.	22 not 23
25.	animal/ not human/
26.	nonhuman/
27.	exp Animal Experiment/
28.	exp Experimental Animal/
29.	animal model/
30.	exp Rodent/
31.	(rat or rats or mouse or mice).ti.
32.	or/24-31
33.	16 not 32
34.	limit 33 to English language
35.	health economics/
36.	exp economic evaluation/
37.	exp health care cost/
38.	exp fee/
39.	budget/
40.	funding/
41.	budget*.ti,ab.
42.	cost*.ti.
43.	(economic* or pharmaco?economic*).ti.
44.	(price* or pricing*).ti,ab.
45.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
46.	(financ* or fee or fees).ti,ab.
47.	(value adj2 (money or monetary)).ti,ab.
48.	or/35-47

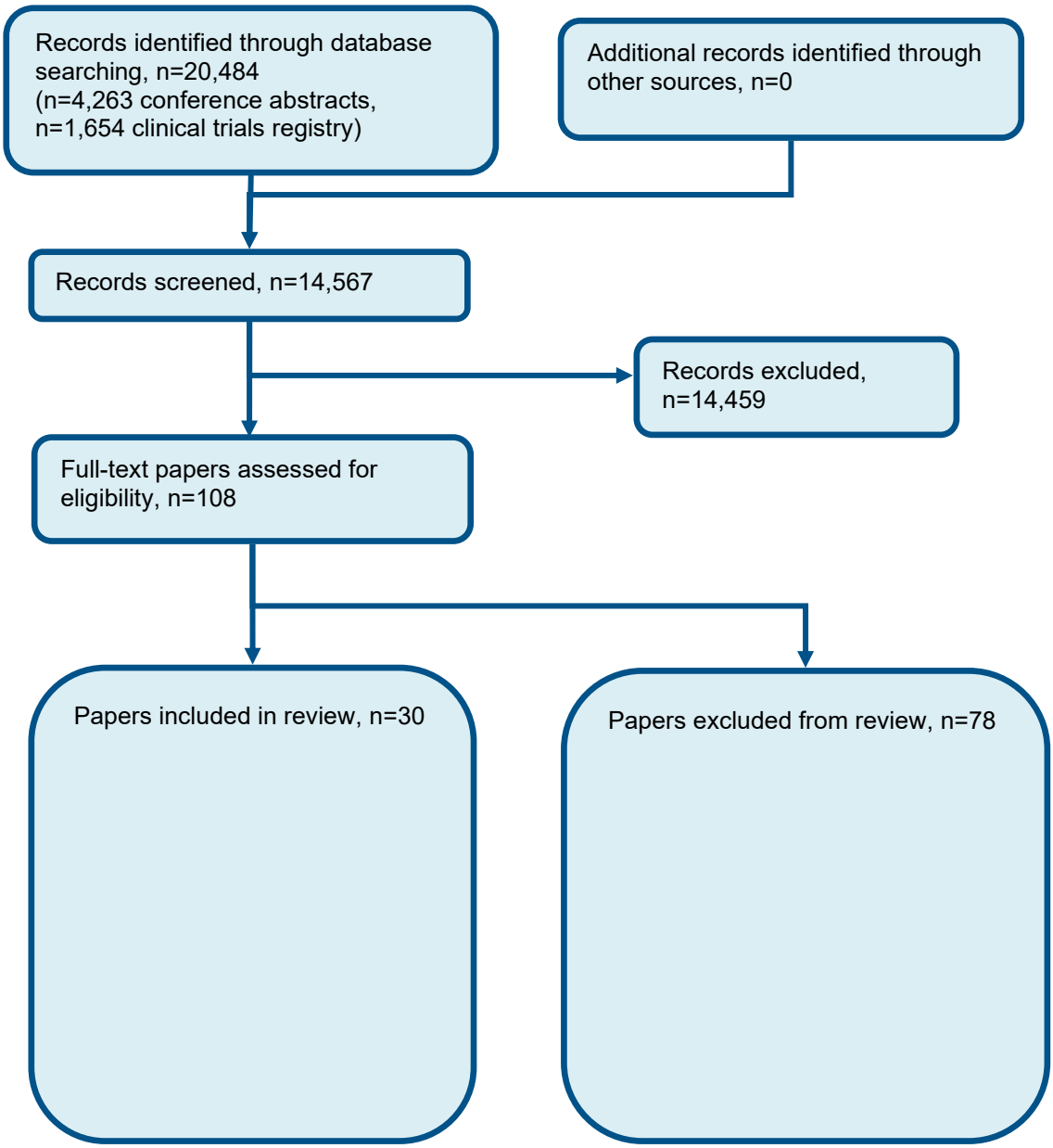
49.	34 and 48

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Fatigue Syndrome, Chronic
#2.	(chronic fatigue or fatigue syndrome*)
#3.	((myalgic adj (encephalomyelitis or encephalopathy)))
#4.	(((ME adj CFS) or (CFS adj ME)))
#5.	(post viral fatigue or post viral syndrome* or viral fatigue syndrome* or PVFS)
#6.	#1 OR #2 OR #3 OR #4 OR #5
#7.	(neurasthenic neuroses or epidemic neuromyasthenia or post infectious encephalomyelitis or neurataxia or neuroasthenia)
#8.	(((atypical or simulating or resembling) adj poliomyelitis))
#9.	(chronic epstein Barr virus or chronic mononucleosis)
#10.	(xenotropic murine leukemia virus-related virus)
#11.	(((chronic fatigue and immune dysfunction syndrome*) or cfids or chronic fatigue-fibromyalgia syndrome* or chronic fatigue disorder* or Systemic Exertion Intolerance Disease or SEID or effort syndrome or post infectious fatigue))
#12.	((((akureyri or iceland or tapanui or royal free or royal free hospital) adj disease*) or ((yuppie or yuppy or tapanui) adj flu)))
#13.	#7 OR #8 OR #9 OR #10 OR #11 OR #12
#14.	#6 or #13

Appendix C Effectiveness evidence study selection

Figure 1: Flow chart of clinical study selection for the review of pharmacological interventions



Appendix D Effectiveness evidence

Study	Arnold 2015[4]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in USA; Setting: Single outpatient research centre setting
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Revised CDC definition of CFS
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 18-65; met revised CDC criteria for CFS: at least 6 months of persistent disabling fatigue, 4 or more of impaired memory/concentration, sore throat, tender glands, aching or stiff muscles, multi joint pain, new headaches, unrefreshing sleep and post-exertional fatigue; General fatigue score of >+13 on MFI at baseline
Exclusion criteria	Other medical disorders that could explain the fatigue; psychiatric disorders; substance abuse disorders; women who were pregnant or breastfeeding; women of childbearing potential not using contraceptives; people deemed to be refractory to treatment; people whose response was deemed to be influenced by current or future disability compensation issues; serious unstable medical illness; abnormal TSH levels; uncontrolled narrow angle glaucoma; acute liver injury/severe cirrhosis; suicidal risk; known non-responder to duloxetine; known hypersensitivity to duloxetine; any treatment with investigational drug within 30 days of screening; use of any medications or herbal agents with CNS effects (except occasional sedating antihistamines); treatment with analgesics except OTC NSAIDs and paracetamol; unconventional or alternative therapies.

Recruitment/selection of patients	consecutive referral or via advertisement
Age, gender and ethnicity	Age - Mean (range): duloxetine/placebo: 43/44.3. Gender (M:F): 13.3:86.7. Ethnicity: Duloxetine/placebo: white 86.7%/83.3%; African-American 13.3%/13.3%; other 0%/3.3%
Further population details	-
Extra comments	Duloxetine/placebo: MFI general fatigue 17.3/17.3; MFI general fatigue 14.8/13.9; MFI reduced activity 14.3/14.5;MFI reduced motivation 12.3/12.6; MFI mental fatigue 15.4/15.5; Brief Pain Inventory average pain severity 4/3.8; HADS anxiety 8.1/8.8; HADS depression 6.3/9; Total CDC symptom inventory 63.7/67.8; SF36 physical functioning 63.1/55.9; role physical 25/18.5; social functioning 67.6/54.2; bodily pain 50.2/46.3; mental health 71/55.4; role emotional 74.1/38.3; vitality 17.6/16.7; general health 49.9/52.1; CGI severity - moderate 86.2%/90%
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM is not a compulsory feature
Interventions	(n=30) Intervention 1: antidepressants - serotonin-norepinephrine reuptake inhibitors. Duloxetine hydrochloride - 30mg once a day for 1 week, then 60 mg once a day for the next 3 weeks. The dose then increased to 90mg per day for next 4 weeks (as tolerated). If highest doses not tolerated doses could be as low as 60mg per day. Duration 12 weeks. Concurrent medication/care: At the end of 12 weeks patients had a 1 week tapering phase in which the drug was reduced by 30mg daily until discontinuation. Indirectness: No indirectness (n=30) Intervention 2: placebo. Identical placebo given in same way as study drug. Duration 12 weeks. Concurrent medication/care: At the end of 12 weeks patients had a 1 week tapering phase in which the placebo was reduced by 30mg daily until discontinuation. Indirectness: No indirectness
Funding	Study funded by industry (Eli Lilly and Company Investigator-Initiated Trial program.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SEROTONIN-NOREPINEPHRINE REUPTAKE INHIBITORS versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF-36 physical functioning at 12 weeks; MD; 6.8 (95%CI -8.5 to 22, SF-36 physical functioning 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 mental health at 12 weeks; MD; -1.1 (95%CI -11.8 to 9.5, SF-36 mental health 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 role emotional at 12 weeks; MD; 4.4 (95%CI -24.2 to 32.9, SF-36 role emotional 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 role physical at 12 weeks; MD; 11 (95%CI -9 to 30.9, SF-36 role physical 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 bodily pain at 12 weeks; MD; 11.4 (95%CI -0.5 to 23.2, SF-36 bodily pain 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 vitality at 12 weeks; MD; 3.3 (95%CI -10.3 to 17, SF-36 vitality 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 general health at 12 weeks; MD; 0 (95%CI -10.8 to 10.7, SF-36 general health 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: SF-36 social functioning at 12 weeks; MD; 0.7 (95%CI -14.7 to 16, SF-36 social functioning 0-100; High=Top is good outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

Protocol outcome 2: Fatigue at longest follow up available

- Actual outcome for adults; severity mixed or unclear: MFI-20 general fatigue at 12 weeks; MD; -1 (95%CI -2.8 to 0.7, Multidimensional fatigue inventory-20 general fatigue subscale range not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: MFI-20 physical fatigue at 12 weeks; MD; -0.9 (95%CI -2.7 to 0.9, Multidimensional fatigue inventory-20 physical fatigue subscale range not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: MFI-20 reduced activity at 12 weeks; MD; 0 (95%CI -1.8 to 1.8, Multidimensional fatigue inventory-20 reduced activity subscale range not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: MFI-20 reduced motivation at 12 weeks; MD; -0.8 (95%CI -2.6 to 1.1, Multidimensional fatigue inventory-20 reduced motivation subscale range not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: MFI-20 mental fatigue at 12 weeks; MD; -2.5 (95%CI -4.4 to -0.6, Multidimensional fatigue inventory-20 mental fatigue subscale range not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

Protocol outcome 3: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: HADS - anxiety at 12 weeks; MD; -0.9 (95%CI -2.4 to 0.6, Hospital anxiety and depression scale anxiety subscale 0-21, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: HADS - depression at 12 weeks; MD; 0.94 (95%CI 0.72 to 1.23, Hospital anxiety and depression scale depression subscale 0-21, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

Protocol outcome 4: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Brief Pain Inventory - severity at 12 weeks; MD; -0.73 (95%CI -1 to -0.54, Brief pain inventory severity subscale 0-10, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Brief Pain Inventory - interference at 12 weeks; MD; -0.7 (95%CI -0.96 to -0.51, Brief pain inventory interference subscale 0-10, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0

Protocol outcome 5: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Most frequently reported treatment-emergent adverse events at 12 weeks; Group 1: 131/29, Group 2: 62/30; Comments: Total number of reported treatment-emergent adverse events. Adverse events were those reported by at least 5% of patients in the treatment group. The list of adverse events reported was: nausea, somnolence, dizziness, headache, dry mouth, insomnia, constipation, cold virus, decreased appetite, diarrhoea, light headedness, anxiety, vivid dreams, increased urination, increased yawning, jittery, increased sweating, chills, depression, fever, hot flush, increased appetite, irritability,

pruritus, muscle fasciculation, abdominal pain, sinus infection, vaginal infection, weight gain.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 6: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: Clinical Global Impression of Severity at 12 weeks; MD; -0.1 (95%CI -0.3 to 0), Clinical Global Impression of Severity 1-7, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Clinical Global Impression of Improvement at 12 weeks; MD; -0.8 (95%CI -1.7 to 0), Clinical Global Impression of Improvement 1-7, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: CDC symptom inventory at 12 weeks; MD; -2.7 (95%CI -15.5 to 10.1), CDC symptom inventory scale not reported, High=Top is poor outcome. Comments: Model based estimate is the MD in changes from baseline, with adjustment for other covariates.);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Systematically higher SF-36 indices for the duloxetine group. Other outcomes similar.; Blinding details: Self-administered so possible lack of assessor blinding may not have been overly important; Group 1 Number missing: 10, Reason: unclear; Group 2 Number missing: 4, Reason: unclear

Protocol outcomes not reported by the study	Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Blacker 2004[8]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=434)

Countries and setting	Conducted in USA; Setting: Most patients recruited from primary care centres, and some from tertiary care centres
Line of therapy	Not applicable
Duration of study	Follow up (post intervention): 20 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Modified US centers for disease control and prevention diagnosis for CFS
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	age 18-65 years; modified US CDC diagnosis for CFS; illness duration; 7 years; those with concurrent fibromyalgia also included.
Exclusion criteria	Concurrent psychiatric diagnoses; any inpatient psychiatric care; previous suicide; IBS; peptic ulcer; severe asthma; endocrine or metabolic disease; HIV; neurological disease; sensitivity to cholinergic agents; exposure to organophosphates; Gulf war syndrome; participation in CBT or GET programmes during the study; pregnancy; concomitant medication during trial except minor analgesics; antidepressants or cholinergics or antihypertensives or corticosteroids or antihistamines within 3 months prior to trial onset; other psychotropic medication within 6 weeks prior to study onset; Domperidone was allowed for anti-emetic use
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 37-39.1. Gender (M:F): 34:66. Ethnicity: White 96%; Black 1.4%; Indian subcontinent 0.5%; Asian 0.025%; Hispanic 2%
Further population details	-
Extra comments	baseline values not provided
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature

Interventions	(n=352) Intervention 1: Galantamine hydrobromide. 3 x 2.5mg per day or 3 x 5mg per day or 3 x 7.5mg per day or 3 x 10mg per day. Duration 16 weeks. Concurrent medication/care: Titrated over a 3-8 week period, commencing at 2.5mg/day, with weekly increments of 2.5 to 7.5mg depending on dose. Target dose maintained for final 8 weeks at least. Indirectness: No indirectness (n=82) Intervention 2: placebo. 3 x daily. Duration 16 weeks. Concurrent medication/care: Titration details not clear. Indirectness: No indirectness
Funding	Study funded by industry (Shire Pharmaceutical Development Ltd)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GALANTAMINE HYDROBROMIDE versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Chalder fatigue rating scale-physical at 20 weeks; Mean; , Comments: Only mean change from baseline given for the placebo (9.86) and the 4 dose sub-groups (8.77 to 11.02). No measures of variance so not possible to estimate 95% CIs;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Baselines given as range across all arms; result is change from baseline.; Group 1 Number missing: 9, Reason: unclear; Group 2 Number missing: 3, Reason: unclear
- Actual outcome for adults; severity mixed or unclear: Chalder fatigue rating scale-mental at 20 weeks; Mean; , Comments: Only mean change from baseline given for the placebo (6.8) and the 4 dose sub-groups (5.89 to 7.74). No measures of variance so not possible to estimate 95% CIs;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Baselines given as range across all arms; result is change from baseline.; Group 1 Number missing: 9, Reason: unclear; Group 2 Number missing: 3, Reason: unclear

Protocol outcome 2: Cognitive function at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Computerised cognitive test at 20 weeks; Mean; , Comments: For each of the sub-tests only mean changes from baseline were given without any measure of variance. The values are not given here, as they cannot be usefully used in a meta-analysis; not possible to estimate 95% CIs;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Baselines given as range across all arms; result is change from baseline.; Group 1 Number missing: 9, Reason: unclear; Group 2 Number missing: 3, Reason: unclear

Protocol outcome 3: Sleep quality at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Pittsburgh Sleep Quality Index - total score at 20 weeks; Mean; , Comments: Only mean change from baseline

given for the placebo (-2.02) and the 4 dose sub-groups (-2.28 to -1.43). No measures of variance so not possible to estimate 95% CIs;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Baselines given as range across all arms; result is change from baseline.; Group 1 Number missing: 9, Reason: unclear; Group 2 Number missing: 3, Reason: unclear

Protocol outcome 4: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: Clinician global impression scores - no change or worse (>=3) at 20 weeks; Group 1: 169/280, Group 2: 47/67
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: No baseline details given; Group 1 Number missing: 72, Reason: mostly adverse events but generally unclear if related to outcome; Group 2 Number missing: 15, Reason: adverse events but generally unclear if related to outcome

Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Blockmans 2006[9]**
Study type	RCT (Patient randomised; Crossover: 1 week (half-life of drug = 2 hours, so likely to be appropriate))
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Belgium; Setting: General Internal medicine Outpatient clinic at a University Hospital in Gasthuisberg, Belgium.
Line of therapy	Not applicable
Duration of study	Follow up (post intervention): 4 weeks for each period
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1994 CDC CFS criteria

Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	1994 CDC criteria; at least 4 minor criteria; at least 6 months of fatigue; ambulant; concentration problems mandatory;
Exclusion criteria	Any abnormalities in biochemical investigation (such as FBC, protein electrophoresis ionogram, calcium, phosphorous, renal function, liver function, glycaemia, muscle enzymes, antinuclear factor, cortisol, thyroid function, hepatitis B and C serology, urine microscopy, chest X-ray and abdominal US); primary psychiatric disorders; addition problems; <18 years; history of stomach/duodenal ulcers, arterial hypertension, glaucoma, DM, cardiac arrhythmia; Tourette's syndrome; use of beta blockers, antidepressant or antipsychotic medication; ongoing pregnancy
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 40 (8). Gender (M:F): 15:45. Ethnicity: unclear
Further population details	-
Extra comments	Median fatigue duration 36 months (IQR 22-74); weight 71.3kg; sbp 128(15) mmHg; dbp 81 mmHg; HR 72(8); sleepless ness 67%; dry mouth 38%; dizziness 70%; akathisia 70%; abdominal pain 53%; chest pain 43%.
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=60) Intervention 1: sympathomimetic/central antihypertensive drugs - methylphenidate. 10 mg twice daily (8am and 2pm). Taken for 1 month. Duration 1 months. Concurrent medication/care: Washout period of 1 week (half-life of drug is 2 hours so over 1 week there would be only 1 / [2 to the power of 84] remaining - thus it is an appropriate duration). Patients who stopped the treatment during the first period but who returned after 4 weeks were allowed to start therapy with the second compound. Indirectness: No indirectness (n=60) Intervention 2: placebo. Taken twice daily, but unclear if identical in appearance to study drug. Duration 1 month. Concurrent medication/care: Washout period of 1 week (half-life of drug is 2 hours so over 1 week there would be only 1 / [2 to the power of 84] remaining - thus it is an appropriate duration). Patients who stopped the treatment during the first period but who returned after 4 weeks were allowed to start therapy with the second compound. Indirectness: No indirectness

	Comments: The same 60 patients took both drugs, but in a random order. Thus about half would have had the study drug in the first period, whilst the other half would have had the placebo first. A washout period of 1 week was used before each patient took the alternative compound in the second period of 4 weeks.
Funding	Funding not stated (No report of conflicts of interest or funding)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: METHYLPHENIDATE (RITALIN) (KPAX002) versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF-36 Physical composite at 4 weeks; Group 1: mean 52.8 (SD 19); n=60, Group 2: mean 51.2 (SD 18.7); n=60; SF36 physical composite 0-100, High=Top is good outcome. Unclear whether a more appropriate paired analysis was performed. However the current analysis will be conservative in that it will tend to underestimate the precision.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF-36 mental composite at 4 weeks; Group 1: mean 51.8 (SD 16.4); n=60, Group 2: mean 47.3 (SD 16.7); n=60; SF36 physical composite 0-100, High=Top is good outcome. Unclear whether a more appropriate paired analysis was performed. However the current analysis will be conservative in that it will tend to underestimate the precision.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: CIS fatigue total score at 4 weeks; Group 1: mean 102.8 (SD 22.4); n=60, Group 2: mean 112.5 (SD 11.3); n=60; Checklist Individual Strength – fatigue 20-140, High=Top is poor outcome. Unclear whether a more appropriate paired analysis was performed. However the current analysis will be conservative in that it will tend to underestimate the precision.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: HADS Depression at 4 weeks; Group 1: mean 7.3 (SD 3.8); n=60, Group 2: mean 7.7 (SD 3.7); n=60; Hospital anxiety and depression scale depression subscale 0-21, High=Top is poor outcome. Unclear whether a more appropriate paired analysis was performed. However the current analysis will be conservative in that it will tend to underestimate the precision.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: HADS anxiety at 4 weeks; Group 1: mean 8.3 (SD 3.8); n=60, Group 2: mean 8.7 (SD 4.7); n=60; Hospital anxiety

and depression scale anxiety subscale 0-21, High=Top is poor outcome. Unclear whether a more appropriate paired analysis was performed. However the current analysis will be conservative in that it will tend to underestimate the precision. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Adverse events at longest follow up available - Actual outcome for adults; severity mixed or unclear: sleeplessness at 4 weeks; Group 1: 21/60, Group 2: 23/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: dry mouth at 4 weeks; Group 1: 34/60, Group 2: 18/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: dizziness at 4 weeks; Group 1: 30/60, Group 2: 38/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: Akathisia at 4 weeks; Group 1: 29/60, Group 2: 34/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: Abdominal pain at 4 weeks; Group 1: 28/60, Group 2: 23/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: chest pain at 4 weeks; Group 1: 17/60, Group 2: 25/60 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Fluge 2011[22]**

Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in Norway; Setting: Department of oncology, single (tertiary referral) centre
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: diagnosis of CFS by a neurologist, according to the Fukuda 1994 criteria; pre-treatment evaluation included thorough clinical examination, standard laboratory tests and further diagnostic tests if pre-treatment evaluation revealed any relevant abnormality that could explain the severe fatigue
Stratum	adults; severity mixed or unclear: age 18–65 years, meeting Fukuda 1994 criteria
Subgroup analysis within study	Not applicable: NA
Inclusion criteria	a diagnosis of CFS by a neurologist, according to the Fukuda 1994 criteria; age 18–65 years; written informed consent
Exclusion criteria	fatigue not fulfilling CFS criteria; previous malignant disease (except basal cell carcinoma and cervical dysplasia); previous long-term immunosuppressive treatment; previous Rituximab treatment; endogenous depression; lack of ability to adhere to protocol; evidence of on-going infection
Recruitment/selection of patients	Most of the participants were recruited from patients referred to Department of Neurology
Age, gender and ethnicity	Age - Mean (SD): Rituximab 37.3 (11.5) years, placebo 31.5 (11.6) years. Gender (M:F): 9/21. Ethnicity: not reported
Further population details	-
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM is not a compulsory feature
Interventions	(n=15) Intervention 1: immunomodulatory drugs - rituximab. Rituximab 500 mg/m² (maximum 1000 mg), diluted in saline to a concentration of 2 mg/ml, or an equal volume of saline, were given twice two weeks apart, with nurse

	surveillance and according to local guidelines used for treating B-cell lymphomas. Infusion bags had double plastic covers to avoid content identification by nurse or patient. Duration 2 weeks. Concurrent medication/care: No additional Rituximab infusions, or other intervention, were given during follow-up. All patients were given oral cetirizine 10 mg, paracetamol 1 g, and dexamethasone 8 mg prior to infusion. Indirectness: No indirectness; Indirectness comment: NA (n=15) Intervention 2: placebo. An equal volume if saline, given twice two weeks apart, with nurse surveillance and according to local guidelines used for treating B-cell lymphomas. Infusion bags had double plastic covers to avoid content identification by nurse or patient. Duration 2 weeks. Concurrent medication/care: No additional infusions, or other interventions, were given during follow-up. All patients were given oral cetirizine 10 mg, paracetamol 1 g, and dexamethasone 8 mg prior to infusion. Indirectness: No indirectness; Indirectness comment: NA
Funding	Academic or government funding (Helse Vest (Western Norway Regional Health Authority); legacy of Torstein Hereid)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RITUXIMAB versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF36 physical composite (max % change from baseline) at 10 months ; Group 1: mean 54 (SD 46); n=13, Group 2: mean 26 (SD 17); n=15
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Baselines, mean (SD): Rituximab 24 (5); Placebo 26 (6); Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 mental composite (max % change from baseline) at 10 months ; Group 1: mean 9 (SD 54); n=13, Group 2: mean 5 (SD 32); n=15
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Baselines, mean (SD): Rituximab 46 (11); Placebo 46 (8); Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available;

	Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Hickie 2000[34]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=90)
Countries and setting	Conducted in Australia; Setting: Recruited from infectious diseases and immunology outpatient clinics in Sydney, Australia.
Line of therapy	Not applicable
Duration of study	Intervention time: 6 weeks
Method of assessment of guideline condition	Inadequate method of assessment/diagnosis: Lloyd criteria - not a set of criteria based on expert group consensus
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 18-65; fulfilled diagnostic criteria for CFS by LLoyd et al. (1988) [similar to 1994 CDC comprising chronic, persisting or relapsing fatigue for >6 months with neurological dysfunction including impairment of concentration and/or new onset of short-term memory impairment].
Exclusion criteria	Diagnosis of alternative illness that explains symptoms; steroid medication or other immunomodulatory agents; hepatic dysfunction; recent alcohol or substance abuse; pregnant/breastfeeding/not using contraception.
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (range): 42.3 to 44.9. Gender (M:F): 41:49. Ethnicity: unclear

Further population details	-
Extra comments	moclobemide/placebo: duration of illness 84.2 weeks/90.9 weeks; initial KPI score 74.3/75.9; POMS fatigue 18/18; POMS vigour 8.2/8.8; POMS depression 12.9/14.1; current major depression 30%/40%; current psychological distress 68%/67%; CD4 T cell count 0.87/0.95
Indirectness of population	Serious indirectness: Lloyd 1988 criteria were excluded from the diagnostic criteria review on the basis there was unclear methodology for the development of the criteria and have therefore been downgraded here for indirectness.
Interventions	(n=47) Intervention 1: antidepressants - MAOIs. Moclobemide - a reversible inhibitor of monoamine oxidase (RIMA) - initially given as 150mg tablet twice daily after meals. After 1 week the dose was increased to 2 tablets in morning and 1 tablet at night for a total dose of 450mg/day. This was increased to 600mg/day if tolerated. Duration 6 weeks. Concurrent medication/care: Intermittent night dosages of benzodiazepines allowed for insomnia. Indirectness: No indirectness (n=43) Intervention 2: placebo. Identical 150mg tablet given in same dosages and time points as moclobemide - i.e. initially 300mg/day (in 2 doses) rising to 450mg or 600mg per day. Duration 6 weeks. Concurrent medication/care: Intermittent benzodiazepines allowed for sleep problems. Indirectness: No indirectness
Funding	Funding not stated (No mention of funding or conflicts of interest)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: MAOIS versus PLACEBO

Protocol outcome 1: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Karnofsky performance index (KPI) - measures level of disability at 6 weeks; Group 1: mean 0.86 standard deviation at baseline (SD 1.2); n=40, Group 2: mean 0.58 standard deviation at baseline (SD 1.3); n=37; Karnofsky performance index - measures level of disability scale not reported, High=Top is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 2: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Profile of mood states (POMS) - fatigue at 6 weeks; Group 1: mean -0.05 Units of baseline standard deviation (SD

0.4); n=40, Group 2: mean -0.01 Units of baseline standard deviation (SD 0.3); n=37; Profile of mood states – fatigue 0-28, High=Top is poor outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: - Actual outcome for adults; severity mixed or unclear: Profile of mood states (POMS) - vigour at 6 weeks; Group 1: mean 0.51 units of baseline standard deviation (SD 1.2); n=40, Group 2: mean 0 units of baseline standard deviation (SD 1.1); n=37; Profile of mood states – vigour 0-32, High=Top is good outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: - Actual outcome for adults; severity mixed or unclear: Profile of mood states (POMS) - depression at 6 weeks; Group 1: mean -0.06 standard deviations at baseline (SD 1); n=40, Group 2: mean -0.08 standard deviations at baseline (SD 0.7); n=37; Profile of mood states – depression 0-60, High=Top is poor outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 3: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Globally improved cases at 6 weeks; Group 1: 24/47, Group 2: 14/43 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Kakumanu 2003[37]**
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	1 (n=28)
Countries and setting	Conducted in USA; Setting: University hospital
Line of therapy	Not applicable

Duration of study	Intervention time: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: CDC criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 18-65; CDC criteria for CFS; symptoms of rhinitis
Exclusion criteria	sleep apnoea; obesity; nasal polyps; recent URTI; deviated septum; seasonal allergic rhinitis; asthma; other respiratory diseases
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (range): 46.2(31-62). Gender (M:F): 8:20. Ethnicity: unclear
Further population details	-
Extra comments	perennial rhinitis 54%; nonallergic rhinitis 46%;
Indirectness of population	Very serious indirectness: All participants had rhinitis and 1994 CDC criteria used; PEM is not a compulsory feature
Interventions	(n=21) Intervention 1: oral corticosteroids - fludrocortisone/hydrocortisone/other. NASAL (not oral) corticosteroid (Flunisolide) self-administered with two sprays twice daily. Duration 4 weeks - 8 weeks. Concurrent medication/care: This was a hybrid parallel/cross-over trial design. There were 4 groups of 7 who were treated as follows: active throughout; placebo throughout; active then placebo; placebo then active. Thus 21 had the active treatment at one point. In the analysis the results from these 21 people were aggregated without any apparent adjustments for some having had the other treatment (with the possibility of carryover effects). Indirectness: No indirectness (n=21) Intervention 2: placebo. Saline spray - 2 sprays twice daily. Duration 4 weeks - 8 weeks. Concurrent medication/care: This was a hybrid parallel/cross-over trial design. There were 4 groups of 7 who were treated as follows: active throughout; placebo throughout; active then placebo; placebo then active. Thus 21 had the placebo at some point. In the analysis the results from these 21 people were aggregated without any apparent adjustments for

	some having had the other treatment (with the possibility of carryover effects). Indirectness: No indirectness
Funding	Academic or government funding (GCRC grant)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: FLUDROCORTISONE/HYDROCORTISONE/OTHER versus PLACEBO Protocol outcome 1: Sleep quality at longest follow up available - Actual outcome for adults; severity mixed or unclear: Epworth Sleepiness Scale at 4-8 weeks; MD; -3.18 (95%CI -6.57 to 0.21); Epworth sleepiness scale 0-24, High=poor outcome; Comments: baseline scores not reported Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: - Actual outcome for adults; severity mixed or unclear: Functional Outcomes of Sleep Questionnaire at 4-8 weeks; MD; 0.89 (95%CI -0.9884 to 2.7716); University of Pennsylvania Functional Outcomes of Sleep Quality of Life Survey scale not reported, High=Top is good outcome; Comments: baseline scores not reported Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 2: Fatigue at longest available follow up - Actual outcome for adults; severity mixed or unclear: Chronic Fatigue Syndrome Severity Rating at 4-8 weeks; MD; -3.17 (95%CI -7.48 to 1.14), Units: unclear, High=poor outcome; Comments: Unclear if this is a validated scale; CIs calculated from SE Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Symptom scales at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Lloyd 1990[44]**

Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=49)
Countries and setting	Conducted in Australia; Setting: Unclear.
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 3 months + 3 months
Method of assessment of guideline condition	Partially adequate method of assessment/diagnosis: Diagnosed according to Lloyd 1988 criteria. These criteria were not included in the diagnostic criteria review of this guideline and have therefore been downgraded for indirectness. However, the study states that their own criteria emphasize the same features as the criteria published subsequently by the Centers for Disease Control.
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	A diagnosis of CFS required: (1) a history of at least 6 months in duration of marked exercise-aggravated muscle fatigue, with abnormally prolonged recovery time, associated with typical constitutional and neuropsychiatric symptoms; (2) that CFS was producing frequent medical consultation and a substantial reduction in the ability to participate in usual daily activities when com-pared with the subject's pre-morbid status [for ex-ample, considerable time lost from school or work, and inability to participate in sports]. All patients had chronic and persisting symptomatology, rather than a relapsing and remitting course as sometimes reported for this syndrome.
Exclusion criteria	A physical examination and standardized investigation protocol excluded other chronic infectious or immunodeficiency-related disorders.
Recruitment/selection of patients	Unclear.
Age, gender and ethnicity	Age - Mean (SD): Treatment group 39 (10) years; placebo group 33 (12) years. Gender (M:F): 25 males, 24 females. Ethnicity: Not stated.

Further population details	-
Extra comments	All subjects had normal blood cell counts, renal and liver function tests, muscle enzyme assays, thyroid function tests, antinuclear antibodies, rheumatoid factor, total immunoglobulin levels, and serologic tests for syphilis, hepatitis B, and human immunodeficiency virus.
Indirectness of population	Serious indirectness: CFS diagnostic criteria used (Lloyd 1988) were excluded from the diagnostic criteria review on the basis there was unclear methodology for the development of the criteria. The study states that the criteria emphasize the same features as the criteria published subsequently by the Centers for Disease Control.
Interventions	(n=23) Intervention 1: immunomodulatory drugs - IV immunoglobulin G. High-dose intravenous (IV) immunoglobulin G. Immunoglobulin G (Intragam, Commonwealth Serum Laboratories, Melbourne Australia [based on the formulation of Gamimune N, Cutter Laboratories, Berkeley, California]) was administered intravenously by continuous infusion in a dosage of 2 g (IgG)/kg. Three infusions lasting 24 hours were administered at monthly intervals. Duration 3 24-hour infusions over 3 months. Concurrent medication/care: Not stated. Indirectness: No indirectness (n=26) Intervention 2: placebo. Placebo (10% w/v maltose) was administered intravenously for 24 hours at an equivalent volume to the IgG infusion. Duration 3 24-hour infusions over 3 months. Concurrent medication/care: Not stated. Indirectness: No indirectness
Funding	Academic or government funding (The research was funded by scholarships from the National Health and Medical Research Council of Australia (AL), Canberra, Australia, and the New South Wales Institute of Psychiatry (IH) and Myalgic Encephalomyelitis Society of New South Wales, Sydney, Australia.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IV IMMUNOGLOBULIN G versus PLACEBO

Protocol outcome 1: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Hamilton Depression Scale at 6 months; Group 1: mean 9 (SD 5); n=23, Group 2: mean 10 (SD 3); n=26; Hamilton Depression Scale 0-62 Top=High is poor outcome; Comments: Baselines, mean (SD): Immunoglobulin 10.7 (2.8) Placebo 10.5 (3.4)
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: Zung Self-Rating Depression Scale at 6 months; Group 1: mean 41 (SD 11); n=23, Group 2: mean 40 (SD 12); n=26; Zung Self-Rating Depression Scale 0-80 Top=High is poor outcome; Comments: Baselines, mean (SD): Immunoglobulin 42 (8) Placebo 38 (11) Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Baselines, mean (SD): Immunoglobulin 42 (8) Placebo 38 (11); Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Return to school or work at longest follow up available - Actual outcome for adults; severity mixed or unclear: Resumption of pre-morbid employment status in full-time occupations or housework. at 6 months; Group 1: 6/23, Group 2: 0/26; Comments: Reported that six of the 13 patients (all from immunoglobulin group) who 'responded' (ie had a marked reduction in symptoms and improvement in functional capacity) resumed their pre-morbid employment status in full-time occupations or housework. Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Marked reduction in symptoms and improvement in functional capacity at 6 months; Group 1: 10/23, Group 2: 3/26; Comments: Determined through an evaluation of symptoms and disability by the physician, meeting the criteria for "response". Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Mckenzie 1998[47]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=70)
Countries and setting	Conducted in USA; Setting: Unclear

Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: CDC 1988 criteria for CFS (all met 1994 criteria as well)
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged 18 to 55 years who met the CDC1988 criteria for CFS; illness began over a period of 6 weeks or less; use of birth control; negative pregnancy test
Exclusion criteria	Contraindications to systemic steroids; any other acute or chronic condition that required ongoing or intermittent medication; use of any prescribed and OTC drugs (except paracetamol) in 2-6 weeks before enrolment or during study
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: hydrocortisone 36.7yrs, placebo 38.3 yrs. Gender (M:F): 14:56. Ethnicity: Hydrocortisone/placebo; white 97%/94%, black 3%/0, other 0/6%
Further population details	-
Extra comments	Hydrocortisone/placebo: length of illness 46.9/59.9 months; impaired employment 77%/69%; urine cortisol 192/187;resting serum cortisol 425/397; self-rating wellness score 38.8/37.6; BDI 12/9.9; activity scale 4.7/5; SIP 18.7/17.9; profile of mood states (PMS) anger 5.7/4.7; PMS anxiety 8.1/8; PMS confusion 10.7/10; PMS depression 7.7/4.8; PMS fatigue 19.6/17.8; PMS vigor 7.9/7.3; SQ 90-R general severity index 0.61/0.53; positive symptom distress index 1.7/1.8; positive symptom total 29.4/26.2; Hamilton Depression rating Scale 9.8/9.4; concurrent mental disorders 74%/94%
Indirectness of population	Serious indirectness – Holmes 1988 and 1994 CDC criteria used; PEM is not a compulsory feature

Interventions	(n=35) Intervention 1: oral corticosteroids -fludrocortisone/hydrocortisone/other. Hydrocortisone pills - dose of 16mg per square metre of body surface per day (20-30mg every morning at 8am and 5mg every day at 2pm). Duration 12 weeks. Concurrent medication/care: None. Indirectness: No indirectness (n=35) Intervention 2: placebo. Identical placebo at same doses as hydrocortisone group. Duration 12 week. Concurrent medication/care: None. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: FLUDROCORTISONE/HYDROCORTISONE/OTHER versus PLACEBO

Protocol outcome 1: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Beck Depression Inventory (BDI) at 12 weeks; Group 1: mean -2.1 (SD 5.1); n=34, Group 2: mean -0.4 (SD 4.1); n=34; Beck Depression Inventory (BDI) 0-63 Top=High is poor outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: BDI 12/9.9; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear
- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - anger at 12 weeks; Group 1: mean -1.6 (SD 3.9); n=34, Group 2: mean -0.8 (SD 3.8); n=34; Profile of Mood States (POMS) 0-48 Top=High is poor outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear
- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - anxiety at 12 weeks; Group 1: mean -0.8 (SD 2.5); n=34, Group 2: mean -2.1 (SD 3.6); n=34; Profile of Mood States (POMS) 0-46 Top=High is poor outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear
- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - confusion at 12 weeks; Group 1: mean -1.1 (SD 3.3); n=34, Group 2: mean -1.4 (SD 2.9); n=34; Profile of Mood States (POMS) 0-28 Top=High is poor outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear
- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - depression at 12 weeks; Group 1: mean -1.6 (SD 4.6); n=34, Group 2: mean 0 (SD 3.8); n=34; Profile of Mood States (POMS) 0-60 Top=High is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - fatigue at 12 weeks; Group 1: mean -3.6 (SD 5.3); n=34, Group 2: mean -1.8 (SD 4.5); n=34; Profile of Mood States (POMS) 0-28 Top=High is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Profile of Mood States - vigour at 12 weeks; Group 1: mean 1.2 (SD 3.3); n=34, Group 2: mean 0.7 (SD 3.3); n=34; Profile of Mood States (POMS) 0-32 Top=High is good outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Symptom checklist-90-R general severity index at 12 weeks; Group 1: mean -0.1 (SD 0.2); n=34, Group 2: mean -0.1 (SD 0.2); n=34; Symptom checklist-90-R general severity index scale not reported, High=Top is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Symptom checklist-90-R positive symptom distress index at 12 weeks; Group 1: mean 0 (SD 0.3); n=34, Group 2: mean -0.1 (SD 0.3); n=34; Symptom checklist-90-R positive symptom distress index scale not reported, High=Top is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Symptom checklist-90-R positive symptom total at 12 weeks; Group 1: mean -2.6 (SD 10.8); n=34, Group 2: mean -2.4 (SD 11.5); n=34; Symptom checklist-90-R positive symptom total scale not reported, High=Top is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear

- Actual outcome for adults; severity mixed or unclear: Hamilton Depression rating Scale at 12 weeks; Group 1: mean -0.8 (SD 3.8); n=32, Group 2: mean 0.1 (SD 2.9); n=33; Hamilton Depression Rating Scale 0-52 Top=High is poor outcome

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 3, Reason: unclear; Group 2 Number missing: 2, Reason: unclear

Protocol outcome 2: Adverse events at longest follow up available

- Actual outcome for adults; severity mixed or unclear: Any adverse reaction at 12 weeks; Group 1: 31/35, Group 2: 27/35;

Adverse reactions included fatigue, depressed mood, difficulty with concentration, increased appetite, weight gain and more. Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 0, Reason: 0; Group 2 Number missing: 0, Reason: 0 Protocol outcome 3: Activity levels at longest follow up available - Actual outcome for adults; severity mixed or unclear: Activity scale at 12 weeks; Group 1: mean 0.3 (SD 1.1); n=34, Group 2: mean 0.7 (SD 1.4); n=34; Activity scale unclear Top=High is good outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: hydrocortisone group had better status generally; Group 1 Number missing: 1, Reason: unclear; Group 2 Number missing: 1, Reason: unclear Protocol outcome 4: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Wellness scale at 12 weeks; Group 1: mean 6.3 (SD 11.7); n=30, Group 2: mean 1.7 (SD 8.8); n=35; Wellness scale 0-100 Top=High is good outcome Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Wellness score 38.8/37.6; Group 1 Number missing: 5, Reason: no pre-treatment scores; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: Sickness Impact Profile at 12 weeks; Group 1: mean -2.5 (SD 6.4); n=33, Group 2: mean -2.2 (SD 6.8); n=34; Sickness Impact Profile not reported, Top=High is poor outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: SIP 18.7/17.9; Group 1 Number missing: 2, Reason: unclear; Group 2 Number missing: 1, Reason: unclear	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Montoya 2018[52]**
Study type	RCT (Patient randomised; Parallel)

Number of studies (number of participants)	1 (n=135)
Countries and setting	Conducted in USA; Setting: Conducted at 4 sites in USA
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1994 CDC criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 18-59; met CDC 1994 criteria; complained of alertness and concentration difficulties; otherwise in good health based on medical history and screening; willing not to use any nutritional, herbal, or caffeine/pseudoephedrine containing compounds
Exclusion criteria	Pregnancy; active substance abuse; major depression; active medical conditions for which methylphenidate hydrochloride is contraindicated; daily anxiolytics; daily use of >1 antidepressant; use of MAOs, CNS stimulants and narcotic opioids; abnormal laboratory test values; ECG abnormalities
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age: Drug/placebo: 42.8/42.3 years. Gender (M:F): 36:92. Ethnicity: Drug/placebo: white 90%/91%; Asian 3%/0; African American 2%/8%; Other 5%/2%
Further population details	-
Extra comments	Drug/placebo: duration of CFS symptoms >=10 years 48%/46%; mean CIS total score 112.2/112.4
Indirectness of population	Serious population indirectness: 1994 CDC criteria used; PEM not a compulsory feature

Interventions	(n=67) Intervention 1: sympathomimetic/central antihypertensive drugs - methylphenidate. 5mg methylphenidate daily for week 1 and 10mg twice daily for weeks 2 to 12. Mitochondrial modulator (nutritional supplement) given as 4 tablets twice daily. The combination of these two agents is called KPAX002. KPAX002 is comprised of a low dosage of methylphenidate hydrochloride, combined with nutrients believed to modulate mitochondrial function. Duration 12 weeks. Concurrent medication/care: Taken with breakfast and lunch. Indirectness: No indirectness (n=65) Intervention 2: placebo. Placebo version of KPAX002 treatment. Unclear if this meant both placebo versions of methylphenidate and mitochondrial modulator, or just the former. Likely to be both. . Duration 12 weeks. Concurrent medication/care: Taken with breakfast and lunch. Indirectness: No indirectness
Funding	Funding not stated (No conflicts of interest statement made)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: METHYLPHENIDATE (RITALIN) (KPAX002) versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Checklist Individual Strength (CIS) total score at 12 weeks; Group 1: mean -16.9 (SD 23.52); n=63, Group 2: mean -13.8 (SD 22.15); n=65; Checklist Individual Strength scale 20-140, High=Top is poor outcome
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Change values analysed so any unreported baseline discrepancies may not create sig bias; Group 1 Number missing: 4, Reason: did not meet ITT criteria of at least 1 complete treatment; Group 2 Number missing: 3, Reason: did not meet ITT criteria of at least 1 complete treatment

Protocol outcome 2: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: AEs leading to discontinuation at 12 weeks; Group 1: 8/63, Group 2: 3/65
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Change values analysed so any unreported baseline discrepancies may not create sig bias; Group 1 Number missing: 4, Reason: did not meet ITT criteria of at least 1 complete treatment; Group 2 Number missing: 3, Reason: did not meet ITT criteria of at least 1 complete treatment
- Actual outcome for adults; severity mixed or unclear: Serious AEs (pyelonephritis) at 12 weeks; Group 1: 1/63, Group 2: 0/65
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Change values analysed so any unreported baseline discrepancies may not create sig bias; Group 1 Number

missing: 4, Reason: did not meet ITT criteria of at least 1 complete treatment; Group 2 Number missing: 3, Reason: did not meet ITT criteria of at least 1 complete treatment	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Morriss 2002[56]**
Study type	RCT (Patient randomised; Crossover: 2 weeks)
Number of studies (number of participants)	1 (n=10)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient clinic for CFS at a general hospital in UK
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1994 Fukuda
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	CFS diagnosed by Fukuda criteria; no non-CFS diagnoses accounting for symptoms
Exclusion criteria	ICD-10 Depressive episode; psychotropic medication, oral contraceptives, steroids, thyroxine, bromocriptine or anti-hypertensive medication in previous 15 days; age <18 years or above 60 years; BMI <15, >30; migraine; pregnancy or breast feeding.

Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 46(7.6). Gender (M:F): 1:1. Ethnicity: unclear
Further population details	-
Extra comments	BMI 24.8; NART IQ 118.4; Total fatigue 31.7; Cognitive failures questionnaire total score 57.8; HADS depression 5.5; HADS anxiety 5.5; somatosensory amplification 27.5; duration of CFS 75mo
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=10) Intervention 1: sympathomimetic/central antihypertensive drugs - clonidine. Clonidine IV infusion 2.5 micrpog/kg in 10ml normal saline over 5 minutes. Duration One-off treatment. Concurrent medication/care: Heparinised cannula used for infusion. Cross-over and randomised order for clonidine/placebo with washout of 2 weeks. Indirectness: No indirectness (n=10) Intervention 2: placebo. IV infusion of 10ml normal saline over 5 mins. Duration One-off treatment. Concurrent medication/care: Heparinised cannula used for IV. Cross-over and randomised order for clonidine/placebo with washout of 2 weeks. Indirectness: No indirectness
Funding	Academic or government funding (MRC)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CLONIDINE versus PLACEBO

Protocol outcome 1: Cognitive function at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Stockings of Cambridge - minimum moves at 30 minutes; Group 1: mean 9 (SD 2.18); n=9, Group 2: mean 10.22 (SD 2.39); n=9.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Stockings of Cambridge - initial thinking time (s) at 30 minutes; Group 1: mean 7.99 (SD 4.34); n=9, Group 2: mean 9.27 (SD 4.13); n=9

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Stockings of Cambridge - subsequent thinking time (s) at 30 minutes; Group 1: mean 1.38 (SD 2.46); n=9, Group 2: mean 1.89 (SD 3.07); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Rapid Visual Information Processing - reaction time (s) at 30 minutes; Group 1: mean 5 (SD 1.52); n=9, Group 2: mean 5.15 (SD 1.22); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Intradimensional (IDS) set sift/extradimensional (EDS) set shift: IDS errors at 30 minutes; Group 1: mean 0.44 (SD 0.73); n=9, Group 2: mean 0.22 (SD 0.44); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Intradimensional (IDS) set sift/extradimensional (EDS) set shift: EDS errors at 30 minutes; Group 1: mean 1.78 (SD 1.56); n=9, Group 2: mean 4.44 (SD 6.64); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Spatial working memory: between-search errors at 30 minutes; Group 1: mean 7.09 (SD 4.21); n=9, Group 2: mean 9.26 (SD 6.82); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: Spatial working memory: strategy score at 30 minutes; Group 1: mean 31.56 (SD 5.96); n=9, Group 2: mean 31.78 (SD 6.38); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.
- Actual outcome for adults; severity mixed or unclear: pattern recognition - number correct at 30 minutes; Group 1: mean 22.3 (SD 1.3); n=9, Group 2: mean 21.4 (SD 2.2); n=9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome:

No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used. - Actual outcome for adults; severity mixed or unclear: spatial recognition - number correct at 30 minutes; Group 1: mean 15.2 (SD 2.9); n=9, Group 2: mean 15.3 (SD 2.1); n=9 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used. - Actual outcome for adults; severity mixed or unclear: spatial span - length at 30 minutes; Group 1: mean 6.4 (SD 1.26); n=9, Group 2: mean 6.1 (SD 1.2); n=9 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used. - Actual outcome for adults; severity mixed or unclear: Delayed matching to sample 2-s delay at 30 minutes; Group 1: mean 6.56 (SD 1.69); n=9, Group 2: mean 7.78 (SD 1.39); n=9 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used. - Actual outcome for adults; severity mixed or unclear: Paired associate learning - sets completed at 30 minutes; Group 1: mean 8.89 (SD 0.33); n=9, Group 2: mean 8.89 (SD 0.33); n=9 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.; Group 2 Number missing: 1, Reason: Patient placed on exclusion medication by GP between the tests. Only received placebo but data for both arms not used.	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Olson 2003[61]**
Study type	RCT (Patient randomised; Parallel)

Number of studies (number of participants)	1 (n=20)
Countries and setting	Conducted in Australia; Setting: Newcastle Sleep Disorders Centre, Australia
Line of therapy	Not applicable
Duration of study	Intervention time: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fukuda
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Fukuda criteria; normal results for overnight sleep study; mean daytime latency of >7 minutes;
Exclusion criteria	history of alcohol or other substance abuse; history of epilepsy; history of MI; current hypertension; cardiac arrhythmias; angina pectoris; coeliac disease; psychiatric disorders other than depression; use of anti-depressant drugs
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range: 17-72. Gender (M:F): 7:13. Ethnicity: unclear
Further population details	-
Extra comments	Dexa/placebo; length of illness 7.1yrs/5.6yrs; mean sleep latency 12.9mins/13mins; member of patient support group 10%/0%; employed 80%/80%; age 32.1/39.7
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=10) Intervention 1: Amphetamines - dexamphetamine. Dexamphetamine 5mg twice daily for first week. Dose increased to 10mg twice daily if indicated at start of 2nd week. Increment repeated if appropriate at start of 3rd week. This dose continued for a further 4 weeks. Duration 6 weeks. Concurrent medication/care: None. Indirectness: No indirectness

	(n=10) Intervention 2: placebo. Identical doses and strategies for increase as study drug. Duration 6 weeks. Concurrent medication/care: None. Indirectness: No indirectness
Funding	Equipment / drugs provided by industry (Sigma Pharmaceuticals)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DEXAMPHETAMINE versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF36 Physical composite at 6 weeks; Group 1: mean 6.9 (SD 13.97); n=10, Group 2: mean 5.2 (SD 10.76); n=10; SF36 physical composite 0-100, High=Top is good outcome; Comments: sds estimated from 95% CIs given for each group in the paper
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 Mental composite at 6 weeks; Group 1: mean 4.2 (SD 11.46); n=10, Group 2: mean 3.9 (SD 12.86); n=10; SF36 mental composite 0-100, High=Top is good outcome; Comments: sds estimated 95% CIs given for each group in the paper
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue Severity Scale at 6 weeks; Group 1: mean -1.45 (SD 1.09); n=10, Group 2: mean -0.03 (SD 1.11); n=10; Fatigue Severity Scale scale not reported, High=Top is poor outcome
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Sleep quality at longest follow up available
- Actual outcome for adults; severity mixed or unclear: sleep latency at 6 weeks; Group 1: mean 13 Time taken to fall asleep when light turned off (mins) (SD 5.45); n=10, Group 2: mean 11.8 Time taken to fall asleep when light turned off (mins) (SD 3.77); n=10; Comments: sds estimated from 95% CIs given for each group in the paper. Final values used as groups very similar at baseline.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 4: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: AEs - anorexia at 6 weeks; Group 1: 5/10, Group 2: 1/10
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Pardini 2011[63]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Italy; Setting: Single centre through referrals from clinicians and through self-referrals
Line of therapy	Not applicable
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fukuda
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Fukuda criteria; routine laboratory tests within normal ranges; no neurological or psychiatric conditions
Exclusion criteria	None reported
Recruitment/selection of patients	consecutive

Age, gender and ethnicity	Age - Mean (SD): 31.9 (1.8). Gender (M:F): 18:22. Ethnicity: unclear
Further population details	-
Extra comments	Baseline details: Amisulpride/fluoxetine: FSS 50.5/52.4; VAS pain 59.9/55.9; HADS A 5.3/5.3; HADs D 4.9/5.1; SF-12 41.3/41.7; CGI-S 4.9/4.6; mean disease duration 2.5 years/2.9 years
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM is not a compulsory feature
Interventions	(n=20) Intervention 1: Antipsychotics. Amisulpride (a substituted benzamide) is an atypical antipsychotic. 25 mb bid. Duration 12 weeks. Concurrent medication/care: None. Indirectness: No indirectness (n=20) Intervention 2: antidepressants - selective serotonin reuptake inhibitors. Fluoxetine 20 mg uid. Duration 12 weeks. Concurrent medication/care: None. Indirectness: No indirectness
Funding	No funding (No conflicts of interest statement)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ANTIPSYCHOTICS versus SELECTIVE SEROTONIN REUPTAKE INHIBITORS

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF-12 at 12 weeks; Group 1: mean 53.2 (SD 4.8); n=20, Group 2: mean 37.6 (SD 4.9); n=20; SF12 0-100 Top=High is good outcome; Comments: Baseline values very similar (41.3/41.7) slightly favouring fluoxetine, so this does not create bias explaining final follow up result in favour of amisulpride
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue Severity Scale at 12 weeks; Group 1: mean 36.3 (SD 8.6); n=20, Group 2: mean 48.9 (SD 4.9); n=20; Fatigue severity scale 9-63 Top=High is poor outcome; Comments: Slight difference at baseline (50.5/52.4) but not enough to explain result at follow up
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Psychological status at longest follow up available - Actual outcome for adults; severity mixed or unclear: HADS - anxiety at 12 weeks; Group 1: mean 4.5 (SD 1); n=20, Group 2: mean 4.9 (SD 1); n=20; Hospital anxiety and depression scale 0-21 Top=High is poor outcome; Comments: groups same at baseline (5.3/5.3) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for adults; severity mixed or unclear: HADS - depression at 12 weeks; Group 1: mean 4.3 (SD 0.9); n=20, Group 2: mean 4.2 (SD 1); n=20; Hospital anxiety and depression scale 0-21 Top=High is poor outcome; Comments: Similar baseline values (4.9/5.1) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Pain at longest follow up available - Actual outcome for adults; severity mixed or unclear: VAS pain at 12 weeks; Group 1: mean 40.5 (SD 13.1); n=20, Group 2: mean 53.1 (SD 8.3); n=20; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Different at baseline, but favouring fluoxetine at baseline (59.9/55.9) so the baseline difference did not create bias towards the observed 12 week effect in favour of amisulpride. Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 5: Adverse events at longest follow up available - Actual outcome for adults; severity mixed or unclear: FIBSER - global burden at 12 weeks; Group 1: mean 0.8 (SD 0.7); n=20, Group 2: mean 0.6 (SD 0.8); n=20; Frequency, Intensity, and Burden of Side Effects Rating Scale not reported Top=High is poor outcome; Comments: Measures overall burden of AEs Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 6: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Clinical Global Impression Severity (CGI-S) at 12 weeks; Group 1: mean 2.9 (SD 0.6); n=20, Group 2: mean 4.2 (SD 1.1); n=20; Clinical global impression severity 1-7 Top=High is poor outcome; Comments: Amilsupride worse at baseline (4.9/4.6) so baseline discrepancy does not explain benefit for amisulpride at follow up. Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	Peterson 1990[65]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in USA; Setting: All treatments were administered in individual rooms in the Drug Evaluation Unit, Hennepin County Medical Center.
Line of therapy	Not applicable
Duration of study	Intervention time: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Patients were diagnosed according to the CFS diagnostic criteria of Holmes 1988
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable:
Inclusion criteria	A diagnosis of CFS was established after thorough medical, psychometric, and psychiatric evaluations did not establish another explanation for chronic fatigue, and after the other criteria for a case definition of CFS were met (Holmes 1988).
Exclusion criteria	The psychometric assessment was performed, consisted of three standardized questionnaires (Beck Depression Inventory, Zung Self-Rating Anxiety Scale, and Symptom Checklist-90). Seventeen patients scored within the normal range on all three tests; 13 patients had abnormal scores on one or more psychometric tests, all of whom were interviewed by a psychiatric co-investigator (CS) who found no evidence of underlying psychopathology as an explanation of chronic fatigue.
Recruitment/selection of patients	Patients were recruited from a CFS Research Program established at Hennepin County Medical Center, Minneapolis, Minnesota, in July 1988.
Age, gender and ethnicity	Age - Mean (SD): 40.8 (11.2). Gender (M:F): 22 females, 8 males. Ethnicity: Not stated.

Further population details	-
Extra comments	Duration of illness in years, mean (SD): 3.8 (2.2)
Indirectness of population	Serious indirectness: Holmes 1988 criteria used; PEM is not a compulsory feature
Interventions	(n=15) Intervention 1: immunomodulatory drugs - IV immunoglobulin G. Patients were scheduled to receive a total of six infusions of IV IgG (1 g/kg body weight, Gammagard ®, Hyland Division, Baxter Healthcare Corp., Glendale, California). The IV IgG solution was prepared according to the package insert; each millilitre of solution contained 50 mg of IgG. The fusions were initiated at a rate of 0.5 mL/kg/hour and increased as per the IV IgG package insert to a maximum of 4 mL/kg/hour. All treatments were administered in individual rooms in the Drug Evaluation Unit, Hennepin County Medical Center. Treatments were given at intervals of 30 +/- 3 days. The first infusion was administered on Study Day 0 and the sixth infusion on Study Day 150. Duration Once per month for six months (6 infusions). Concurrent medication/care: During the course of the study, patients were permitted to take vitamins, nonsteroidal anti-inflammatory agents, decongestants, antihistamines, oral contraceptives, antibiotics, and other medications as prescribed by their primary physicians. Indirectness: No indirectness (n=15) Intervention 2: placebo. Participants in the placebo group received the same course of IV but IgG was replaced with an exactly correlating volume of a 1% albumin solution as placebo. The albumin solution was made using albumin USP 25% (Buminate, Baxter Healthcare Corp.) dissolved in normal saline; the resulting solution contained 10 mg/mL of albumin. Duration Once per month for six months (6 infusions). Concurrent medication/care: During the course of the study, patients were permitted to take vitamins, nonsteroidal anti-inflammatory agents, decongestants, antihistamines, oral contraceptives, antibiotics, and other medications as prescribed by their primary physicians. Indirectness: No indirectness
Funding	Study funded by industry (This work was supported in part by a grant from Baxter Healthcare Corp., Glendale, California. Dr. Lurie is a Henry J. Kaiser Family Foundation Faculty Scholar in General Internal Medicine)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IV IMMUNOGLOBULIN G versus PLACEBO

Protocol outcome 1: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Physical functioning on the Medical Outcome Study Short Form at 150 days (final treatment day); Group 1: mean

56 (SD 23.2); n=14, Group 2: mean 51.8 (SD 22.2); n=14; Medical Outcome Study Short Form 0-100 Top=High is good outcome; Comments: Baselines, mean (SD): IV IgG 63.1 (25.9), Placebo 66.1 (21.0) Risk of bias: All domain - High, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Serious adverse event; Group 2 Number missing: 1, Reason: Serious adverse event Protocol outcome 2: Psychological status at longest follow up available - Actual outcome for adults; severity mixed or unclear: Mental health on the Medical Outcome Study Short Form at 150 days (final treatment day); Group 1: mean 58.3 (SD 17.4); n=14, Group 2: mean 62.9 (SD 13.3); n=14; Medical Outcome Study Short Form 0-100 Top=High is good outcome; Comments: Baselines, mean (SD): IV IgG 63.7 (17.1), Placebo 59.7 (13.4) Risk of bias: All domain - High, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 1, Reason: Serious adverse event; Group 2 Number missing: 1, Reason: Serious adverse event Protocol outcome 3: Adverse events at longest follow up available - Actual outcome for adults; severity mixed or unclear: Major adverse events at 150 days (final treatment day); Group 1: 3/15, Group 2: 3/15; Comments: Nature of adverse events unclear. One participant in each group dropped out of the study as a result of their major adverse experience. Risk of bias: All domain - High, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Peterson 1998[64]**
Study type	RCT (Patient randomised; Crossover: 6 weeks)
Number of studies (number of particants)	1 (n=25)
Countries and setting	Conducted in USA; Setting: Patients on registries of a research programme in Minneapolis, or a CFS clinic

Line of therapy	Not applicable
Duration of study	Intervention time: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Holmes and Fukuda criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of CFS
Exclusion criteria	Fatigue severity during the preceding month of <5 on a 0-10 VAS; taking fludrocortisone or another medication that could confound interpretation of the results
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 39.7 (10.9). Gender (M:F): Define. Ethnicity: White: 100%
Further population details	-
Extra comments	Mean treatment duration 7 years; acute infectious onset 88%; Fatigue VAS 7.4; SF36 Physical 43.2;
Indirectness of population	Serious indirectness: Holmes 1988 and 1994 CDC criteria used; PEM is not a compulsory feature
Interventions	(n=25) Intervention 1: oral corticosteroids - fludrocortisone/hydrocortisone/other. Initial dose of fludrocortisone acetate was 0.1mg via 1 tablet orally. Dose doubled if no AEs reported after 2 weeks of treatment. Duration 6 weeks. Concurrent medication/care: Patients told not to make any dietary changes (including salt intake) during study. Cross-over study: patients randomised to order of drug/placebo with 6 week washout period. Indirectness: No indirectness (n=25) Intervention 2: placebo. identical tablets taken at same dosing regimen as study drug. Duration 6 weeks. Concurrent medication/care: Patients told not to make any dietary changes (including salt intake) during study. Cross-over study: patients randomised to order of drug/placebo with 6 week washout period. . Indirectness: No indirectness

Funding	Academic or government funding (Minnesota CFS association, Institute for Research and Education of Health System Minnesota, Minneapolis Medical Research Foundation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: FLUDROCORTISONE/HYDROCORTISONE/OTHER versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
-Actual outcome for adults; severity mixed or unclear: SF-36 general well-being at 6 weeks; Group 1: mean 32.1 (SD 12.4); n=20, Group 2: mean 35.8 (SD 15.9); n=20; SF36 general well-being 0-100, High=Top is good outcome; Comments: 32.9/35.6 at baseline
Risk of bis: All domain – High, Selection – Low, Blinding – Low, Incomplete outcome data – High, Outcome reporting – Low, Measurement – Low, Crossover – Low; Indirectness of outcome: No indirectness; Baseline details: No clinically important difference in outcome at baseline; Group Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
-Actual outcome for adults; severity mixed or unclear: SF-36 social at 6 weeks; Group 1: mean 40.1 (SD 20.4); n=20, Group 2: mean 38.2 (SD 21.4); n=20; SF36 social 0-100, High=Top is good outcome; Comments: 33.6/38.2 at baseline
Risk of bis: All domain – High, Selection – Low, Blinding – Low, Incomplete outcome data – High, Outcome reporting – Low, Measurement – Low, Crossover – Low; Indirectness of outcome: No indirectness; Baseline details: No clinically important difference in outcome at baseline; Group Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 emotional well-being at 6 weeks; Group 1: mean 72.6 (SD 13.9); n=20, Group 2: mean 68.8 (SD 15.4); n=20; SF36 emotional well-being 0-100, High=Top is good outcome; Comments: 73.9/73.9 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: same for outcome at baseline; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 emotional role limitation at 6 weeks; Group 1: mean 87.8 (SD 22.8); n=20, Group 2: mean 87.8 (SD 25.4); n=20; SF36 emotional role 0-100, High=Top is good outcome; Comments: 84.2/91.2 baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 physical at 6 weeks; Group 1: mean 49.7 (SD 20.4); n=20, Group 2: mean 42.1 (SD 21.4); n=20; SF36 physical 0-100, High=Top is good outcome; Comments: 43.2/43.7 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 role physical at 6 weeks; Group 1: mean 13.2 (SD 19.3); n=20, Group 2: mean 25 (SD 34.4); n=20; SF36 role physical 0-100, High= Top is good outcome; Comments: 23.7/15.8 at baseline

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 energy or fatigue at 6 weeks; Group 1: mean 20.3 (SD 14.5); n=20, Group 2: mean 18.2 (SD 16.2); n=20; SF36 energy or fatigue 0-100, High=Top is good outcome; Comments: 19/14.7 baseline
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: SF-36 pain at 6 weeks; Group 1: mean 49.9 (SD 25.2); n=20, Group 2: mean 50.5 (SD 22.1); n=20; SF36 pain 0-100, High=Top is good outcome; Comments: 46.1/49.2 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue VAS at 6 weeks; Group 1: mean 7.5 (SD 1.2); n=20, Group 2: mean 7.5 (SD 2.2); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 7.4/7.1 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 3: Cognitive function at longest follow up available
- Actual outcome for adults; severity mixed or unclear: inability to concentrate VAS at 6 weeks; Group 1: mean 5.2 (SD 2.5); n=20, Group 2: mean 5.8 (SD 2.6); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 6.1/6.1 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: same for outcome at baseline; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: forgetfulness VAS at 6 weeks; Group 1: mean 4.7 (SD 2.7); n=20, Group 2: mean 5.6 (SD 2.3); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 5.9/6.2 at baseline
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to

affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: confusion VAS at 6 weeks; Group 1: mean 4.3 (SD 2.7); n=20, Group 2: mean 4.4 (SD 2.4); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 5.1/5.4 at baseline
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: reaction time (s) at 6 weeks; Group 1: mean 0.35 (SD 0.07); n=20, Group 2: mean 0.36 (SD 0.08); n=20; Comments: 0.35/0.37 at baseline
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 4: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Positive and negative effect scale (PANAS) positive affect at 6 weeks; Group 1: mean 22.7 (SD 8.3); n=20, Group 2: mean 21.7 (SD 6.7); n=20; PANAS 10-50; High=Top is good outcome; Comments: 22.9/22.7 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: same for outcome at baseline; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 5: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: muscle pain VAS at 6 weeks; Group 1: mean 5.8 (SD 3.1); n=20, Group 2: mean 5.9 (SD 2.4); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 6.1/5.9 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome
- Actual outcome for adults; severity mixed or unclear: joint pains VAS at 6 weeks; Group 1: mean 4.8 (SD 3.8); n=20, Group 2: mean 5.1 (SD 2.9); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 5.1/4.3 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 6: Sleep quality at longest follow up available

- Actual outcome for adults; severity mixed or unclear: Unrefreshing sleep VAS at 6 weeks; Group 1: mean 7.7 (SD 2); n=20, Group 2: mean 8.2 (SD 1.8); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 8.2/7.1 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 7: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: adverse events causing withdrawal from study at 6 weeks; Group 1: 0/20, Group 2: 2/20; Comments: racing pulse and severe headache
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: same for outcome at baseline; Group 1 Number missing: 5, Reason: withdrawal due to worsening of symptoms, family problems, rest unclear; Group 2 Number missing: 5, Reason: ovarian surgery unrelated to treatment, rest unclear
- Actual outcome for adults; severity mixed or unclear: adverse events at 6 weeks; Group 1: 4/20, Group 2: 4/20; Comments: fludrocortisone arm: chest tightness/severe headache, severe headache, others unclear (6 events, some patients experienced multiple events); placebo arm: racing pulse, severe headache, others unclear 5 events, some patients experienced multiple events).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: same for outcome at baseline; Group 1 Number missing: 5, Reason: withdrawal due to worsening of symptoms, family problems, rest unclear; Group 2 Number missing: 5, Reason: ovarian surgery unrelated to treatment, rest unclear
Protocol outcome 8: Activity level at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Distance until exhausted at 6 weeks; Group 1: mean 2.7 (SD 1); n=20, Group 2: mean 2.7 (SD 1.3); n=20; Distance before exhausted 1-5; High=Top is good outcome; Comments:2.5/2.5 at baseline; 1=1 block, 2=1 to 3 blocks, 3=3 to 8 blocks, 4=1 to 3 miles, 5=3 miles or more
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: No difference in outcome at baseline; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 9: Exercise performance measure at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Treadmill time (mins) at 6 weeks; Group 1: mean 22.8 (SD 9.2); n=20, Group 2: mean 20.2 (SD 11.5); n=20; Comments: 19.3/20 at baseline
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome

Protocol outcome 10: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: headaches VAS at 6 weeks; Group 1: mean 6 (SD 2.6); n=20, Group 2: mean 6 (SD 2.4); n=20; Visual analogue

scale 0-10; High=Top is poor outcome; Comments: 6/6.2 at baseline Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). Favours drug so overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome - Actual outcome for adults; severity mixed or unclear: painful lymph nodes VAS at 6 weeks; Group 1: mean 3.5 (SD 3.3); n=20, Group 2: mean 3.7 (SD 3.5); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 4/3.9 at baseline Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome - Actual outcome for adults; severity mixed or unclear: sore throat VAS at 6 weeks; Group 1: mean 3.1 (SD 2.1); n=20, Group 2: mean 3.3 (SD 3); n=20; Visual analogue scale 0-10; High=Top is poor outcome; Comments: 3.2/3 at baseline Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There were differences for the outcome at baseline, and the final outcome values not adjusted (p values for adjusted analysis though). But favours placebo so not overestimating treatment effect; Group 1 Number missing: 3, Reason: 3 withdrawn due to worsening symptoms - likely to affect outcome; Group 2 Number missing: 1, Reason: ovarian surgery - unlikely related to outcome	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Physical functioning at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Care needs at longest follow up available; Impact on families and carers at longest follow up available

Study	**Randall 2005[68]**
Study type	RCT (Patient randomised; Crossover: 2 weeks; half-life is 15 hours, so will have dropped by a factor of 2 to the power 22 so to < 1/4,200,000 of the starting dose; therefore appropriate in terms of the drug in system)
Number of studies (number of participants)	1 (n=14)
Countries and setting	Conducted in United Kingdom; Setting: unclear
Line of therapy	Not applicable

Duration of study	Intervention time: 20 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fukuda
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Fukuda criteria for CFS; age 18-70; restless legs syndrome rating scale <10; MMSE >=26; surgically sterile, 2 years post-menopausal, non-pregnant, non-lactating, using a method of birth control.
Exclusion criteria	Any clinical condition explaining chronic fatigue; current major depressive disorder; LV hypertrophy; symptomatic mitral valve prolapse; hypertension (sbp >160 mmHg); and disorder interfering with drug absorption, distribution, metabolism or excretion; history of alcohol/drug abuse; sight, hearing or movement problems; colour blindness; >8 cups of coffee per day; familiarity with the cognitive tests used in study;
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 41.2(3.3). Gender (M:F): 7:7. Ethnicity: unclear
Further population details	-
Extra comments	MMSE 29.2; Epworth sleepiness 9.1; MSLT mean sleep latency 16.4min; HADS A 5.1; HADS D 5.1; CGI-S 4.1; illness duration 5.4yrs; caffeine 3.3 cups/day; alcohol 4.1 units per week; sbp 120.6; dbp 78; pulse 72.2; RR 16.6; weight 75.7kg; temperature 36.6C.
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=14) Intervention 1: modafinil. 200mg modafinil; dose increased slowly at 3 day intervals starting at 100mg until required dose reached. Duration 20 days. Concurrent medication/care: Patients used a medication diary. Indirectness: No indirectness

	(n=14) Intervention 2: modafinil. 400mg. Dose increased from stating dose of 100mg at 3 day intervals as tolerated. Duration 20 days. Concurrent medication/care: medication diary given. Indirectness: No indirectness (n=14) Intervention 3: placebo. Identical doses of placebo. Duration 20 days. Concurrent medication/care: Medication diary used. Indirectness: No indirectness
Funding	Study funded by industry (Cephalon UK - unrestricted grant)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: MODAFINIL versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF36 vitality at 20 days; Group 1: mean 29.6 (SD 26.55); n=14, Group 2: mean 26.1 (SD 23.94); n=14; SF36 vitality 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 vitality at 20 days; Group 1: mean 21.4 (SD 20.57); n=14, Group 2: mean 26.1 (SD 23.94); n=14; SF36 vitality 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 Physical Function at 20 days; Group 1: mean 55.4 (SD 29.12); n=14, Group 2: mean 53.6 (SD 27.3); n=14; SF36 physical function 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 Physical Function at 20 days; Group 1: mean 48.6 (SD 29.55); n=14, Group 2: mean 53.6 (SD 27.3); n=14; SF36 physical function 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia,

nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 Physical role limitation at 20 days; Group 1: mean 10.7 (SD 27.3); n=14, SF36 physical role 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 Physical role limitation at 20 days; Group 1: mean 19.2 (SD 32.16); n=14, Group 2: mean 21.4 (SD 32.16); n=14; SF36 physical role 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 emotional role limitation at 20 days; Group 1: mean 66.1 (SD 47.5); n=14, Group 2: mean 95.2 (SD 11.97); n=14; SF36 emotional role 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 emotional role limitation at 20 days; Group 1: mean 85.7 (SD 31.4); n=14, Group 2: mean 95.2 (SD 11.97); n=14; SF36 emotional role 0-100, High=Top is good outcome; Comments: 400mg modafinil

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 mental health at 20 days; Group 1: mean 68 (SD 21.7); n=14, Group 2: mean 74.9 (SD 12.34); n=14; SF36 mental health 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 mental health at 20 days; Group 1: mean 69.2 (SD 19.07); n=14, Group 2: mean 74.9 (SD 12.34); n=14; SF36 mental health 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0

- Actual outcome for adults; severity mixed or unclear: SF36 pain at 20 days; Group 1: mean 59.5 (SD 32.91); n=14, Group 2: mean 57.2 (SD 30.67); n=14; SF36 pain 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 pain at 20 days; Group 1: mean 50 (SD 32.16); n=14, Group 2: mean 57.2 (SD 30.67); n=14; SF36 pain 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 social at 20 days; Group 1: mean 43.7 (SD 26.57); n=14, Group 2: mean 43.7 (SD 30.68); n=14; SF36 social 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 social at 20 days; Group 1: mean 38.9 (SD 32.16); n=14, Group 2: mean 43.7 (SD 30.68); n=14; SF36 social 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 general health at 20 days; Group 1: mean 50.1 (SD 23.57); n=14, Group 2: mean 49.2 (SD 21.7); n=14; SF36 general health 0-100, High=Top is good outcome; Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: SF36 general health at 20 days; Group 1: mean 47.5 (SD 19.83); n=14, Group 2: mean 49.2 (SD 21.7); n=14; SF36 general health 0-100, High=Top is good outcome; Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Chalder's physical fatigue at 20 days; Group 1: mean 12.6 (SD 7.1); n=14, Group 2: mean 13.6 (SD 7.85); n=14; Chalder's physical fatigue 0-21, High=Top is poor outcome. Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Chalder's physical fatigue at 20 days; Group 1: mean 14.1 (SD 4.49); n=14, Group 2: mean 13.6 (SD 7.85); n=14; Chalder's physical fatigue 0-21, High=Top is poor outcome. Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Chalder's mental fatigue at 20 days; Group 1: mean 7.2 (SD 3.74); n=14, Group 2: mean 7.4 (SD 2.99); n=14; Chalder's mental fatigue 0-21, High=Top is poor outcome. Comments: 200mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Chalder's mental fatigue at 20 days; Group 1: mean 8.4 (SD 2.24); n=14, Group 2: mean 7.4 (SD 2.99); n=14; Chalder's mental fatigue 0-21, High=Top is poor outcome. Comments: 400mg modafinil. A more appropriate paired analysis [ie MD(SE)] not available. However the current analysis will be conservative in that it will tend to underestimate the precision. sds calculated from SE in paper
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: adverse effects including headaches, light headedness, disturbed balance, palpitations, insomnia, nausea, increased tendency to cry and constipation (patient 1); headaches, heavy legs and lethargy (patient 2); headaches (patient 3).; Group 2 Number missing: 0

Protocol outcome 3: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Any adverse events at 20 days; Group 1: 9/14, Group 2: 8/14; Comments: 200mg modafinil
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Any adverse events at 20 days; Group 1: 12/14, Group 2: 8/14; Comments: 400mg modafinil
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Mortality at longest follow up available; Cognitive function at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up; Psychological status at longest follow up available; Physical function at longest follow up available; Pain at longest follow up available

Study	**RituxME trial: Fluge 2019[23]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=152)
Countries and setting	Conducted in Norway; Setting: 4 university hospitals and 1 general hospital in Norway
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Canadian consensus criteria; patients where the workup uncovers other pathology as a possible cause of symptoms were excluded
Stratum	adults; severity mixed or unclear: aged 18 to 65 years; ME/CFS according to Canadian consensus criteria; mild or mild/moderate 40%, moderate 30%, moderate/severe and severe 30%; patients with very severe ME/CFS (WHO function class IV), who were totally bedridden and in need of care were excluded
Subgroup analysis within study	Not applicable: NA
Inclusion criteria	Patients with ME/CFS according to Canadian criteria of 2003; disease duration: 2-15 years; for patients with mild ME/CFS disease duration must be a minimum of 5 years; severity: mild, mild/moderate, moderate, moderate/severe and severe ME/CFS; age: 18-65 years; signed informed consent

Exclusion criteria	Patients with fatigue, who do not comply with the diagnostic Canadian criteria for ME/CFS or disease duration < 24 months or > 15 years; patients where the workup uncovers other pathology as a possible cause of symptoms; patients with very severe ME/CFS (WHO function class IV), who are totally bedridden and in need of care; pregnancy or breast feeding; previous cancer (except basal cell carcinoma of the skin or cervix dysplasia); previous long-term systemic treatment with immunosuppressive agents (Imurel, Sandimmun, Cellcept), except steroid treatments for e.g. obstructive lung disease or other autoimmune diseases like ulcerative colitis; serious endogenous (primary) depression; lack of ability to complete the study including follow-up; known serious multi-allergy, clinically assessed with an elevated risk of allergic reactions during rituximab infusion; reduced kidney function (creatinine > 1.5 x reference area); reduced liver function (bilirubin > 1.5 x reference area, or transaminase > 1.5 x reference area); known HIV-positivity, previous hepatitis B or hepatitis C, or reason to suspect other ongoing and clinically relevant infection; known immunodeficiency disorders with an elevated risk involved in therapeutic B lymphocyte depletion, e.g. hypogammaglobulinemia
Recruitment/selection of patients	referrals from physicians or direct requests from patients or their relatives to be evaluated for future clinical trials
Age, gender and ethnicity	Age - Mean (SD): Rituximab 37.8 (11.4), Placebo 35.5 (11.2) years. Gender (M:F): 27/124. Ethnicity: not reported
Further population details	-
Extra comments	NA
Indirectness of population	No indirectness: NA
Interventions	(n=77) Intervention 1: immunomodulatory drugs - rituximab. Induction treatment with 2 infusions, 2 weeks apart, of rituximab (MabThera, Roche), 500 mg/m2 of body surface area (maximum of 1000 mg). In the maintenance phase, patients received a 500-mg fixed dose of rituximab at 3, 6, 9, and 12 months. Duration 12 months. Concurrent medication/care: One hour before infusions, all patients received premedication with1g of oral acetaminophen, 10 mg of cetirizine, and 8mg of dexamethasone. Indirectness: No indirectness; Indirectness comment: NA Comments: NA (n=75) Intervention 2: placebo. Induction treatment with 2 infusions, 2 weeks apart, of 500 mg/m2 of body surface area (maximum of 1000 mg) saline with added human albumin (Flexbumin [Baxalta] or Albunorm [Octapharma]), 0.4 mg/mL, to ensure no visible difference from the active comparator. In the maintenance phase, patients received a 500-mg fixed dose of saline with human albumin at 3, 6, 9, and 12 months. Duration 12 months. Concurrent medication/care: One hour before infusions, all patients received premedication with1gof oral acetaminophen, 10 mg

	of cetirizine, and 8mg of dexamethasone. Indirectness: No indirectness; Indirectness comment: NA Comments: NA
Funding	Academic and government funding (The Norwegian Research Council, Norwegian Regional Health Trusts, Kavli Trust, MEandYou Foundation, and Norwegian ME Association.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RITUXIMAB versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue severity scale at 18 months ; MD; -0.07 (95%CI -3.21 to 3.08) (p value: 0.68) NA Fatigue severity scale 9-63 Top=High is poor outcome, Comments: Baseline values: Rituximab 59.1 (6.7), placebo 59.88 (3.3);
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew
- Actual outcome for adults; severity mixed or unclear: Fatigue score (NRS) at 16-20 months ; MD; -0.06 (95%CI -0.51 to 0.39) (p value : 0.79) NA Fatigue NRS 0-10 Top=High is poor outcome, Comments: Baseline values: rituximab 3, placebo 3;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew

Protocol outcome 2: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF36 physical functioning at 24 months ; MD; 1.24 (95%CI -7.38 to 9.86) (p value : 0.68) NA SF36 physical functioning 0-100 Top=High is good outcome, Comments: Baseline values: Rituximab 35.24 (21.9), placebo 32.45 (19.1);
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew
- Actual outcome for adults; severity mixed or unclear: Function level % at 16-20 months ; MD; -0.68 (95%CI -5.9 to 4.54) (p value: 0.31) NA function percentage 0-100 Top=High is good outcome, Comments: Baseline values: rituximab 20.14 (11.5), placebo 18.37 (8.8);
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew

Protocol outcome 3: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Adverse events with possible or probable relation to intervention at 24 months ; Group 1: 26/77, Group 2: 12/74
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew - Actual outcome for adults; severity mixed or unclear: Serious adverse events with possible or probable relation to intervention at 24 months ; Group 1: 8/77, Group 2: 0/74; Comments: 4 admissions in 2 patients were due to febrile neutropenia, 2 admissions in 1 patient were due to dizziness and nausea, and 1 admission in 1 patient was due to headache and gastroenteritis. Two patients had infusion-related reactions, 1 of whom was also admitted for tests because of noncardiac chest pain. One patient was admitted for examination for involuntary movements, and another for a transient facial paresis. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew - Actual outcome for adults; severity mixed or unclear: Suspected unexpected serious adverse reaction at 24 months ; Group 1: 2/77, Group 2: 1/74; Comments: 1 metrorrhagia with hysterectomy and 1 suspected but unconfirmed coronary disease in the rituximab group, 1 transient paresis in left extremities in the placebo group Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew Protocol outcome 4: Activity levels at longest follow up available - Actual outcome for adults; severity mixed or unclear: Mean steps per 24 hours at 17-21 months ; MD; -127 (95%CI -1004 to 749) (p value : 0.58) number of steps , Comments: Baseline values: rituximab 3297 (2047), placebo 3233 (2099); Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: NA; Group 1 Number missing: 1, Reason: patient withdrawal; Group 2 Number missing: 2, Reason: 1 patient excluded, did not receive intervention (withdrew consent), 1 patient withdrew
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Roerink 2017[70]**
Study type	RCT (Patient randomised; Parallel)

Number of studies (number of participants)	1 (n=50)
Countries and setting	Conducted in Netherlands; Setting: Outpatient department, referred from regional hospitals, CFS treatment centers and a Dutch patient advocacy foundation.
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 24 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: CDC criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	female patients with CFS fulfilling CDC criteria; 18-59 years; maximum fatigue duration of 10 years or recent progression of symptoms; minimum score of 40 on fatigue severity scale of CIS; SIP score of at least 700.
Exclusion criteria	Females who are pregnant or nursing, intend to get pregnant during the study, use or have used psychotropic medication in the past month, received a live vaccine during the last 4 weeks, had substance abuse in the past 3 months, have had symptoms more than 10 years, are taking any medication except oral contraceptives and/or paracetamol, have current engagement in CFS research, do not have the ability to understand the nature and the extent of the trial and the procedure required, have psychiatric conditions (major depression, psychosis, eating disorders, anxiety disorders, bipolar disease and post-traumatic stress disorder).
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Median: anakinra 30, placebo 32. Gender (M:F): No male patients. Ethnicity: unclear
Further population details	-
Extra comments	anakinra/placebo: illness duration 44/39 months; BMI 25/25; CIS fatigue 52/51; SIP 1647/1706; SF36 social functioning 33/39; SF36 physical function 48/56; SCL-90 152/148; VAS max pain 7/7; mean cdc symptoms 7/6

Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=25) Intervention 1: pro-inflammatory cytokine antagonists - Anakinra. Anakinra (Interleukin-1 receptor antagonist) 100mg subcutaneously per day. Duration 4 weeks. Concurrent medication/care: Each participant provided with a box containing 28 syringes and supplies of drug. Patients instructed by physician on how to administer. Daily alarm used to assist compliance, along with adherence monitoring. Indirectness: No indirectness (n=25) Intervention 2: placebo. Identical placebo given in identical doses intramuscularly. Duration 4 weeks. Concurrent medication/care: Each participant provided with a box containing 28 syringes and supplies of drug. Patients instructed by physician on how to administer. Daily alarm used to assist compliance, along with adherence monitoring. Indirectness: No indirectness
Funding	Academic or government funding (Interleukin Foundation. Drugs provided by Swedish Orphan Biovitrium)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ANAKINRA versus PLACEBO

Protocol outcome 1: Mortality at longest follow up available
- Actual outcome for adults; severity mixed or unclear: death at 24 weeks; Group 1: 0/25, Group 2: 0/25
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: CIS fatigue at 24 weeks; MD; 1.3 (95%CI -5.3 to 8); Checklist individual strength fatigue 8-56, High=Top is poor outcome; Comments: Based on ANCOVA, adjusting for baseline;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF36 physical function at 24 weeks; MD; -4 (95%CI -15.1 to 7.1, SF36 physical function 0-100, High=Top is good outcome. Comments: Based on ANCOVA, adjusting for baseline;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 4: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Symptom Checklist 90 at 24 weeks; MD; 3 (95%CI -8.6 to 14.6); Symptom checklist 90 scale 90-450, High=Top is poor outcome; Comments: Based on ANCOVA, adjusting for baseline;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 5: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: VAS maximum pain score at 24 weeks; MD; 0.34 (95%CI -1.1 to 1.7); visual analogue scale 0-10, High=Top is poor outcome; Comments: Based on ANCOVA, adjusting for baseline;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 6: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: adverse events at 24 weeks; Group 1: 24/25, Group 2: 14/25
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for adults; severity mixed or unclear: withdrawal due to adverse events at 24 weeks; Group 1: 1/25, Group 2: 0/25
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 7: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: Sickness Impact profile at 24 weeks; MD; 91.2 (95%CI -275.8 to 458.1); Sickness impact profile 0-5799, High=Top is poor outcome; Comments: Based on ANCOVA, adjusting for baseline;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Baseline details: Though some baseline discrepancies all outcomes adjusted for baseline; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life at longest follow up available; Cognitive function at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Rowe 2001[77]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in USA; Setting: Two tertiary referral centres in the USA
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 11 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1994 CDC
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 18-50; CDC 1994 criteria
Exclusion criteria	History of conditions/drugs that could be exacerbated by fludrocortisone or tilt table testing; previous fludrocortisone use at doses > 106 mg/day; enrolment in another CFS study; psychiatric conditions requiring therapy; alcohol or substance abuse
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 36.2 to 37.3. Gender (M:F): 33:66. Ethnicity: White 98%; no other information provided

Further population details	-
Extra comments	FCA/placebo: age >30 76%/82%; currently working 56%/53%; on disability 20%/8%; duration of CFS 6.9y/6y; at least moderate severity of illness (score of >65 on global wellness scale)
Indirectness of population	Very serious indirectness: All participants had neurally-mediated hypotension and 1994 CDC criteria used; PEM not a compulsory feature.
Interventions	(n=50) Intervention 1: oral corticosteroids - fludrocortisone/hydrocortisone/other. Fludrocortisone starting at a dose of 0.025 mg/day (1 capsule) for a week, then 0.05 mg/day (2 capsules) for the following week, and eventually increased to 0.1 mg/day (4 capsules) for remaining 7 weeks. Duration 9 weeks. Concurrent medication/care: Patients advised to drink at least 2L of fluid per day and to keep normal NaCl intake to their usual levels. Both groups also had KCl tablets 10mEq/day for duration of treatment. If AEs emerged, doses were reduced to previously tolerated levels. Indirectness: No indirectness (n=50) Intervention 2: placebo. identical capsules containing methylcellulose only given exactly as the study drug in the same dose increments. Duration 9 weeks. Concurrent medication/care: Patients advised to drink at least 2L of fluid per day and to keep normal NaCl intake to their usual levels. Both groups also had KCl tablets 10mEq/day for duration of treatment. If AEs emerged, doses were reduced to previously tolerated level. Indirectness: No indirectness
Funding	Academic or government funding (National institute of Allergy and Infectious Diseases; NIH; CFIDAA Inc)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: FLUDROCORTISONE/HYDROCORTISONE/OTHER versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Wood Mental Fatigue Inventory at 11 weeks; Group 1: mean 14.1 (SD 10.9); n=38, Group 2: mean 13.3 (SD 9.6); n=45; Wood mental fatigue inventory 0-36; High=poor outcome; Comments: 16.3/18.3 at baseline, which supports placebo as better
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5
- Actual outcome for adults; severity mixed or unclear: POMS vigour subscale at 11 weeks; Group 1: mean 8.8 (SD 6.1); n=38, Group 2: mean 8.6 (SD 6.7); n=45; POMS

vigour subscale 0-32, High=Top is good outcome; Comments: 7.9/6.7 at baseline - favours study drug and thus explains follow up result
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5
- Actual outcome for adults; severity mixed or unclear: POMS fatigue subscale at 11 weeks; Group 1: mean 16.2 (SD 7.3); n=38, Group 2: mean 16.4 (SD 7.9); n=45; POMS fatigue subscale 0-28, High=Top is poor outcome; Comments: 19.6/21.3 - favours study drug at baseline so explains follow up result
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcome 2: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: SF 36 Physical function at 11 weeks; Group 1: mean 58.9 (SD 21.9); n=38, Group 2: mean 51.4 (SD 27.8); n=45; SF36 physical function 0-100, High=Top is good outcome; Comments: 54.8/45.1 at baseline so favours study drug which explains follow up result
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcome 3: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Beck Depression Inventory at 11 weeks; Group 1: mean 10.4 (SD 7.2); n=38, Group 2: mean 10.8 (SD 6.8); n=45; Beck depression inventory 0-63, High=Top is poor outcome; Comments: Baseline 14.7/15, so baseline discrepancy explains follow up result
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5
- Actual outcome for adults; severity mixed or unclear: SF 36 mental health at 11 weeks; Group 1: mean 68.6 (SD 19.1); n=38, Group 2: mean 69.8 (SD 16.3); n=45; SF36 mental health 0-100, High=Top is good outcome. Comments: 63.7/66.3 - favours placebo so this may explain follow up result favouring placebo
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcome 4: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: adverse effects at 11 weeks; Group 1: 23/38, Group 2: 32/45; Comments: 61% of drug patients and 71% of

placebo patients had at least 1 adverse effect. Denominators not stated so assumed it is the denominators given for efficacy outcomes.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcome 5: Activity levels at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Duke activity Status Index at 11 weeks; Group 1: mean 9.2 (SD 10.6); n=38, Group 2: mean 6.7 (SD 7.3); n=45; Duke activity status index 0-58.2, High=Top is good outcome. Comments: 7.8/5 at baseline - favours study drug and this explains follow up result
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcome 6: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: Wellness Score at 11 weeks; Group 1: mean 3.8 (SD 11.5); n=38, Group 2: mean 2.7 (SD 10); n=45; Wellness score scale not reported, High=Top is good outcome; Comments: 46.8/40.7 at baseline – favours study drug
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 12, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 12; Group 2 Number missing: 5, Reason: list of reasons given but these include some of those where data were available so not able to report reasons for these specific 5

Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Snorrason 1996[79]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=49)

Countries and setting	Conducted in Iceland; Setting: Outpatient clinics of the National University Hospital of Iceland and a rheumatological
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Inadequate method of assessment/diagnosis: Not using a recognised consensus-based set of criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	CFS, defined as symptoms of fatigue for >50% of time and lasting >6 months, major sleep disturbances and myalgia; minor psychiatric symptoms allowed
Exclusion criteria	Medical conditions known to produce symptoms of fatigue; major psychiatric diagnosis
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 43.44 to 44.46. Gender (M:F): 7:42. Ethnicity: unclear
Further population details	-
Extra comments	Galnathamine/placebo: duration of illness 13.68/11.79 years
Indirectness of population	Serious indirectness: Downgraded for unclear CFS criteria.
Interventions	(n=25) Intervention 1: Galantamine hydrobromide. 10mg 3x daily, reached by a schedule of escalating dosage. Duration 8 weeks. Concurrent medication/care: Optional cross-over design added to parallel group RCT design - patients could cross-over after 2 weeks if failed to improve or had symptoms worsening. Indirectness: No indirectness (n=24) Intervention 2: placebo. placebo 3 x daily. Duration 8 weeks. Concurrent medication/care: Optional cross-over design added to parallel group RCT design - patients could cross-over after 2 weeks if failed to improve or had

	symptoms worsening. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GALANTAMINE HYDROBROMIDE versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue VAS at 2 weeks; Group 1: mean 7.25 (SD 2.1); n=25, Group 2: mean 7.11 (SD 1.35); n=24; Fatigue visual analogue scale range not reported, High=Top is poor outcome; Comments:7.72/7.41 at baseline
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 2: Cognitive function at longest follow up available
- Actual outcome for adults; severity mixed or unclear: memory VAS at 2 weeks; Group 1: mean 5.63 (SD 3.16); n=25, Group 2: mean 4.72 (SD 2.46); n=24; Memory visual analogue scale, High=Top is Comments: baseline 4.86/5.22
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 3: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Myalgia VAS at 2 weeks; Group 1: mean 7.52 (SD 1.97); n=25, Group 2: mean 7.99 (SD 1.26); n=24; Myalgia visual analogue scale, High=Top is poor outcome; Comments: 8.57/8.56 at baseline
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 4: Sleep quality at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Sleep disturbance VAS at 2 weeks; Group 1: mean 7 (SD 2.35); n=25, Group 2: mean 6.66 (SD 2.49); n=24; Sleep disturbance visual analogue scale, High=Top is poor outcome; Comments: baseline 7.52/7.77 - goes against follow up scores
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 5: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: AEs dizziness VAS at 2 weeks; Group 1: mean 4.26 (SD 2.77); n=25, Group 2: mean 3.54 (SD 3.12); n=24; Comments: 3.95/2.95 at baseline

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 6: Return to school or work at longest follow up available - Actual outcome for adults; severity mixed or unclear: Work capacity/satisfaction on VAS at 2 weeks; Group 1: mean 4.92 (SD 2.15); n=25, Group 2: mean 5.09 (SD 1.67); n=24; Work capacity/satisfaction visual analogue scale, High=Top is good outcome; Comments: baseline 4.81/5.25 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Psychological status at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Steinberg 1996[84]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in USA; Setting: patients recruited from a patient CFS registry in USA
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 2 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Holmes criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable

Inclusion criteria	CFS defined by Holmes criteria
Exclusion criteria	None
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 36.2(11.4). Gender (M:F): 7:23. Ethnicity: unclear
Further population details	-
Extra comments	Atopic history 73.3%; immediate skin test reactivity 53.3%
Indirectness of population	Serious indirectness: Holmes 1988 criteria used; PEM not a compulsory feature
Interventions	(n=15) Intervention 1: Antihistamines - terfenadine. 60mg terfenadine twice daily. Duration 2 months. Concurrent medication/care: Permitted to take oral contraceptives, antibiotics, vitamins, aspirin, NSAIDS, beta adrenergic agonists; not permitted to take antihistamines, decongestants, tricyclic antidepressants or ENT anti-inflammatory agents. Indirectness: No indirectness (n=15) Intervention 2: placebo. placebo twice daily. Duration 2 months. Concurrent medication/care: Permitted to take oral contraceptives, antibiotics, vitamins, aspirin, NSAIDS, beta adrenergic agonists; not permitted to take antihistamines, decongestants, tricyclic antidepressants or ENT antiinflammatory agents. Indirectness: No indirectness
Funding	Study funded by industry (Marion Merrell Dow Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TERFENADINE versus PLACEBO

Protocol outcome 1: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: modified Medical Outcome Study Short Form - physical functioning at 2 months; Group 1: mean 63.1 (SD 17.52); n=14, Group 2: mean 69.66 (SD 18.09); n=14; Medical Outcome Study Short Form - physical functioning 0-100, High=Top is good outcome; Comments: 60.32/64.53 at baseline, favouring placebo - this may explain follow up results
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 1, Reason: perception of no improvement; Group 2 Number missing: 1, Reason: perception of no improvement Protocol outcome 2: Psychological status at longest follow up available - Actual outcome for adults; severity mixed or unclear: modified Medical Outcome Study Short Form - mental health at 2 months; Group 1: mean 63.89 (SD 21.36); n=14, Group 2: mean 74.62 (SD 15.31); n=14; Medical Outcome Study Short Form - mental health 0-100, High =Top is good outcome; Comments: 64.29/77.18 at baseline - explains follow up result Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: see results section; Group 1 Number missing: 1, Reason: perception of no improvement; Group 2 Number missing: 1, Reason: perception of no improvement	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Straus 1988[86]**
Study type	RCT (Patient randomised; Crossover: 37 days)
Number of studies (number of participants)	1 (n=27)
Countries and setting	Conducted in USA; Setting: National Institutes of Health
Line of therapy	Not applicable
Duration of study	Intervention time: 37 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Holmes CDC criteria

Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	CFS by Holmes CDC criteria; no other medical diagnosis explaining symptoms; titres of antibodies to diffuse or restricted early antigens of Epstein barr virus of >=1:40 or to lack antibodies to EBNA <1:2)
Exclusion criteria	None reported
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): 34.1(7.5). Gender (M:F): 8:19. Ethnicity: unclear
Further population details	-
Extra comments	Duration of illness 6.8yrs; years of education 14.9yrs; vocationally disabled 12/27; working part time 10/27; single or divorced 14/27; with children 7/27
Indirectness of population	Serious indirectness: Holmes 1988 criteria used; PEM not a compulsory feature
Interventions	(n=27) Intervention 1: antiviral drugs - acyclovir. IV acyclovir (500mg per square metre) infused over a period of 60 minutes in 150ml of saline every 8 hrs for 7 days of hospitalisation. Vigorous oral hydration encouraged. Then discharged to take 800mg acyclovir tablets for 30 days. Duration 37 days. Concurrent medication/care: Cross-over - order of acyclovir/placebo randomised. Indirectness: No indirectness (n=27) Intervention 2: placebo. IV placebo (500mg per square metre) infused over a period of 60 minutes in 150ml of saline every 8 hrs for 7 days of hospitalisation. Vigorous oral hydration encouraged. Then discharged to take 800mg placebo tablets for 30 days. Duration 37 days. Concurrent medication/care: Cross-over - order of acyclovir/placebo randomised. Indirectness: No indirectness
Funding	Study funded by industry (Burroughs Wellcome)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACYCLOVIR versus PLACEBO

Protocol outcome 1: Psychological status at longest follow up available
- Actual outcome for young people; severity mixed or unclear: POMS anxiety at 37 days; MD; 2.92 (95%CI 0.6334 to 5.2066); Profile of mood states anxiety 0-36, High=Top is poor outcome; Comments: paired analysis.;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0
- Actual outcome for young people; severity mixed or unclear: POMS depression at 37 days; MD; 3.97 (95%CI 0.6946 to 7.2454); Profile of mood states depression 0-60, High=Top is poor outcome; Comments: Paired analysis;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0
- Actual outcome for young people; severity mixed or unclear: POMS anger at 37 days; MD; 2.30 (95%CI -0.1308 to 4.7308); Profile of mood states anger 0-48, High=Top is poor outcome; Comments: paired analysis;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0
- Actual outcome for young people; severity mixed or unclear: POMS vigour at 37 days; MD; -2.05 (95%CI -4.6456 to 0.5456); Profile of mood states vigour 0-32, High=Top is good outcome; Comments: paired analysis;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0
- Actual outcome for young people; severity mixed or unclear: POMS fatigue at 37 days; MD; 1.26 (95%CI -1.006 to 3.526); Profile of mood states fatigue 0-28, High=Top is poor outcome; Comments: paired analysis;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0
- Actual outcome for young people; severity mixed or unclear: POMS confusion at 37 days; MD; 1.83 (95%CI 0.5734 to 3.0866); Profile of mood states confusion 0-28, High=Top is poor outcome; Comments: paired analysis;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0

Protocol outcome 2: Adverse events at longest follow up available - Actual outcome for young people; severity mixed or unclear: AEs - reversible renal failure at 37 days; Group 1: 3/27, Group 2: 0/27 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Activity levels at longest follow up available - Actual outcome for young people; severity mixed or unclear: rest (hrs/day) at 37 days; MD; -0.05 (95%CI -0.8328 to 0.7328, Comments: paired analysis); Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this would have greatly affected the mean POMs and wellness scores; Group 2 Number missing: 0 Protocol outcome 4: Symptom scales at longest available follow up - Actual outcome for young people; severity mixed or unclear: Wellness score at 37 days; MD; -1.08 (95%CI -7.2806 to 5.1206); Wellness scale not reported; Comments: paired analysis; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: renal failure - had they been included this may have affected the mean POMs and wellness scores; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Strayer 2012[88]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=234)
Countries and setting	Conducted in USA; Setting:

Line of therapy	Not applicable
Duration of study	Intervention + follow up: Stage I (extracted here) 42 weeks. Stage II (placebo group crossover) 24 weeks.
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Patients were diagnosed with CFS according to CDC diagnosis criteria (Holmes 1988)
Stratum	adults; severity mixed or unclear: Patients were stratified according to treadmill duration (≥9 minutes vs >9 minutes) then randomised.
Subgroup analysis within study	Stratified then randomised
Inclusion criteria	1. A diagnosis of CFS, as defined by the Center for Disease Control (1988 CDC case definition) ≥12 months; 2. Age range: ≥18 years old, ≤60 years old; 3. Males or non-pregnant, non-lactating females: Females must be of non-child-bearing potential (either post-menopausal for two years or surgically sterile, including tubal ligation) or using an effective means of contraception (birth control pills, intrauterine device, diaphragm). Females who are less than two years post-menopausal, those with tubal ligations, and those using contraception must have a negative serum pregnancy test within the two weeks prior to the first study medication infusion. Females of child-bearing potential agree to use an effective means of contraception from four weeks prior to the baseline pregnancy test until four weeks after the last study medication infusion. 4. A reduced quality of life as determined by a documented KPS of 40 to 60 on three occasions, each at least 14 days apart, during the twelve weeks immediately preceding the start of study drug infusions. The KPS must be rounded in increments of ten. 5. Ability to walk (minimum of 20 seconds) on the moving treadmill (grade=0%; belt speed=1 mph) on a minimum of two occasions during the twelve weeks immediately preceding study entry. 6. Laboratory documentation (baseline or historical following onset of CFS) of a negative antinuclear antibody or a negative anti-ds DNA, a negative rheumatoid factor, and an erythrocyte sedimentation rate. 7. Laboratory documentation that the patient is euthyroid (patients on thyroid replacement therapy must be on a stable dose during the eight week washout period) based on thyroid profile (T4, T3, TSH, T3 uptake and Free T4 index) performed during baseline. 8. Ability to provide written informed consent indicating awareness of the investigational nature of this study.

Exclusion criteria	1. Inability to return to the investigator's site for scheduled infusions and evaluations during Stages 1 and 2 of the study. 2. Chronic or intercurrent acute medical disorder or disease making implementation or interpretation of the protocol or results difficult or unsafe. 3. Pregnant or lactating females. 4. Treatment with any of the following therapies within the eight weeks immediately preceding the start of study baseline or during baseline: systemic glucocorticoids (ie, hydrocortisone, prednisone, etc.) or mineralocorticoids (ie, fludrocortisone [Florinef], etc.), interferons, interleukin-2, systemic antivirals, gamma globulin, or investigational drugs and experimental agents not yet approved for use in the United States. The patient was to give written consent prior to discontinuation of any drugs listed under this criterion. 5. Prior participation in a study of Poly I:C12U. 6. Medical necessity, as determined by the patient's private doctor or the principal investigator, to continue aspirin (ASA) or non-steroidal anti-inflammatory (NSAID) drugs for 20 consecutive days or for more than 10% of the study duration (i.e., 28 total days for Stage 1 and 17 total days for Stage 2). 7. Ability to exercise over 18 minutes during any of the baseline ET procedures. 8. Evidence or history of any exclusion criteria for the ET testinga. Previous documented evidence of myocardial infarction or recent significant change in the resting electrocardiogram (ECG) suggesting infarction or other acute cardiac events.b. Current symptoms of coronary insufficiency (i.e., angina pectoris and/or ST segment depression on ECG).c. Evidence of uncontrolled atrial or frequent or complex ventricular ectopy, or myocardial conduction defect which would increase the risk of syncope (for example, second degree or higher A-V block).d. History of congestive heart failure, suspected or known dissecting aneurysm, recent systemic or pulmonary embolus, severe valvular heart disease, ventricular aneurysm, active or suspected myocarditis or pericarditis, thrombophlebitis or intracardiac thrombi, or acute infection.e. Evidence of moderate or severe obstructive pulmonary disease.f. Resting diastolic blood pressure >115 mm Hg or resting systolic blood pressure >200 mm Hg.g. Uncontrolled metabolic disease (e.g., diabetes, thyrotoxicosis, or myxedema).h. Concurrent use of any beta blockers and/or bronchodilators which cannot remain at a stable dosage level during the eight- (8-) week washout period and continuing during baseline and Stages 1 and 2. 9. History of alcohol or other substance abuse within two (2) years before the onset of the chronic fatigue and/or at any time afterward. 10. History of suicidal ideation or a suicide attempt within two (2) years of baseline. 11. Any past or current diagnosis of a major depressive disorder with psychotic or melancholic features; bipolar affective disorders; schizophrenia of any subtype; delusional disorders of any subtype; dementias of any subtype; anorexia nervosa; or bulimia nervosa.
Recruitment/selection of patients	353 patients initially signed consents. 46 failed to meet entry criteria and 67 decided to withdraw from the study prior to completing baseline procedures. 240 patients were randomized.

Age, gender and ethnicity	Age - Mean (SD): Rintatolimod group 43.4 (9.2), placebo group 43.5 (10.1). Gender (M:F): 170 females, 64 males. Ethnicity: Not stated.
Further population details	-
Extra comments	Duration of CFS symptoms, mean (SD) years: rintatolimod group 9.6 (5.36), placebo group 9.7 (6.08)
Indirectness of population	Serious indirectness: Holmes 1988 criteria used; PEM is not a compulsory feature
Interventions	(n=117) Intervention 1: immunomodulatory drugs - rintatolimod (Ampligen). Patients initially received a 200 mg IV dose of rintatolimod twice weekly for two weeks. Following this, a 400 mg dose of rintatolimod was administered twice weekly for 40 weeks. Duration 42 weeks. Concurrent medication/care: Not stated.. Indirectness: No indirectness (n=117) Intervention 2: placebo. Patients initially received a 200 mg IV dose of placebo (physiological saline) twice weekly for two weeks. Following this, a 400 mg dose of saline placebo was administered twice weekly for 40 weeks. . Duration 42 weeks. Concurrent medication/care: Not stated. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded, designed, and analyzed by Hemispherx Biopharma with oversight by the Food & Drug Administration (FDA) including statistical analysis. Following completion of FDA audits, the decision to publish was made.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RINTATOLIMOD (AMPLIGEN) versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults - severe: Vitality score (SF36) at 42 weeks; Group 1: mean 10, Group 2: mean 10
Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Over 10% participants missing overall. Insufficient variance data reported for analysis; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.; Group 2 Number missing: 9, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.

Protocol outcome 2: Physical functioning at longest follow up available
- Actual outcome for adults - severe: Karnofsky Performance Score (KPS) at 42 weeks; Group 1: mean 55, Group 2: mean 50

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Over 10% participants missing overall. Insufficient variance data for analysis; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.; Group 2 Number missing: 9, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.
- Actual outcome for adults - severe: Activities of Daily Living (ADL) at 42 weeks; Group 1: mean 72.4, Group 2: mean 69.4
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Over 10% participants missing overall. Insufficient variance data for analysis; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.; Group 2 Number missing: 9, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.

Protocol outcome 3: Adverse events at longest follow up available
- Actual outcome for adults - severe: Serious Adverse Events possibly or probably treatment-related at 42 weeks; Group 1: 1/117, Group 2: 2/117; Comments: Rintatolimod group: suicide attempt; placebo group: difficulty breathing/chest tightness, epilepsia partialis continua/seizures. There were 15 SAEs in rintatolimod group including non-treatment related: cerebral aneurysm, depression, anxiety, suicidal thoughts, upper respiratory tract infection, headache, suicide attempt, abscess, abdominal pain (x2), pulmonary embolism, uterine fibroids, cerebrovascular accident, parasthesia, abdominal pain with gastric distention. There were 7 SAEs in placebo group including non-treatment related: abdominal pain, difficulty breathing with chest tightness, accidental injury, epilepsia partialis continua, cholelithiasis, anxiety, depression
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Over 10% participants missing overall.; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 4: Exercise performance measure at longest follow up available
- Actual outcome for adults - severe: Treadmill exercise duration in seconds at 42 weeks; Group 1: mean 672 (SD 314.1); n=100, Group 2: mean 616 (SD 286.7); n=108; Comments: Baselines, mean (SD): Rintatolimod 576 (257.5) Placebo 588 (234.4)
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Over 10% participants missing overall.; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.; Group 2 Number missing: 9, Reason: Dropout due to: CFS symptoms, lack of efficacy, adverse events, other medical reasons and other non-medical reasons.

Protocol outcomes not reported by the study	Mortality at longest follow up available; Fatigue at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	Sulheim 2014[91]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120 CFS patients (excluding 68 healthy controls))
Countries and setting	Conducted in Norway; Setting: The Department of Paediatrics at Oslo University Hospital
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 8 weeks + 22 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Patients were diagnosed with CFS according to NICE diagnostic criteria for CFS in children/adolescents.
Stratum	young people; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	In agreement with clinical guidelines, the study applied a broad case definition requiring 3 months of unexplained disabling, chronic/relapsing fatigue of new onset. The study did not require that patients meet any other accompanying symptom criteria.
Exclusion criteria	Referring units were required to confirm that the patient did not have any medical or psychiatric disorder that might explain the fatigue and that they had experienced no concurrent demanding life event.
Recruitment/selection of patients	All 20 hospital paediatric departments in Norway, as well as primary care paediatricians and general practitioners, were invited to refer patients with CFS aged 12 to 18 years consecutively to the Department of Paediatrics at Oslo University Hospital, a national referral center in Norway for young patients with CFS.
Age, gender and ethnicity	Age - Mean (SD): 15.4. Gender (M:F): 34 males, 86 females. Ethnicity: Scandinavian: 118 (98%) Non-Scandinavian: 2 (1.7%)
Further population details	

Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: sympathomimetic/central antihypertensive drugs - clonidine. Tablets containing 25µg of clonidine hydrochloride (Catapresan; Boehringer Ingelheim) were enclosed in orange opaque, demolition-resistant lactose capsules (Apoteket Produktion and Laboratorier). Clonidine lowers blood pressure dose dependently, possibly increasing the risk of adverse effects in patients with CFS who already experience orthostatic intolerance. Therefore, clonidine dosages were chosen to yield plasma concentrations within the lower range of what is considered clinically effective. Based on a previous pilot study, the following algorithm was used: (1) Patient weight greater than 35 kg: 2 capsules twice daily for 8 weeks (ie, clonidine, 50 µg twice daily, in the intervention group); and (2) Patient weight less than 35kg: 1 capsule twice daily for 8 weeks (ie, clonidine, 25 µg twice daily, in the intervention group). Therapy was initiated 1 week after the baseline investigational program. One-half of the dose was given during the first 3 days to minimize introductory adverse effects. After 8 weeks of the full dose, the dose was halved for 1 additional week to avoid rebound effects, after which treatment was discontinued. At therapy initiation, each patient was supplied with a defined number of capsules. The residual number at week 8 was counted, and an index of adherence was calculated. Clonidineplasma concentration was measured at weeks 3 and 8. Duration 8 weeks. Concurrent medication/care: Not stated. Indirectness: No indirectness (n=60) Intervention 2: placebo. Empty capsules were used as placebo comparators. Duration 8 weeks. Concurrent medication/care: Not stated. Indirectness: No indirectness
Funding	Other (This study was funded by Health South–East Hospital Trust, the University of Oslo,Oslo and Akershus University College of Applied Sciences, the Norwegian Competence Network of Paediatric Pharmacotherapy, Simon Fougner Hartmann's Family Foundation, and Eckbo's Family Foundation.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CLONIDINE versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for young people; severity mixed or unclear: Chalder Fatigue Questionnaire (CFQ) total sum score at 30 weeks; MD; 0.5 (95%CI -14.7 to 15.7); Chalder Fatigue Questionnaire scale not reported, High=Top is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason:

Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 2: Physical functioning at longest follow up available
- Actual outcome for young people; severity mixed or unclear: Fatigue Disability Index (FDI) total sum score at 30 weeks; Mean; 0.2 (95%CI -13.3 to 13.6); Fatigue Disability Index scale not reported, High=Top is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 3: Cognitive function at longest follow up available
- Actual outcome for young people; severity mixed or unclear: Digit span backward test total at 30 weeks; Mean; -0.5 (95%CI -1.2 to 0.1);
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 4: Pain at longest follow up available
- Actual outcome for young people; severity mixed or unclear: BPI average pain score at 30 weeks; Mean; 0.4 (95%CI -0.4 to 1.1); Brief pain inventory 0-10, High=Top is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 5: Sleep quality at longest follow up available
- Actual outcome for young people; severity mixed or unclear: KSQ insomnia score at 30 weeks; Mean; 0.1 (95%CI -0.3 to 0.4); Karolinska sleep questionnaire insomnia scale not reported, High=Top is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 6: Adverse events at longest follow up available
- Actual outcome for young people; severity mixed or unclear: Adverse effects, self-reported at 9 weeks; Group 1: 43/57, Group 2: 33/51; Comments: Events are total number of participants who experience one or more of the following adverse effects: drowsiness, dry mouth, unwellness, constipation, sleepiness, loose stool, rash,

itching, sadness, headache, breast development in men, dry nasal mucus membranes, hallucinations, confusion, nightmares, "tingling" in extremities, pain in the extremities, dizziness when rising, blurred vision, dry eyes, oedema in the extremities, and 'other'.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: Serious indirectness, Comments: Some adverse effects are poorly defined, e.g. "unwellness" and "other"; Baseline details: Unclear - most reported baseline characteristic tables compare CFS patients and healthy controls at baseline, rather than between groups. Groups were stratified according to duration of CFS (>18< months) before randomisation.; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 7: Activity levels at longest follow up available
- Actual outcome for young people; severity mixed or unclear: Steps per day measured by accelerometer at 30 weeks; Mean; 119 (95%CI -796 to 1035);
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcome 8: Symptom scales at longest available follow up
- Actual outcome for young people; severity mixed or unclear: CFS symptom inventory hypersensitivity score at 30 weeks; Mean; -0.03 (95%CI -0.4 to 0.3); CFS symptom inventory hypersensitivity score scale not reported, High=Top is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 3, Reason: 5 dropouts reported, but unclear which 3 dropouts were not interviewed regarding adverse effects. Reasons for dropout: 2 headache, 1 syncope, 1 declined to take capsules, 1 declined participation.; Group 2 Number missing: 9, Reason: Reasons for dropout: 5 declined participation, 1 declined to take capsules, 1 suspected suicidality, 1 abdominal discomfort, 1 incurrent illness.

Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Psychological status at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**The 2010[93]**
Study type	RCT (Patient randomised; Parallel)

Number of studies (number of participants)	1 (n=67)
Countries and setting	Conducted in Netherlands; Setting: Patients referred to CFS specialist clinic at a University Hospital
Line of therapy	Not applicable
Duration of study	Intervention time: 10 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1994 CDC
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	CDC 1994 criteria; scoring above clinical cut-off on the CIS fatigue sub-scale and SIP 8
Exclusion criteria	Current psychiatric morbidity; pregnancy/lactating; lactose intolerance; psychotropic drug use; experimental medications
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 34.7 to 35.8. Gender (M:F): 20:47. Ethnicity: unclear
Further population details	-
Extra comments	Ondansetron/placebo: CIS fatigue 49.4/50; SIP 8 1375/1359; CDC symptoms 7.4/6.8
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=33) Intervention 1: 5-HT3 receptor antagonists - Ondansetron. Ondansetron (8 mg tablets). 16mg per day in 2 doses. Duration 10 weeks. Concurrent medication/care: None. Indirectness: No indirectness (n=34) Intervention 2: placebo. identical placebo - 2 tablets taken per day. Duration 10 weeks. Concurrent

	medication/care: None. Indirectness: No indirectness
Funding	Study funded by industry (GlaxoSmithKline)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ONDANSETRON versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: CIS fatigue at 12 weeks; Group 1: mean 44 (SD 11.1); n=33, Group 2: mean 45.4 (SD 11.5); n=34; Checklist individual strength fatigue subscale 8-56, High=Top is poor outcome; Comments: 49.4/50 at baseline - so may partially explain follow up direction of effect. ANCOVA (adjusting for baseline) p=0.73, which concurs with this.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: See results section for values at baseline. Because of small numbers, random mixing probably inadequate to allow good baseline comparability, despite good methodology in other respects; Group 1 Number missing: 2 (but imputed so included in analysis), Reason: increased feeling of malaise; Group 2 Number missing:

Protocol outcome 2: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Complaints of constipation at 12 weeks; Group 1: 3/33, Group 2: 1/34
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: See results section for values at baseline. Because of small numbers, random mixing probably inadequate to allow good baseline comparability, despite good methodology in other respects; Group 1 Number missing: 2 (but imputed so included in analysis), Reason: increased feeling of malaise; Group 2 Number missing:
- Actual outcome for adults; severity mixed or unclear: Increased feeling of malaise at 12 weeks; Group 1: 2/33, Group 2: 0/34
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: See results section for values at baseline. Because of small numbers, random mixing probably inadequate to allow good baseline comparability, despite good methodology in other respects; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 3: Activity levels at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Actometer (objective accelerometer-based method of measuring activity) at 12 weeks; Group 1: mean 55 (SD 15.5); n=33, Group 2: mean 60.6 (SD 17.9); n=34; Comments: 54.1/58.4 at baseline - so explains follow up result. ANCOVA p=0.9
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Baseline details: See results section for values at baseline. Because of small numbers, random mixing probably inadequate to allow good baseline comparability, despite good methodology in other respects; Group 1 Number missing: 2 (but imputed so included in analysis), Reason: increased feeling of malaise; Group 2 Number missing:

Protocol outcome 4: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Sickness Impact Profile (SIP) 8 at 12 weeks; Group 1: mean 1063 (SD 525.5); n=33, Group 2: mean 1172 (SD 694.6); n=34; Sickness Impact Profile (SIP) 8 0-5799, High=Top is poor outcome; Comments: 1375/1359 at baseline - so does not explain follow up benefit for ondansetron. ANCOVA p =0.3, however Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: See results section for values at baseline. Because of small numbers, random mixing probably inadequate to allow good baseline comparability, despite good methodology in other respects; Group 1 Number missing: 2 (but imputed so included in analysis), Reason: increased feeling of malaise; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	**Vercoulen 1996[97]**
Study type	Systematic Review
Number of studies (number of participants)	1 (n=107)
Countries and setting	Conducted in Netherlands; Setting: Randomly selected from CFS database to outpatient clinic of a university hospital in netherlands.
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 14 weeks (6 weeks treatment + 8 weeks follow up)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: London (Sharpe) criteria
Stratum	adults; severity mixed or unclear

Subgroup analysis within study	Not applicable
Inclusion criteria	London criteria for CFS; fatigue > 1year; CIS fatigue score 35 or more; Depressed patients had to have BDI of 16 or more; non-depressed patients BDI of 10 or less
Exclusion criteria	Any alternative illness that could explain symptoms; psychiatric diagnosis besides major depressive disorder in depressed patients; any psychiatric disorder in non-depressed patients; pregnancy or lactation; lack of contraception in women of childbearing age; previous exposure to fluoxetine in a trial; previous lack of response to fluoxetine; recent trial participation; any prescribed medications except incidental analgesics that could not be stopped; current psychotherapy
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 37.8 to 29.9. Gender (M:F): 23:73. Ethnicity: unclear
Further population details	-
Extra comments	Median (range) duration of CFS: 5.5 (1-30); Married/cohabiting: 68/96; Currently working 17/96. Paper stratified to depressed and non-depressed patients. However the results relevant to this review have been given in a form that covers both strata.
Indirectness of population	Serious indirectness: Oxford criteria used; PEM is not a compulsory feature
Interventions	(n=54) Intervention 1: antidepressants - selective serotonin reuptake inhibitors. Fluoxetine 20mg once daily. Duration 8 weeks. Concurrent medication/care: None. Indirectness: No indirectness (n=53) Intervention 2: placebo. Placebo given once daily. Duration 8 weeks. Concurrent medication/care: None. Indirectness: No indirectness
Funding	Study funded by industry (Eli Lilly, Netherlands)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS versus PLACEBO	

Protocol outcome 1: Fatigue at longest follow up available - Actual outcome for adults; severity mixed or unclear: CIS fatigue at 8 weeks; MD -0.164 (95% CI -0.64 - 0.31); Checklist individual strength fatigue subscale 8-56, High=Top is poor outcome; Fluoxetine change from baseline - placebo change from baseline; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events Protocol outcome 2: Psychological status at longest follow up available - Actual outcome for adults; severity mixed or unclear: Beck depression inventory at 8 weeks; MD -0.186 (95% CI -0.35 - -0.02); Beck depression inventory 0-63, High=Top is poor outcome; Fluoxetine change from baseline - placebo change from baseline; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events Protocol outcome 3: Adverse events at longest follow up available - Actual outcome for adults; severity mixed or unclear: adverse events - tremor at 14 weeks; Group 1: 18/45, Group 2: 13/51; Placebo results reported as 30 (26%) in paper, which must be a typo, as 30/51 would be 60%, which seems unlikely in the placebo group. Therefore the 26% has been taken as the more likely figure, which yields a numerator of 13. But the possibility of an error here should be realised; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events - Actual outcome for adults; severity mixed or unclear: adverse events - perspiration at 14 weeks; Group 1: 30/45, Group 2: 20/51; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events Protocol outcome 4: Symptom scales at longest follow up available - Actual outcome for adults; severity mixed or unclear: worsening of symptoms at 14 weeks; Group 1: 7/45, Group 2: 12/51; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events - Actual outcome for adults; severity mixed or unclear: improvement in symptoms at 14 weeks; Group 1: 8/45, Group 2: 5/51; Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 9, Reason: adverse events; Group 2 Number missing: 2, Reason: adverse events	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	Vollmer-conna 1997[99]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=99)
Countries and setting	Conducted in Australia; Setting: 2 collaborating centres in Australia - 1 in Sydney and 1 in Australia
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Inadequate method of assessment/diagnosis: Criteria not part of the group of criteria recognised by our review as 'consensus based' - Schluederberg criteria
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of CFS based on medical history, a through physical examination, and laboratory assessment.
Exclusion criteria	Pregnancy; on NASAIDs, steroids, immunomodulatory agents, choline esterase inhibitors; previously received immunologic therapy; recent history of asthma
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Range of means: 38-41. Gender (M:F): 24:75. Ethnicity: not reported
Further population details	-
Extra comments	ImG low/ImG med/ImG high/placebo: disease duration (yrs) 6/7/5/7; Immune cells CD4+ (billions/L)0.99/0.98/0.77/0.96; CD8+ (billions/L) 0.65/0.55/0.52/0.57; DTH response 14/9/13/11; Karnovsky score

	73/70/67/71; QAL score 477/522/481/396; POMS depression 16.8/11.3/18.6/15.9; POMS confusion 9.4/5.7/9.6/9.3; POMS fatigue 20.1/17.7/16/21.3; POMS energy -13/-9.3/-7.3/-16; non sedentary activity (h/day) 5/5/5/5
Indirectness of population	Serious indirectness: unclear criteria used. Schluederberg 1992 publication was not included in the diagnostic criteria review as it presented a review of the CDC 1988 criteria rather than an original set of criteria.
Interventions	(n=22) Intervention 1: immunomodulatory drugs - IV immunoglobulin G. IV immunoglobulin 0.5 g/kg. Immunoglobulin was Intragram. 3 infusions each lasting 24 hours at monthly intervals. Duration 3 months. Concurrent medication/care: None. Indirectness: No indirectness (n=28) Intervention 2: immunomodulatory drugs - IV immunoglobulin G. IV immunoglobulin 1 g/kg. Immunoglobulin was Intragram. 3 infusions each lasting 24 hours at monthly intervals. Duration 3 months. Concurrent medication/care: None. Indirectness: No indirectness (n=23) Intervention 3: immunomodulatory drugs - IV immunoglobulin G. IV immunoglobulin 2g/kg. Immunoglobulin was Intragram. 3 infusions each lasting 24 hours at monthly intervals. Duration 3 months. Concurrent medication/care: None. Indirectness: No indirectness (n=26) Intervention 4: placebo. Identical placebo solution given IV. 3 infusions each lasting 24 hours at monthly intervals. Duration 3 months. Concurrent medication/care: None. Indirectness: No indirectness
Funding	Study funded by industry (Commonwealth Serum laboratories Also CFS society of NSW)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IV IMMUNOGLOBULIN G versus PLACEBO

Protocol outcome 1: Quality of life at longest follow up available
- Actual outcome for adults; severity mixed or unclear: QAL, POMS depression, POMS confusion, POMS fatigue, POMS energy at 6 months; All these outcomes : no data provided.
QAL: NS between-group effect (p>0.13) for the 3 different Ig doses versus placebo
POMS energy: NS between-group effect (p>0.75) for the 3 different Ig doses versus placebo
Other POMS outcomes (depression, confusion, fatigue): NS
Hours in non-sedentary activity: NS

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 2: Physical functioning at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Karnofsky scale at 6 months; Median (IQR)
Low dose 0.5 g Ig 80 (70-80) placebo 77.5 (70-80) [NS]
Baseline discrepancy: 77.5/70 which partially explains result;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for adults; severity mixed or unclear: Karnofsky scale at 6 months; Median (IQR)
Medium dose 1 g Ig 80 (70-80) placebo 77.5 (70-80) [NS]
Baseline discrepancy: 70/70 ;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for adults; severity mixed or unclear: Karnofsky scale at at 6 months; Median (IQR)
Low dose 2 g Ig 75 (70-80) placebo 77.5 (70-80) [NS]
Baseline discrepancy: 70/70 ;
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 3: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Adverse events (high dose) at 3 months; Group 1: 18/23, Group 2: 23/26; Comments: Constitutional symptoms including headaches, worsened fatigue, malaise, and concentration impairment, typically developing 12 to 24 hours after the completion of an infusion and persisting up to 10 days.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Adverse events (medium dose) at 3 months; Group 1: 20/28, Group 2: 23/26; Comments: Constitutional symptoms including headaches, worsened fatigue, malaise, and concentration impairment, typically developing 12 to 24 hours after the completion of an infusion and persisting up to 10 days.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for adults; severity mixed or unclear: Adverse events (low dose) at 3 months; Group 1: 18/22, Group 2: 23/26; Comments: Constitutional symptoms including headaches, worsened fatigue, malaise, and concentration impairment, typically developing 12 to 24 hours after the completion of an infusion and persisting up to 10 days.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at longest follow up available; Fatigue at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Wearden 1998[103]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=136)
Countries and setting	Conducted in United Kingdom; Setting: UK-based hospital.
Line of therapy	Not applicable
Duration of study	Intervention time: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: All participants met the 'Oxford' CFS diagnosis criteria (Sharpe et al 1991).
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	All subjects met operationalised 'Oxford' research criteria (Sharpe et al, 1991) for CFS: (a) a principal complaint of fatigue of at least six months' duration, exacerbated by physical activity (and usually mental activity); (b) impairment in three out of four areas of activity (activities of daily living, occupational, social or leisure activities); (c) no medical cause of fatigue.

Exclusion criteria	Subjects with schizophrenia, bipolar disorder, eating disorder, alcohol or illicit drug misuse, those with current suicidal ideation, a history of ischaemic heart disease or an inability to read and write English were excluded. Pre-menopausal women were required to take precautions against pregnancy during the trial. Subjects taking anti-depressant medication were required to stop and undergo at least a two week washout period before entering the trial. Antidepressants were not withdrawn (and patients were excluded from the trial) if patients were judged to have any significant suicidal risk.
Recruitment/selection of patients	Patients over 18 years of age were recruited from consecutive referrals to a university department of medicine out-patient clinic drawing from across north-west England and north Wales between June 1993 and March 1995.
Age, gender and ethnicity	Age - Mean (SD): 38.7 (10.8) years. Gender (M:F): 97 female, 39 male. Ethnicity: Not stated.
Further population details	-
Extra comments	Four armed trial, including exercise comparison: (1) exercise and fluoxetine, n=33; (2) exercise and placebo drug, n=34; (3) exercise placebo and fluoxetine, n=35; (4) exercise placebo and placebo drug, n=34.
Indirectness of population	Serious indirectness: Oxford criteria used; PEM is not a compulsory feature
Interventions	(n=35) Intervention 1: antidepressants - selective serotonin reuptake inhibitors. Fluoxetine at a fixed daily dose of 20 mg. Plus exercise placebo. Duration 6 months. Concurrent medication/care: The fluoxetine treatment was accompanied by a placebo exercise programme in which a participant activity diaries were reviewed by a physiotherapist. The same was done in the placebo drug group. Subjects were not offered any specific advice on how much exercise they should be taking, but were told to do what they could when they felt capable and to rest when they felt they needed to. Indirectness: No indirectness (n=34) Intervention 2: placebo. Placebo drug and placebo exercise. The placebo to fluoxetine was a capsule of similar taste and appearance, taken daily. Participants in the placebo drug group also received the same exercise placebo as the fluoxetine group. Duration 6 months. Concurrent medication/care: Not stated. Indirectness: No indirectness (n=34) Intervention 3: Graded exercise programme. Subjects were instructed to carry out their preferred aerobic activity (usually walking/jogging, swimming or cycling), for 20 minutes, at least three times per week. The intensity of the activity was initially set at a level which utilised oxygen at approximately 75% of the subject's tested functional

	maximum. Subjects monitored their prescribed exercise programmes on a chart along with pre- and post-exercise heart rates and perceived exertion. Exercise intensity was increased when there was a consistent recorded reduction of 10 beats per minute in post-exercise heart rate for one week and two points on the perceived exertion scale (about three times in six months in an adherent patient). Subjects adhered to the exercise programme if their charts showed that they had performed the required activity, at the required intensity, at least three times per week. Plus placebo drug. Duration 6 months. Concurrent medication/care: Participants in this group also received placebo drug of a capsule of similar taste and appearance to fluoxetine, taken daily. Indirectness: No indirectness (n=33) Intervention 4: Fluoxetine and graded exercise programme, combined intervention. Fluoxetine at a fixed daily dose of 20 mg. Graded exercise intervention: Subjects were instructed to carry out their preferred aerobic activity (usually walking/jogging, swimming or cycling), for 20 minutes, at least three times per week. The intensity of the activity was initially set at a level which utilised oxygen at approximately 75% of the subject's tested functional maximum. Subjects monitored their prescribed exercise programmes on a chart along with pre- and post-exercise heart rates and perceived exertion. Exercise intensity was increased when there was a consistent recorded reduction of 10 beats per minute in post-exercise heart rate for one week and two points on the perceived exertion scale (about three times in six months in an adherent patient). Subjects adhered to the exercise programme if their charts showed that they had performed the required activity, at the required intensity, at least three times per week. Plus placebo drug. Duration 6 months. Concurrent medication/care: Participants in this group received both exercise intervention and fluoxetine. Other background treatment not stated. Indirectness: No indirectness
Funding	Other (The study was funded by a grant from the Linbury Trust.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs plac. Fatigue on the 14-item Chalder fatigue scale at 26 weeks; Group 1: mean -3 (SD 8.15); n=35, Group 2: mean -2.7 (SD 7.77); n=34; Chalder fatigue scale not reported, High=Top is poor outcome; Comments: SDs calculated from reported CI 95% ranges
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 2: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs plac. Depression on the Hospital Anxiety and Depression Scale (HADS) depression subscale at 26 weeks; Group 1: mean -1.7 (SD 3.78); n=35, Group 2: mean -1.3 (SD 2.87); n=34; HADS depression 0-21, High=Top is poor outcome; Comments: SDs calculated from reported CI 95% ranges
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover -

Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 3: Exercise performance measure at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs plac. Functional work capacity (VO2 max, mL O2 per kg per minute; change score) at 26 weeks; Group 1: mean 1 mL O2 kg-1 min-1 (SD 5.8222); n=35, Group 2: mean -0.1 mL O2 kg-1 min-1 (SD 4.8722); n=34
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS versus GRADED EXERCISE PROGRAMME

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs exer. Fatigue on the 14-item Chalder fatigue scale at 26 weeks; Group 1: mean -3 (SD 8.15); n=35, Group 2: mean -5.7 (SD 10.9); n=34; Chalder fatigue scale not reported, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 2: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs exer. Depression on the Hospital Anxiety and Depression Scale (HADS) depression subscale at 26 weeks; Group 1: mean -1.7 (SD 4.01); n=35, Group 2: mean -1.2 (SD 3.49); n=34; HADS depression 0-21, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 3: Exercise performance measure at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs exer. Functional work capacity (VO2 max, mL O2 per kg per minute; change score) at 26 weeks; Group 1: mean 1 mL O2 kg-1 min-1 (SD 5.8222); n=35, Group 2: mean 2.8 mL O2 kg-1 min-1 (SD 5.732); n=34
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS versus SELECTIVE SEROTONIN REUPTAKE INHIBITORS AND GRADED EXERCISE PROGRAMME COMBINED

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs combo. Fatigue on the 14-item Chalder fatigue scale at 26 weeks; Group 1: mean -3 (SD 8.15); n=35, Group 2: mean -6 (SD 10.43); n=33; Chalder fatigue scale not reported, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 2: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs combo. Depression on the Hospital Anxiety and Depression Scale (HADS) depression subscale at 26 weeks; Group 1: mean -1.7 (SD 4.01); n=35, Group 2: mean -2 (SD 3.67); n=34; HADS depression 0-21, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 3: Exercise performance measure at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Flu vs combo. Functional work capacity (VO2 max, mL O2 per kg per minute; change score) at 26 weeks; Group 1: mean 1 mL O2 kg-1 min-1 (SD 5.8222); n=35, Group 2: mean 2 mL O2 kg-1 min-1 (SD 4.2303); n=33
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS AND GRADED EXERCISE PROGRAMME COMBINED versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Combo vs plac. Fatigue on the 14-item Chalder fatigue scale at 26 weeks; Group 1: mean -6 (SD 10.43); n=33, Group 2: mean -2.7 (SD 7.77); n=34; Chalder fatigue scale not reported, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 2: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Combo vs plac. Depression on the Hospital Anxiety and Depression Scale (HADS) depression subscale at 26 weeks; Group 1: mean -2 (SD 3.67); n=33, Group 2: mean -1.3 (SD 2.87); n=34; HADS depression 0-21, High=Top is poor outcome;
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

Protocol outcome 3: Exercise performance measure at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Combo vs plac. Functional work capacity (VO2 max, mL O2 per kg per minute; change score) at 26 weeks; Group 1: mean 2 mL O2 kg-1 min-1 (SD 4.2303); n=33, Group 2: mean -0.1 mL O2 kg-1 min-1 (SD 4.8722); n=34
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SELECTIVE SEROTONIN REUPTAKE INHIBITORS AND GRADED EXERCISE PROGRAMME COMBINED versus GRADED EXERCISE PROGRAMME

Protocol 1: Fatigue at longest follow up available - Actual outcome for adults; severity mixed or unclear: Combo vs exer. Fatigue on the 14-item Chalder fatigue scale at 26 weeks; Group 1: mean -6 (SD 10.43); n=33, Group 2: mean -5.7 (SD 10.9); n=34; Chalder fatigue scale not reported, High=Top is poor outcome; Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5 Protocol outcome 2: Psychological status at longest follow up available - Actual outcome for adults; severity mixed or unclear: Combo vs exer. Depression on the Hospital Anxiety and Depression Scale (HADS) depression subscale at 26 weeks; Group 1: mean -2 (SD 3.67); n=33, Group 2: mean -1.2 (SD 3.49); n=34; HADS depression 0-21, High=Top is poor outcome; Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5 Protocol outcome 3: Exercise performance measure at longest follow up available - Actual outcome for adults; severity mixed or unclear: Combo vs exer. Functional work capacity (VO2 max, mL O2 per kg per minute; change score) at 26 weeks; Group 1: mean 2 mL O2 kg-1 min-1 (SD 4.2303); n=33, Group 2: mean 2.8 mL O2 kg-1 min-1 (SD 5.732); n=34 Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 10; Group 2 Number missing: 5	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Pain at longest follow up available; Sleep quality at longest follow up available; Adverse events at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available ; Symptom scales at longest available follow up

Study	**Young 2013[107]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)

Countries and setting	Conducted in USA; Setting: Suburban research and treatment centre in USA
Line of therapy	Not applicable
Duration of study	Intervention time: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fukuda
Stratum	adults; severity mixed or unclear
Subgroup analysis within study	Not applicable
Inclusion criteria	CFS as diagnosed by Fukuda criteria, plus medical history, clinical interview, brief clinical examination and responses to CF checklist; aged 18-60; Global executive composite score that was 1.5 sds above standardised population mean; able to swallow medication; ability to communicate; capacity to fully comply with procedures and restrictions.
Exclusion criteria	Psychostimulant in past 6 months; positive test for pregnancy; not using accepted forms of contraception during the study; breastfeeding; severe comorbid psychiatric diagnoses; history of psychosis; pervasive medical disorders, severe Axis II disorders; severe substance dependence; chronic/acute medical condition that could be affected by study medication: hypothyroidism, hypertension; fibromyalgia therapy; weight <30 kg or >120kg
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (range): 41 (21-59). Gender (M:F): 1:25. Ethnicity: unclear
Further population details	-
Extra comments	Aged 21-59 with CFS and cognitive complaints
Indirectness of population	Serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature
Interventions	(n=15) Intervention 1: Amphetamines - Lisdexamphetamine. Lisdexamfetamine given as a flexible morning dose (progressing from 30, through 50, and then to 70 mg/day) provided no serious AEs occurred. Duration 6 weeks. Concurrent medication/care: None. Indirectness: No indirectness

	(n=15) Intervention 2: placebo. Identical placebo given in same incremental doses. Duration 6 weeks. Concurrent medication/care: None. Indirectness: No indirectness
Funding	Study funded by industry (Shire)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CISDEXAMPHETAMINE versus PLACEBO

Protocol outcome 1: Fatigue at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Fatigue severity Scale at 6 weeks; Group 1: mean 20.92 'mean improvement' (SD 14.71); n=13, Group 2: mean 5 'mean improvement' (SD 11.73); n=11; Fatigue severity scale 9-63, High=Top is poor outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high.

Protocol outcome 2: Cognitive function at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Behaviour rating Inventory of Executive Function (BRIEF): Global executive composite at 6 weeks; Group 1: mean 21.38 (SD 15.85); n=13, Group 2: mean 3.36 (SD 7.26); n=11; Behaviour rating Inventory of Executive Function scale not reported, High=Top is poor outcome; Comments: results reported are 'improvements'
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high.

Protocol outcome 3: Psychological status at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Hamilton anxiety scale at 6 weeks; Group 1: mean 11.31 'mean improvement' - any positive value represents an improvement (SD 9.74); n=13, Group 2: mean 6.18 'mean improvement' - any positive value represents an improvement (SD 8.28); n=11
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high.

Protocol outcome 4: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: McGill pain questionnaire at 6 weeks; Group 1: mean 10.38 'mean improvement' (SD 8.84); n=13, Group 2: mean 2.54 'mean improvement' (SD 9.53); n=11
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;

	Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high. Protocol outcome 5: Adverse events at longest follow up available - Actual outcome for adults; severity mixed or unclear: Adverse events: headache at 6 weeks; Group 1: 2/15, Group 2: 1/11 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high. - Actual outcome for adults; severity mixed or unclear: Adverse events: dry mouth at 6 weeks; Group 1: 1/15, Group 2: 0/11 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high. - Actual outcome for adults; severity mixed or unclear: Adverse events: insomnia at 6 weeks; Group 1: 1/15, Group 2: 0/11 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high. - Actual outcome for adults; severity mixed or unclear: Discontinuation due to adverse events at 6 weeks; Group 1: 2/15, Group 2: 0/11; Comments: Adverse events: insomnia at visit 3, anxiety at visit 5 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high. Protocol outcome 6: Symptom scales at longest available follow up - Actual outcome for adults; severity mixed or unclear: Clinical Global Improvement - severity at 6 weeks; Group 1: mean 1.92 'mean improvement' (SD 1.5); n=13, Group 2: mean 0.64 'mean improvement' (SD 0.92); n=11 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Due to adverse events; Group 2 Number missing: 4, Reason: These were screen failures and so exclusion not related to outcome. Therefore not counted. Thus differential between groups is 13.33%, which is high.
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Physical functioning at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Study	Zachrisson 2002[109]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	(n=100)
Countries and setting	Conducted in Sweden; Setting: The study was conducted in a special unit at a single hospital
Line of therapy	Unclear
Duration of study	Intervention + follow up: 32 weeks (24 weeks intervention + final follow-up at 32 weeks)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Participants met 1994 CDC criteria for CFS diagnosis (as well as ACR criteria for FM). Investigations prior to study entry included physical examination, vital signs and blood parameters.
Stratum	adults; severity mixed or unclear: Age 18-65 years; severity mixed or unclear (according to global assessment of illness severity measured at baseline 17% were moderately ill, 70% markedly ill, 12% severely ill, 1% most extremely ill)
Subgroup analysis within study	Not applicable: NA
Inclusion criteria	Female; age 18-6 years; met both ACR criteria for fibromyalgia and 1994 CDC criteria for CFS; functional impairment related to these syndromes as documented by ≥6 months of full- or part-time sick leave. Prescribed medications allowed to continue, as long as they were in a steady state.
Exclusion criteria	Pathological values of significance recorded from laboratory results; signs or symptoms of ongoing severe psychiatric or other somatic disorder (patients with a history of depressed mood and earlier treatment with antidepressants were included if the history did not include melancholia or psychotic features; autoimmune or rheumatological disorders.
Recruitment/selection of patients	Consecutive patients referred from primary care centres to special unit at a hospital in Sweden
Age, gender and ethnicity	Age - Mean (range): staph toxoid 49 (26-65); placebo 47 (20-63). Gender (M:F): 0-100. Ethnicity: Not reported

Further population details	-
Extra comments	Staph toxoid/placebo, mean (range): Duration of symptoms (years) 11 (1-40) / 12 (1-36); BMI 27 (18-41) / 26 (18-41). % participants reporting the following features of illness: 47% low grade fevers, 57% sore throat, 87% prone to infections, 63% onset related to regional pain later becoming generalised, 16% onset related to acute infection, 15% onset related to pregnancy/delivery, 6% onset directly related to generalised pain - reported to be no significant differences between groups.
Indirectness of population	Very serious indirectness: 1994 CDC criteria used; PEM not a compulsory feature and all participants also met diagnostic criteria for fibromyalgia.
Interventions	(n=50) Intervention 1: staphylococcus vaccine. Staphylococcus toxoid preparation, Staphypan Berna (SB). Composed of undefined extracts of 2 strains of staphylococci (S. aureus and S. epidermidis), and a preservative compound thiomersal. Injection given subcutaneously in gluteal region by nurse. Drug administered in increasing doses of 0.1ml, 0.2ml, 0.3ml, 0.4ml, 0.6ml, 0.8ml, 0.9ml, and 1.0ml weekly, followed by booster doses of 1.0ml every 4 weeks with final dose given week 24. Drug kept in 1ml ampoules and packed in boxes marked with patient numbers. Both active substance and placebo caused slight local pain and reaction after injection. . Duration 24 weeks. Concurrent medication/care: Patients were allowed to continue with prescribed medication during the study, as long as they were in a steady state. Concomitant medications (reported for study population as a whole): 79% on antidepressants (low doses tricyclics or SSRIs), 42% on hypnotics, 21% on benzodiazepines, 19% on medication for GI problems.. Indirectness: No indirectness; Indirectness comment: NA (n=50) Intervention 2: placebo. Sterile water. Injection given subcutaneously in gluteal region by nurse. Administered in increasing doses of 0.1ml, 0.2ml, 0.3ml, 0.4ml, 0.6ml, 0.8ml, 0.9ml, and 1.0ml weekly, followed by booster doses of 1.0ml every 4 weeks with final dose given week 24. Drug kept in 1ml ampoules and packed in boxes marked with patient numbers. Similar in colour to active treatment. Both active substance and placebo caused slight local pain and reaction after injection. Duration 24 weeks. Concurrent medication/care: Patients were allowed to continue with prescribed medication during the study, as long as they were in a steady state. Concomitant medications (reported for study population as a whole): 79% on antidepressants (low doses tricyclics or SSRIs), 42% on hypnotics, 21% on benzodiazepines, 19% on medication for GI problems. Indirectness: No indirectness; Indirectness comment: NA
Funding	Equipment / drugs provided by industry (The study drug was provided by SSVI, Berne)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: STAPHYLOCOCCUS VACCINE versus PLACEBO

Protocol outcome 1: Pain at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Visual analogue scale of pain at 32 weeks; Group 1: mean 5.9 (SD 2.2); n=49, Group 2: mean 6.2 (SD 1.95); n=49; Visual analogue of pain Unclear Top=High is poor outcome; Comments: Baseline scores, mean (SD): staph toxoid 6.2 (1.69); placebo 6.2 (1.71).
Paper reports line used was 100mm long. Range likely to be 0-10 (cm) instead of 0-100 (mm) based on values reported.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Missing data - missing data (n=4 in each group) supplemented by carrying forward last rated scoring.; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Comparable for age, gender (all patients female), duration of symptoms, BMI. Baseline scores comparable. ; Blinding details: Patient reported outcome; Group 1 Number missing: 1, Reason: n=1 dropped out prior to any assessments (bereavement).; Group 2 Number missing: 1, Reason: n=1 dropped out prior to any assessments (diagnosed with hypertension which required treatment)

Protocol outcome 2: Adverse events at longest follow up available
- Actual outcome for adults; severity mixed or unclear: Most frequent adverse events at 26 weeks; Group 1: 28/50, Group 2: 26/50; Comments: Breakdown of common AEs (staph toxoid/placebo): Headaches 12/3; infections 5/11; skin disorders 4/5; GI problems 2/2; nausea/vomiting 3/1; depression 1/2; cardiovascular problems, palpitations 1/2. AEs were listed irrespective of causal relationship with study drug.
Does not include local reaction at site of injection (100% of participants experienced this, but severity was not rated to preserve blinding)
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Comparable for age, gender (all patients female), duration of symptoms, BMI, baseline scores for other outcomes. ; Blinding details: Nurses who assessed outcomes were different to nurses who administered treatment; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for adults; severity mixed or unclear: Clinical global assessment of side effects at 26 weeks; Mean; Patient rated measure of side effects 1-4 Top=High is poor outcome, Comments: Scale: 1=no side effects, 2=do not significantly interfere with functioning), 3=significantly interfere with functioning, 4=outweigh therapeutic benefit)
Number of patients who experienced side effects: staph toxoid (n=50) 13 (26%); placebo (n=50) 7 (14%)
Number of patients who gave each rating on the scale, staph toxoid (n=50)/placebo (n=50): 1, 37/43; 2, 5/2; 3, 2/0; 4, 6/5 ;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Comparable for age, gender (all patients female), duration of symptoms, BMI, baseline scores for other outcomes. ; Blinding details: Patient reported outcome; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 3: Symptom scales at longest available follow up
- Actual outcome for adults; severity mixed or unclear: Clinical global impression of change at 32 weeks; Group 1: mean 3.7 (SD 1.51); n=49, Group 2: mean 4.4 (SD 1.08); n=49; Clinical global impression of change 1-7 Top=High is poor outcome; Comments: Measured at end of study only

Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Missing data - missing data (n=4 in each group) supplemented by carrying forward last rated scoring. Patients who discontinued before week 26 or were not evaluated for other reasons received the worst score on the clinical global impression of change scale (as only measure at end of study); Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Comparable for age, gender (all patients female), duration of symptoms, BMI.; Blinding details: Observer reported outcome; Group 1 Number missing: 1, Reason: n=1 dropped out prior to any assessments (bereavement).; Group 2 Number missing: 1, Reason: n=1 dropped out prior to any assessments (diagnosed with hypertension which required treatment) - Actual outcome for adults; severity mixed or unclear: Clinical global impression of severity at 32 weeks; Group 1: mean 4.5 (SD 0.52); n=49, Group 2: mean 4.8 (SD 0.62); n=49; Clinical global impression of severity 1-7 Top=High is poor outcome; Comments: Baseline scores, mean (SD): staph toxoid 5.0 (0.63); placebo 5.0 (0.54) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Missing data - missing data (n=4 in each group) supplemented by carrying forward last rated scoring.; Indirectness of outcome: No indirectness, Comments: NA; Baseline details: Comparable for age, gender (all patients female), duration of symptoms, BMI. Baseline scores comparable. ; Blinding details: Observer reported outcome; Group 1 Number missing: 1, Reason: n=1 dropped out prior to any assessments (bereavement).; Group 2 Number missing: 1, Reason: n=1 dropped out prior to any assessments (diagnosed with hypertension which required treatment)	
Protocol outcomes not reported by the study	Quality of life at longest follow up available; Mortality at longest follow up available; Fatigue at longest follow up available; Physical functioning at longest follow up available; Cognitive function at longest follow up available; Psychological status at longest follow up available; Sleep quality at longest follow up available; Activity levels at longest follow up available; Return to school or work at longest follow up available; Exercise performance measure at longest follow up available; Care needs at longest follow up available ; Impact on families and carers at longest follow up available

Appendix E Forest plots

E.1 Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo

Figure 2: Quality of Life: SF36 (max % change from baseline) at 10 months (rituximab)

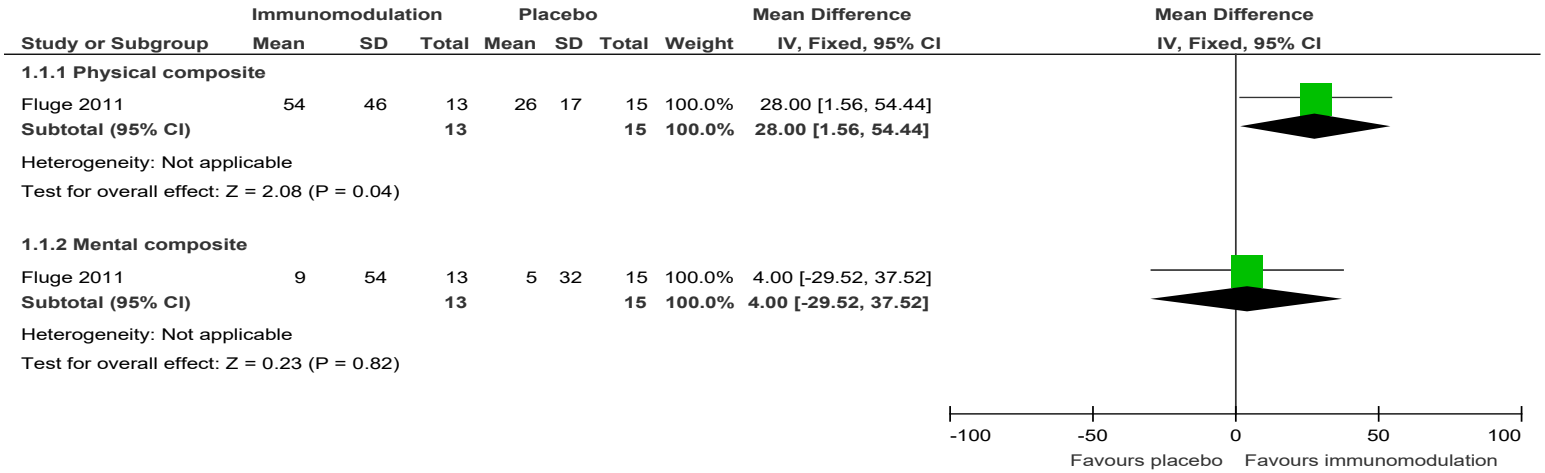


Figure 3: Fatigue/fatigability: Fatigue severity scale at 18 months (rituximab)

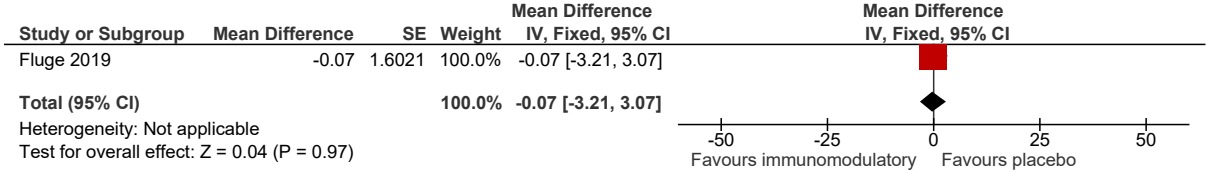


Figure 4: Fatigue/fatigability: Fatigue numeric rating scale 0-10 at 16-20 months (rituximab)

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
Fluge 2019	-0.06	0.2296	100.0%	-0.06 [-0.51, 0.39]
Total (95% CI)			**100.0%**	**-0.06 [-0.51, 0.39]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.26 (P = 0.79)

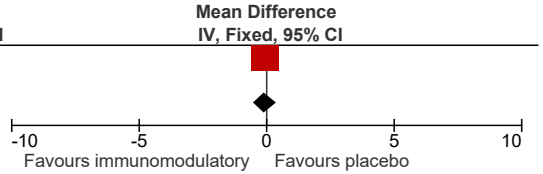


Figure 5: Psychological status: Hamilton Depression Scale at 6 months (IVIG)

Study or Subgroup	Immunomodulation Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lloyd 1990	9	5	23	10	3	26	100.0%	-1.00 [-3.35, 1.35]
Total (95% CI)			**23**			**26**	**100.0%**	**-1.00 [-3.35, 1.35]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.84 (P = 0.40)

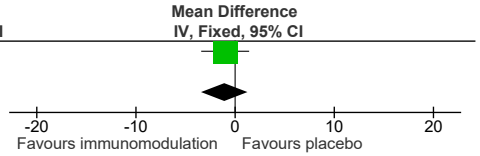


Figure 6: Psychological status: Zung Self-Rating Depression Scale at 6 months (IVIG)

Study or Subgroup	Immunomodulation Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lloyd 1990	41	11	23	40	12	26	100.0%	1.00 [-5.44, 7.44]
Total (95% CI)			**23**			**26**	**100.0%**	**1.00 [-5.44, 7.44]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.30 (P = 0.76)

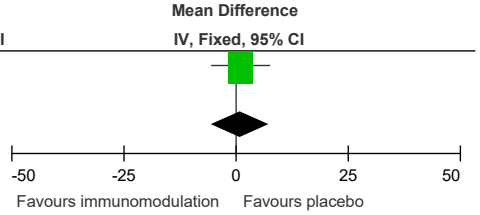


Figure 7: Psychological status: mental health on the Medical Outcome Study Short Form at 150 days (IVIG)

Study or Subgroup	Immunomodulation Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Peterson 1990	58.3	17.4	14	62.9	13.3	14	100.0%	-4.60 [-16.07, 6.87]
Total (95% CI)			**14**			**14**	**100.0%**	**-4.60 [-16.07, 6.87]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

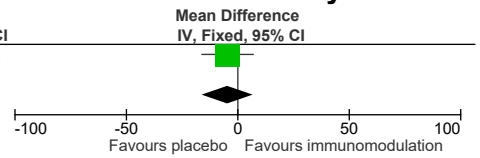


Figure 8: Physical functioning: physical functioning on the Medical Outcome Study Short Form at 150 days (IVIG)

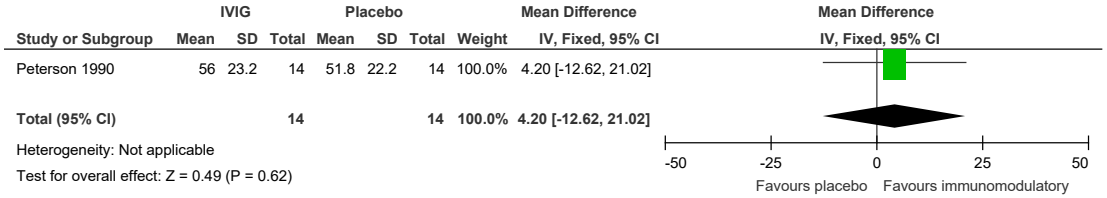


Figure 9: Physical functioning: physical functioning on the SF36 at 24 months (rituximab)

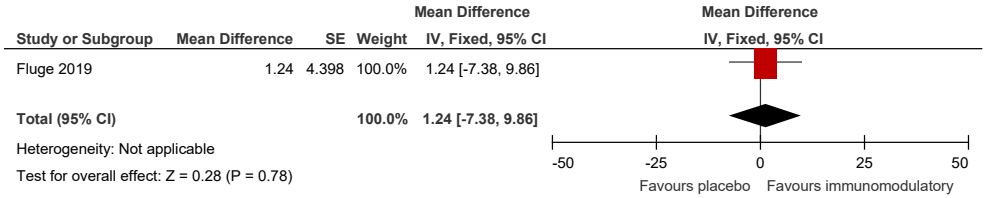


Figure 10: Physical functioning: functional level % at 16-20 months (rituximab)

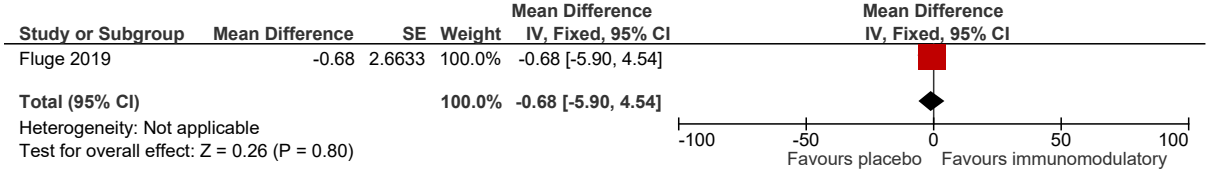


Figure 11: SAEs with possible/probable relation to intervention (rintatolimod)

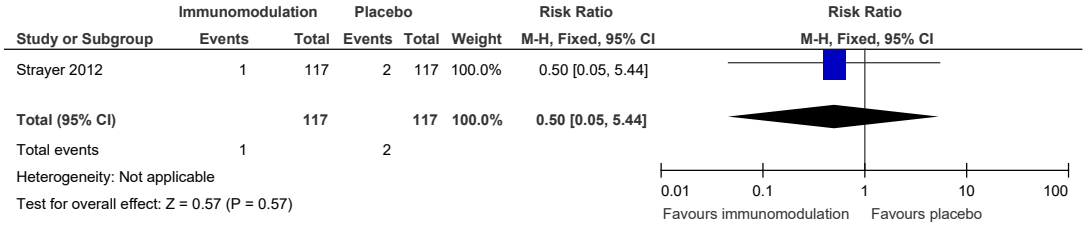


Figure 12: Adverse events: major adverse events (IVIG)

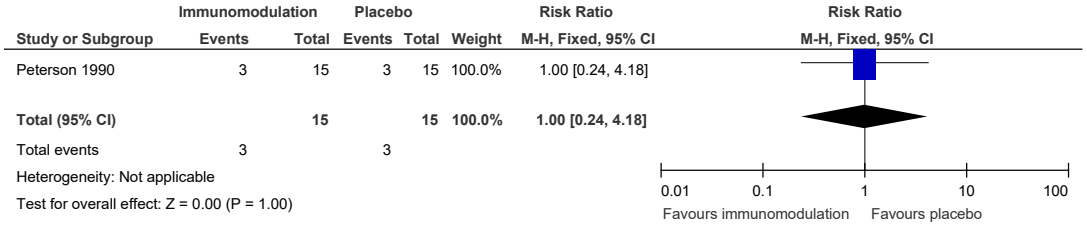


Figure 13: Adverse events: constitutional symptoms (IVIG)

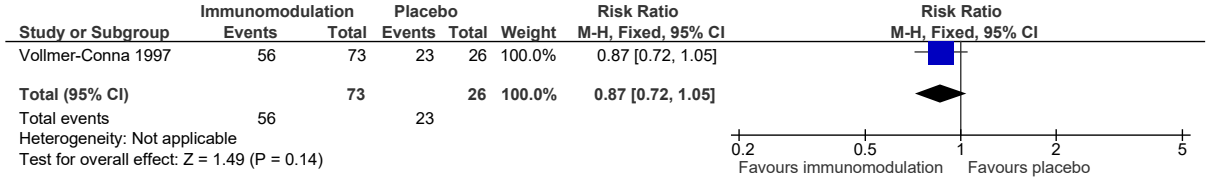


Figure 14: Adverse events: any SAEs with possible/probable relation to intervention (rituximab)

Study or Subgroup	Immunomodulatory Events	Total	Placebo Events	Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Fluge 2019	8	77	0	74	100.0%	7.82 [1.89, 32.35]
Total (95% CI)		**77**		**74**	**100.0%**	**7.82 [1.89, 32.35]**
Total events	8		0			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.84 (P = 0.005)

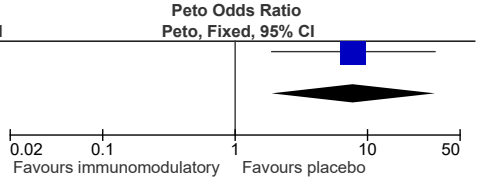


Figure 15: Adverse events: any AEs of at least moderate severity with possible/probable relation to intervention (rituximab)

Study or Subgroup	Immunomodulatory Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Fluge 2019	26	77	12	74	100.0%	2.08 [1.14, 3.81]
Total (95% CI)		**77**		**74**	**100.0%**	**2.08 [1.14, 3.81]**
Total events	26		12			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.38 (P = 0.02)

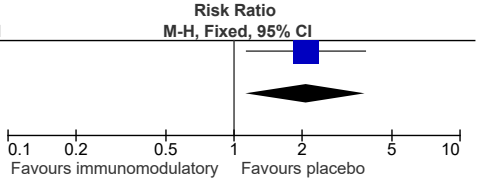


Figure 16: Adverse events: suspected unexpected adverse reactions (rituximab)

Study or Subgroup	Immunomodulatory Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Fluge 2019	2	77	1	74	100.0%	1.92 [0.18, 20.75]
Total (95% CI)		**77**		**74**	**100.0%**	**1.92 [0.18, 20.75]**
Total events	2		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.54 (P = 0.59)

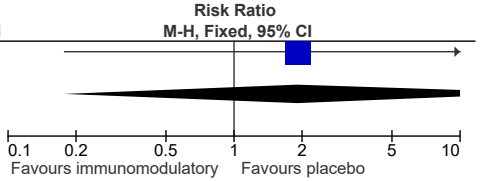


Figure 17: Activity levels: mean steps/24 hrs (rituximab)

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
Fluge 2019	-127	447.4572	100.0%	-127.00 [-1004.00, 750.00]
Total (95% CI)			**100.0%**	**-127.00 [-1004.00, 750.00]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.28 (P = 0.78)

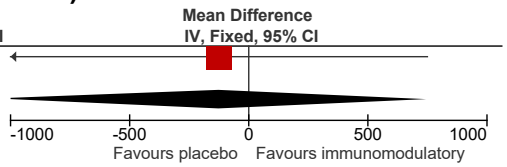


Figure 18: Exercise performance measure: Treadmill exercise duration in seconds at 42 weeks (rintatolimod)

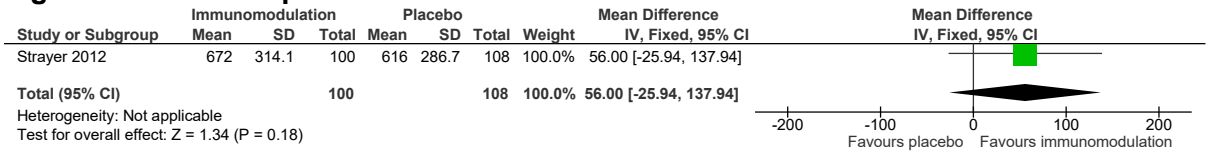


Figure 19: Return to school or work: Resumption of pre-morbid employment status (full-time) at 6 months (IVIG)

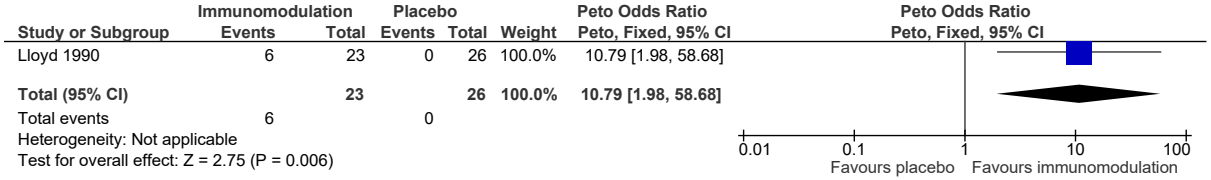
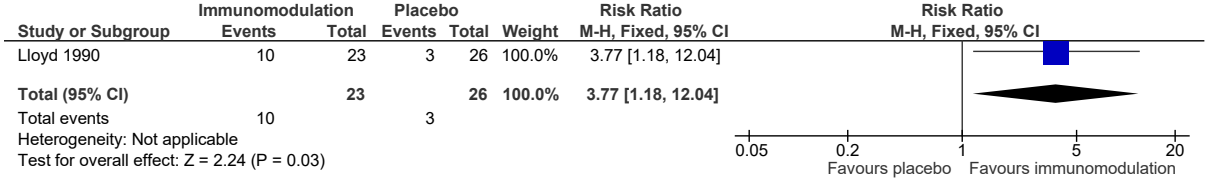


Figure 20: Symptom scales: Marked reduction in symptoms and improvement in functional capacity (IVIG)



E.2 Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo

Figure 21: Quality of life: SF-36 subscales at 12 weeks (duloxetine)

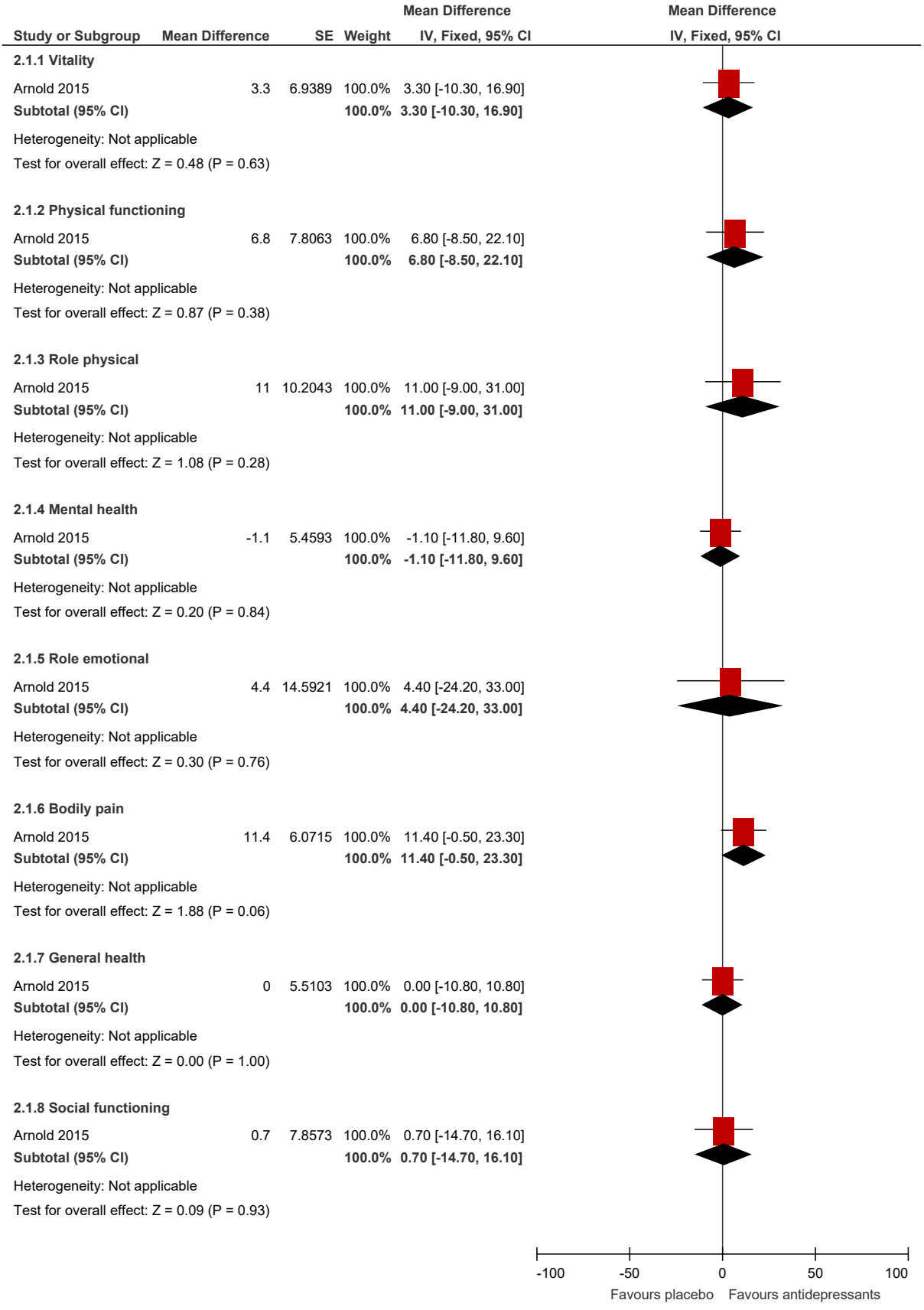


Figure 22: Fatigue: 14-item Chalder fatigue scale at 26 weeks (fluoxetine)

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-3	8.15	35	-2.7	7.77	34	100.0%	-0.30 [-4.06, 3.46]
Total (95% CI)			**35**			**34**	**100.0%**	**-0.30 [-4.06, 3.46]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.16 (P = 0.88)

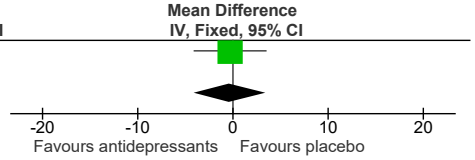


Figure 23: Fatigue: MFI-20 at 12 weeks (duloxetine)

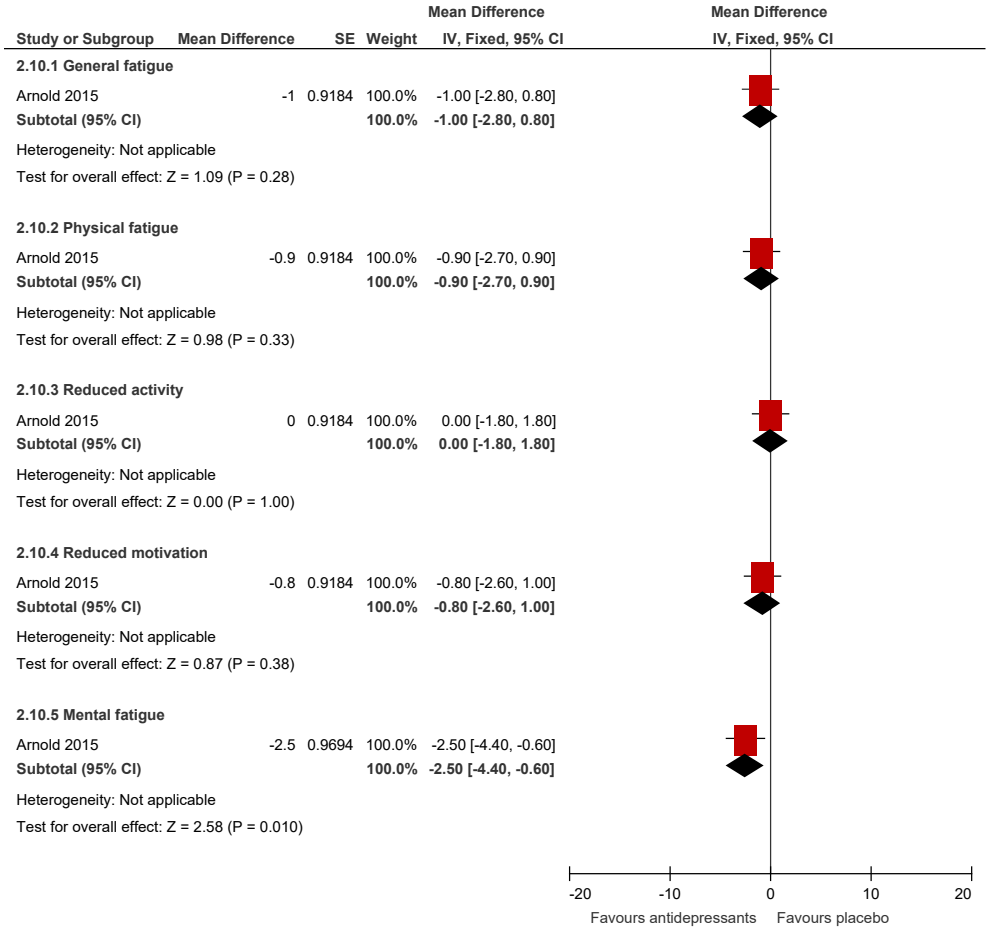


Figure 24: Fatigue: Checklist Individual Strength (CIS) fatigue at 16 weeks (fluoxetine)

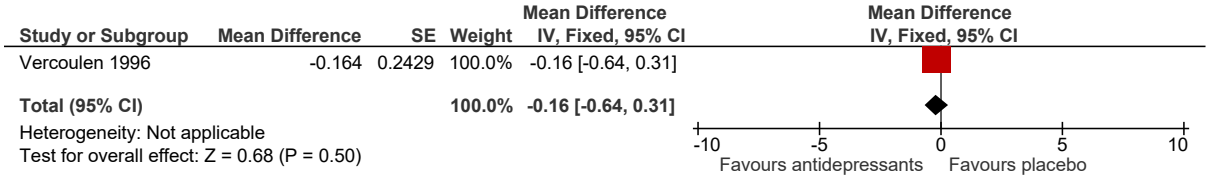


Figure 25: Physical functioning: Karnofsky Performance Index at 6 weeks (moclobemide)

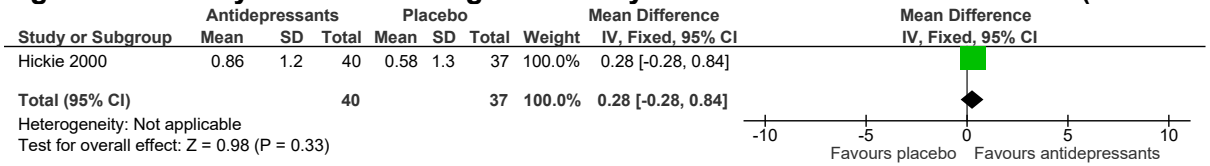
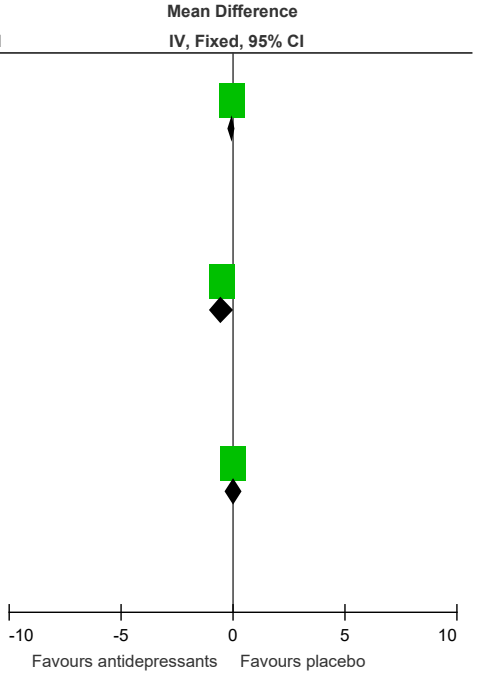


Figure 26: Psychological status: Profile of mood states (POMS) at 6 weeks (moclobemide)

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.17.1 Fatigue								
Hickie 2000	-0.05	0.4	40	-0.01	0.3	37	100.0%	-0.04 [-0.20, 0.12]
Subtotal (95% CI)			**40**			**37**	**100.0%**	**-0.04 [-0.20, 0.12]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.50 (P = 0.62)								
2.17.2 Vigour								
Hickie 2000 (1)	-0.51	1.2	40	0	1.1	37	100.0%	-0.51 [-1.02, 0.00]
Subtotal (95% CI)			**40**			**37**	**100.0%**	**-0.51 [-1.02, 0.00]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
2.17.3 Depression								
Hickie 2000	-0.06	1	40	-0.08	0.7	37	100.0%	0.02 [-0.36, 0.40]
Subtotal (95% CI)			**40**			**37**	**100.0%**	**0.02 [-0.36, 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.10 (P = 0.92)								



Footnotes

(1) Vigour subscale inverted for analysis

Figure 27: Psychological status: HADS depression at 12-26 weeks (change scores) (fluoxetine or duloxetine)

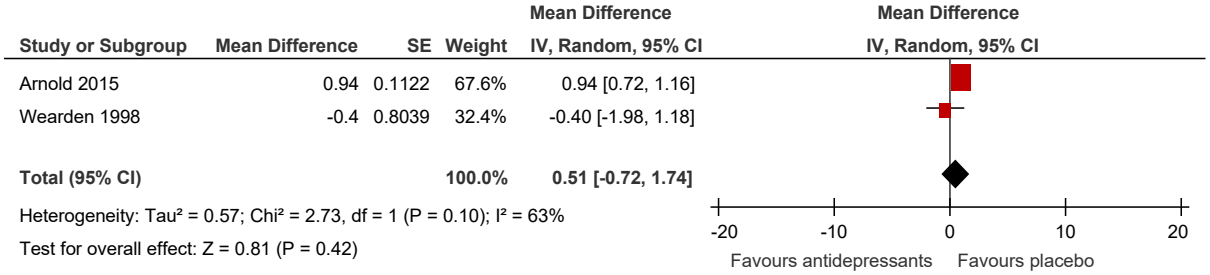


Figure 28: Psychological status: HADS anxiety at 12 weeks (duloxetine)

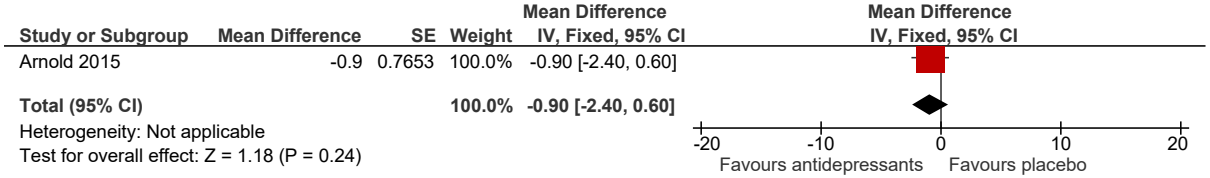


Figure 29: Psychological status: Beck Depression Inventory at 16 weeks (fluoxetine)

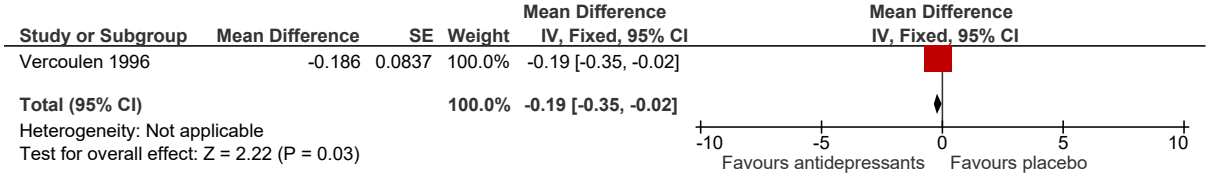


Figure 30: Pain: Brief Pain Inventory at 12 weeks (duloxetine)

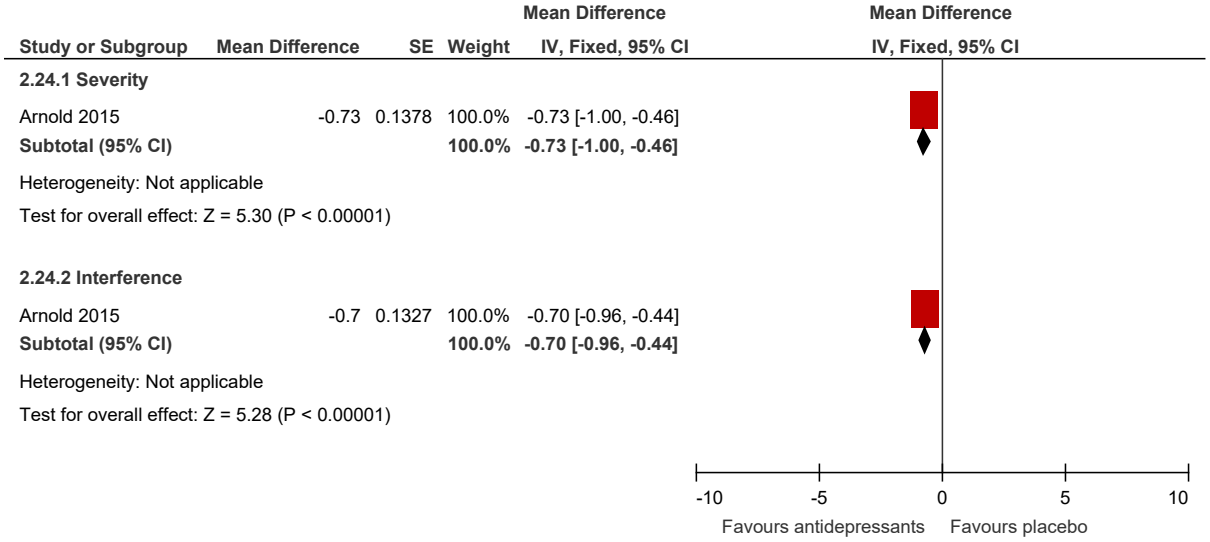


Figure 31: Adverse events (fluoxetine)

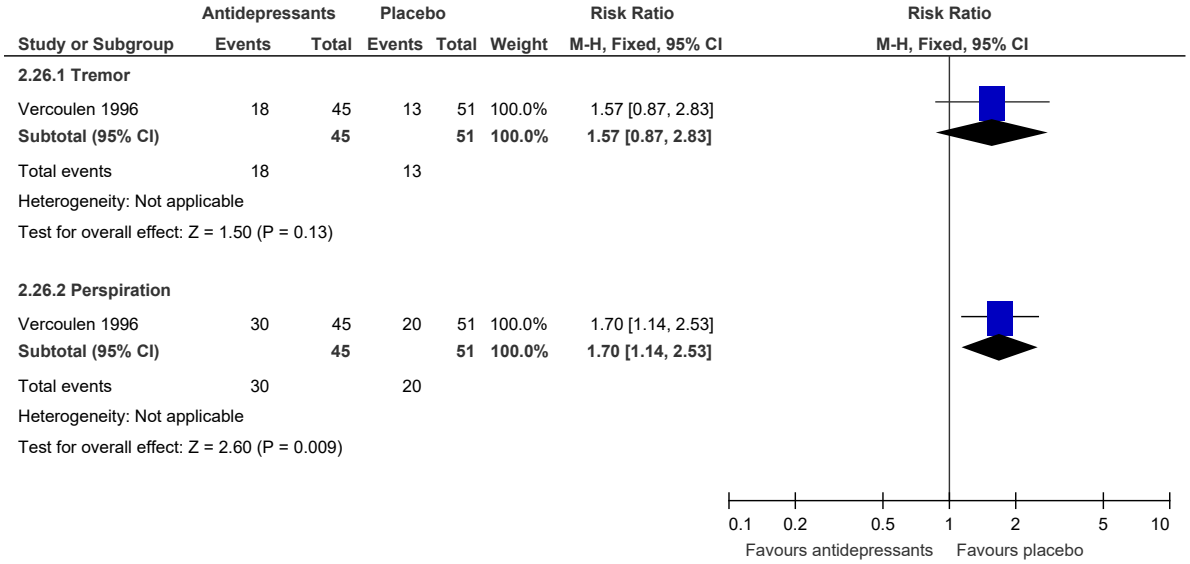


Figure 32: Exercise performance measure: VO2 max (mL O2/kg/min) at 26 weeks (fluoxetine)

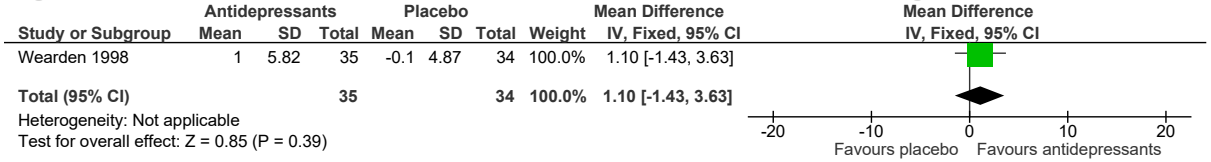


Figure 33: Symptom scales: Clinical Global Impression at 12 weeks (duloxetine)

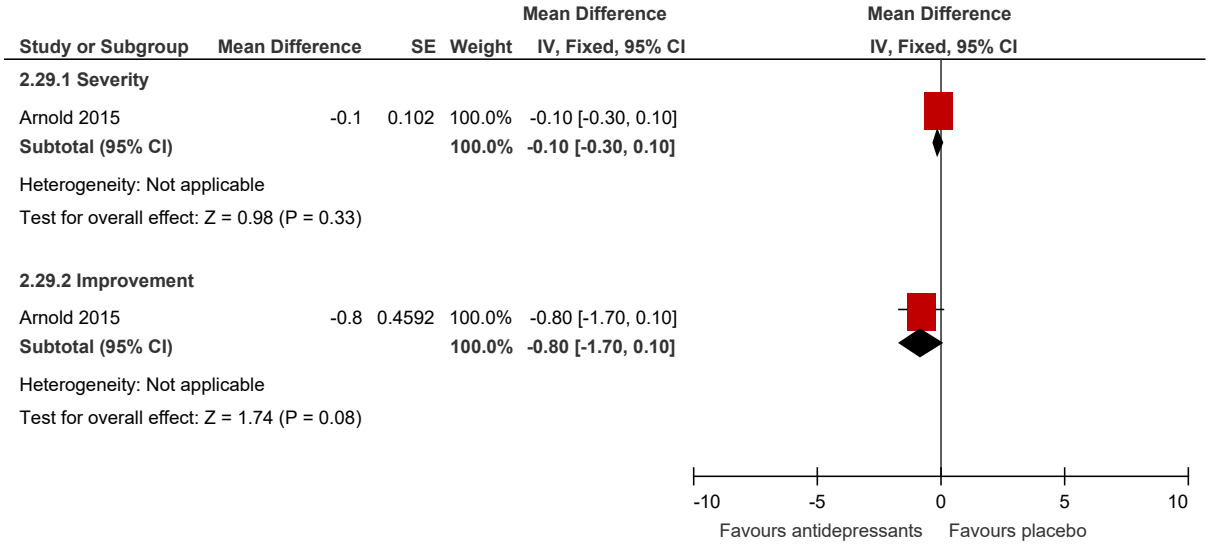


Figure 34: Symptom scales: CDC symptom inventory at 12 weeks (duloxetine)

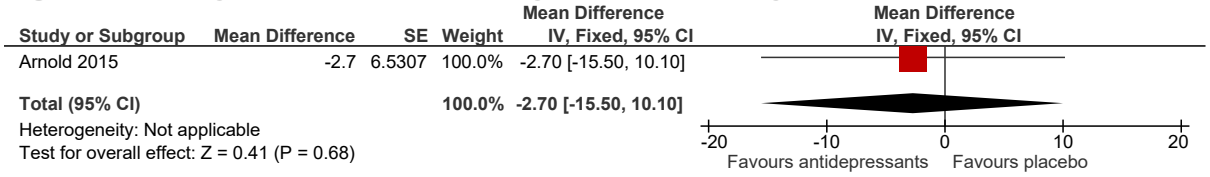


Figure 35: Symptom scales: Improvement of symptoms (patient-reported) at 6-14 weeks (fluoxetine or moclobemide)

Study or Subgroup	Antidepressants Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Hickie 2000	24	47	14	43	75.7%	1.57 [0.94, 2.62]
Vercoulen 1996	8	45	5	51	24.3%	1.81 [0.64, 5.15]
Total (95% CI)		**92**		**94**	**100.0%**	**1.63 [1.02, 2.59]**
Total events	32		19			

Heterogeneity: Chi² = 0.06, df = 1 (P = 0.80); I² = 0%
Test for overall effect: Z = 2.05 (P = 0.04)

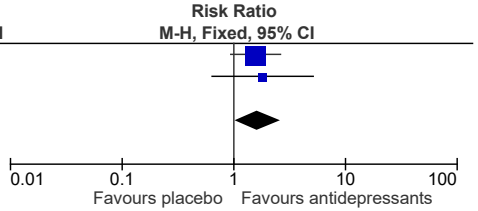
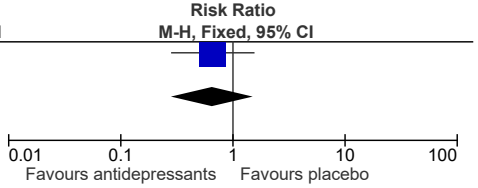


Figure 36: Symptom scales: Worsening of symptoms (patient-reported) at 14 weeks (fluoxetine)

Study or Subgroup	Antidepressants Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Vercoulen 1996	7	45	12	51	100.0%	0.66 [0.28, 1.53]
Total (95% CI)		**45**		**51**	**100.0%**	**0.66 [0.28, 1.53]**
Total events	7		12			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.96 (P = 0.34)



E.3 Antidepressants (fluoxetine) versus graded exercise

Figure 37: Fatigue: 14-item Chalder fatigue scale at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Exercise Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-3	8.15	35	-5.7	10.9	34	100.0%	2.70 [-1.85, 7.25]
Total (95% CI)			**35**			**34**	**100.0%**	**2.70 [-1.85, 7.25]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.16 (P = 0.24)

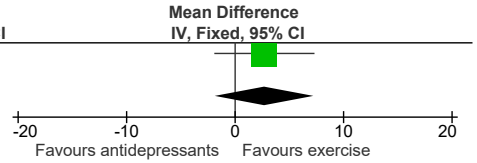


Figure 38: Psychological status: HADS depression at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Exercise Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-1.7	4.01	35	-1.2	3.49	34	100.0%	-0.50 [-2.27, 1.27]
Total (95% CI)			**35**			**34**	**100.0%**	**-0.50 [-2.27, 1.27]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.55 (P = 0.58)

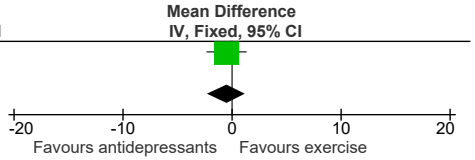
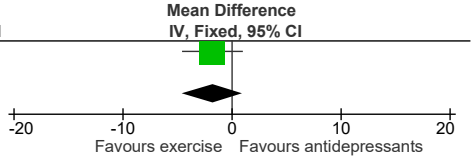


Figure 39: Exercise performance measure: VO2 max (mL O2/kg/min) at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Exercise Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	1	5.82	35	2.8	5.73	34	100.0%	-1.80 [-4.53, 0.93]
Total (95% CI)			**35**			**34**	**100.0%**	**-1.80 [-4.53, 0.93]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.29 (P = 0.20)



E.4 Antidepressants (fluoxetine) versus combined antidepressants & graded exercise

Figure 40: Fatigue: 14-item Chalder fatigue scale at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-3	8.15	35	-6	10.43	33	100.0%	3.00 [-1.47, 7.47]
Total (95% CI)			**35**			**33**	**100.0%**	**3.00 [-1.47, 7.47]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.32 (P = 0.19)

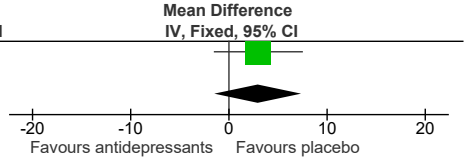


Figure 41: Psychological status: HADS depression at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-1.7	4.01	35	-2	3.67	34	100.0%	0.30 [-1.51, 2.11]
Total (95% CI)			**35**			**34**	**100.0%**	**0.30 [-1.51, 2.11]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.32 (P = 0.75)

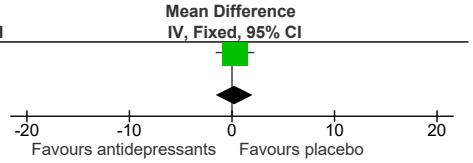


Figure 42: Exercise performance measure: VO2 max (mL O2/kg/min) at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Combined exercise Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	1	5.82	35	2	4.23	33	100.0%	-1.00 [-3.41, 1.41]
Total (95% CI)			**35**			**33**	**100.0%**	**-1.00 [-3.41, 1.41]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.81 (P = 0.42)

E.5 Combined antidepressants (fluoxetine) & graded exercise versus placebo

Figure 43: Fatigue: 14-item Chalder fatigue scale at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-6	10.43	33	-2.7	7.77	34	100.0%	-3.30 [-7.71, 1.11]
Total (95% CI)			**33**			**34**	**100.0%**	**-3.30 [-7.71, 1.11]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.47 (P = 0.14)

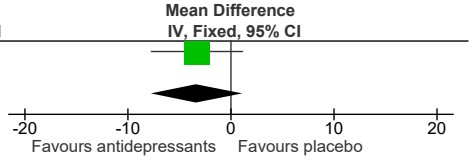


Figure 44: Psychological status: HADS depression at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-2	3.67	33	-1.3	2.87	34	100.0%	-0.70 [-2.28, 0.88]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.70 [-2.28, 0.88]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.87 (P = 0.39)

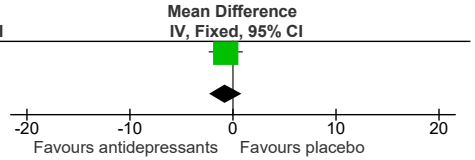
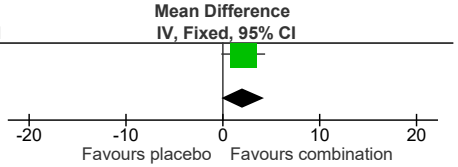


Figure 45: Exercise performance measure: VO2 max (mL O2/kg/min) at 26 weeks

Study or Subgroup	Combination Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	2	4.23	33	-0.1	4.87	34	100.0%	2.10 [-0.08, 4.28]
Total (95% CI)			**33**			**34**	**100.0%**	**2.10 [-0.08, 4.28]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.89 (P = 0.06)



E.6 Combined antidepressants (fluoxetine) & graded exercise versus graded exercise

Figure 46: Fatigue: 14-item Chalder fatigue scale at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-6	10.43	33	-5.7	10.9	34	100.0%	-0.30 [-5.41, 4.81]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.30 [-5.41, 4.81]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.12 (P = 0.91)

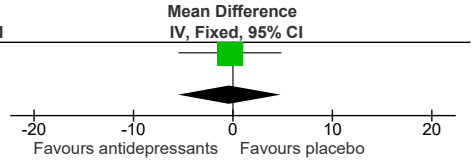


Figure 47: Psychological status: HADS depression at 26 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	-2	3.67	33	-1.2	3.49	34	100.0%	-0.80 [-2.52, 0.92]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.80 [-2.52, 0.92]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.91 (P = 0.36)

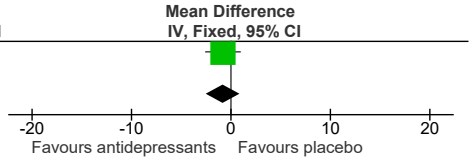
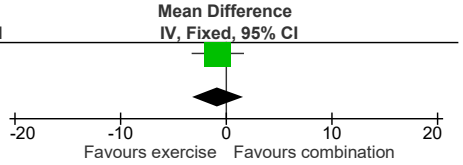


Figure 48: Exercise performance measure: VO2 max (mL O2/kg/min) at 26 weeks

Study or Subgroup	Combination Mean	SD	Total	Exercise Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Wearden 1998	2	4.23	33	2.8	5.73	34	100.0%	-0.80 [-3.21, 1.61]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.80 [-3.21, 1.61]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.65 (P = 0.51)



E.7 Antidepressants (fluoxetine) versus antipsychotics (amisulpride)

Figure 49: Quality of Life: SF12 at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Pardini 2011	37.6	4.9	20	53.2	4.8	20	100.0%	-15.60 [-18.61, -12.59]
Total (95% CI)			**20**			**20**	**100.0%**	**-15.60 [-18.61, -12.59]**

Heterogeneity: Not applicable
Test for overall effect: Z = 10.17 (P < 0.00001)

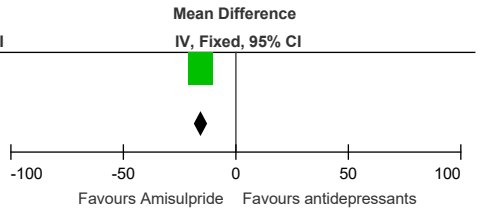


Figure 50: Fatigue: Fatigue Severity Scale at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Pardini 2011	48.9	4.9	20	36.3	8.6	20	100.0%	12.60 [8.26, 16.94]
Total (95% CI)			**20**			**20**	**100.0%**	**12.60 [8.26, 16.94]**

Heterogeneity: Not applicable

Test for overall effect: Z = 5.69 (P < 0.00001)

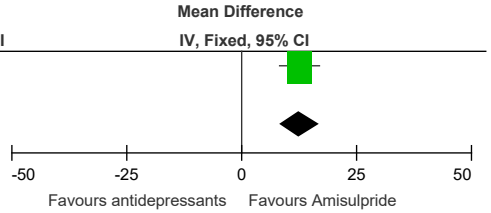


Figure 51: Psychological status: Hospital anxiety and depression scale (HADS) at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
7.3.1 Anxiety								
Pardini 2011	4.9	1	20	4.5	1	20	100.0%	0.40 [-0.22, 1.02]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.40 [-0.22, 1.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.26 (P = 0.21)								
7.3.2 Depression								
Pardini 2011	4.2	1	20	4.3	0.9	20	100.0%	-0.10 [-0.69, 0.49]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**-0.10 [-0.69, 0.49]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.33 (P = 0.74)								

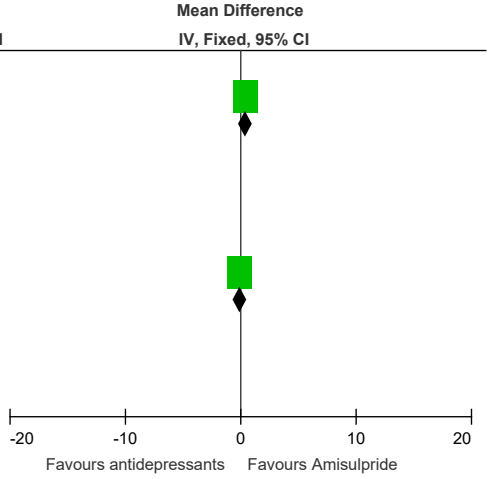


Figure 52: Pain: pain on VAS 0-100 at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Pardini 2011	53.1	8.3	20	40.5	13.1	20	100.0%	12.60 [5.80, 19.40]
Total (95% CI)			**20**			**20**	**100.0%**	**12.60 [5.80, 19.40]**

Heterogeneity: Not applicable

Test for overall effect: Z = 3.63 (P = 0.0003)

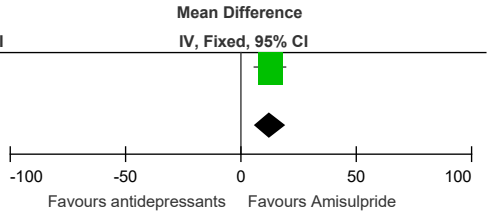


Figure 53: Adverse events: FIBSER global burden at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Pardini 2011	0.6	0.8	20	0.8	0.7	20	100.0%	-0.20 [-0.67, 0.27]
Total (95% CI)			**20**			**20**	**100.0%**	**-0.20 [-0.67, 0.27]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.84 (P = 0.40)

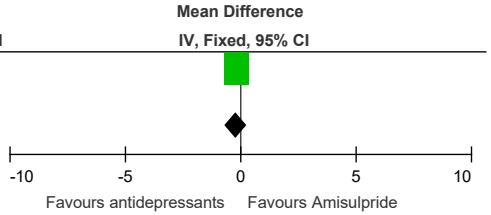
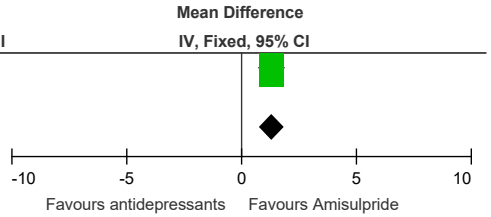


Figure 54: Symptom scales: Clinical Global Impression Severity (CGI-S) at 12 weeks

Study or Subgroup	Antidepressants Mean	SD	Total	Amisulpride Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Pardini 2011	4.2	1.1	20	2.9	0.6	20	100.0%	1.30 [0.75, 1.85]
Total (95% CI)			**20**			**20**	**100.0%**	**1.30 [0.75, 1.85]**

Heterogeneity: Not applicable

Test for overall effect: Z = 4.64 (P < 0.00001)



E.8 Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo

Figure 55: Quality of Life: SF36 at 6 weeks (fludrocortisone)

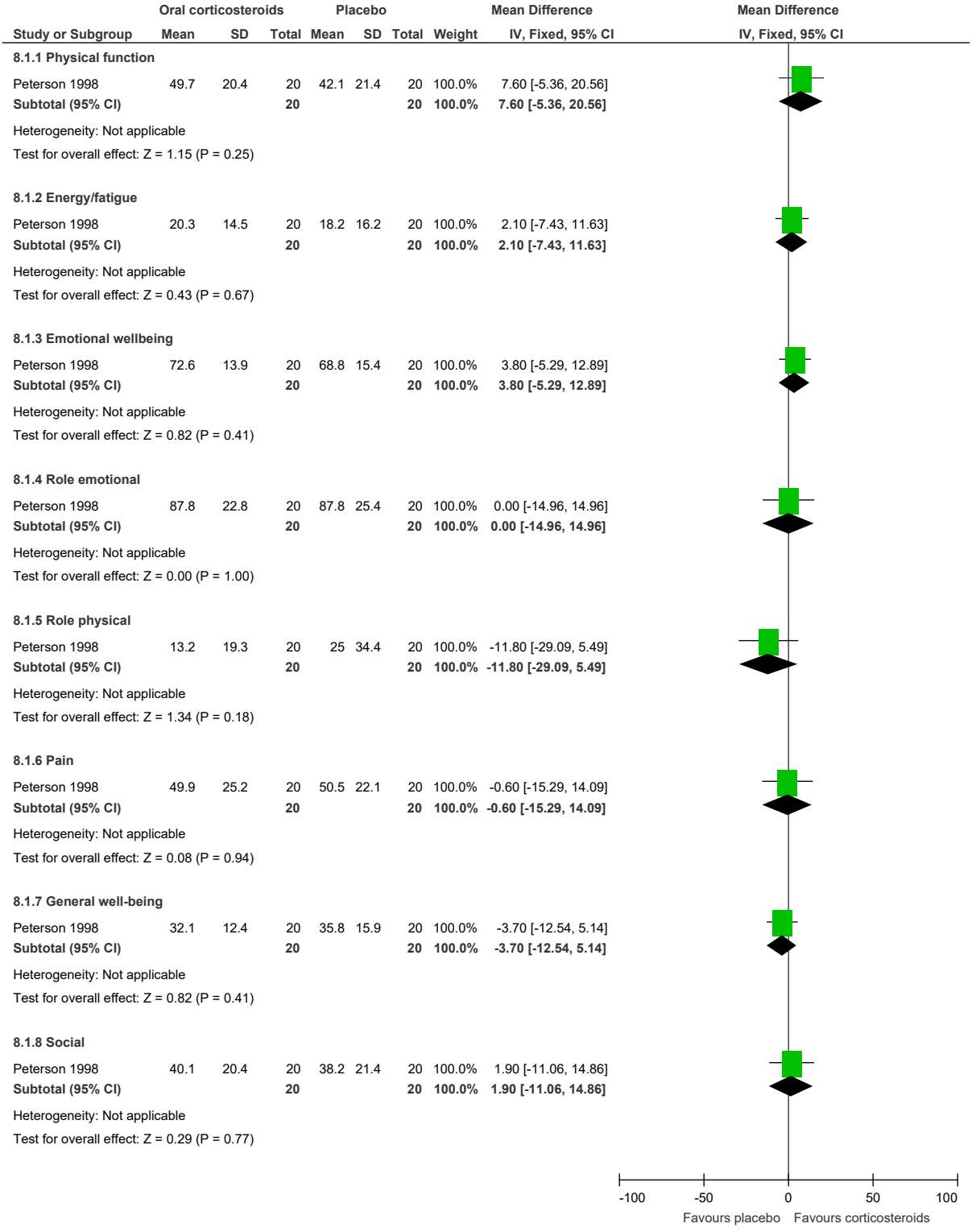


Figure 56: Fatigue: fatigue on VAS 0-10 at 6 weeks (fludrocortisone)

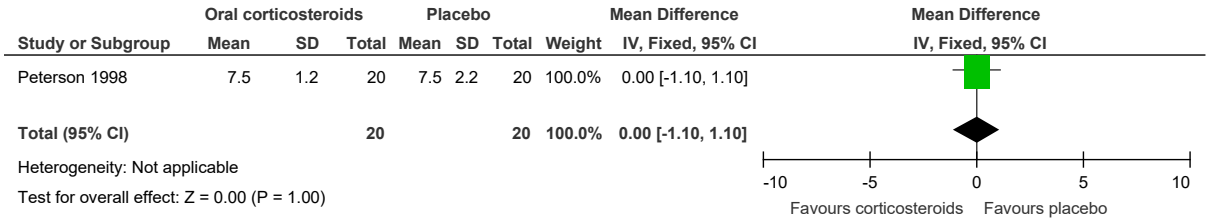


Figure 57: Fatigue: Chronic Fatigue Syndrome Severity Rating at 4-8 weeks (nasal flunisolide)

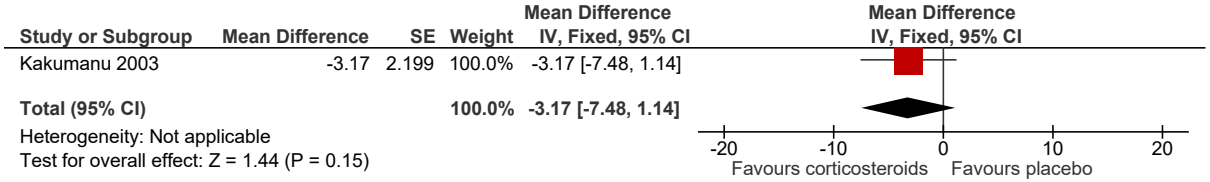


Figure 58: Fatigue: Profile of Mood States – fatigue at 11 weeks (fludrocortisone)

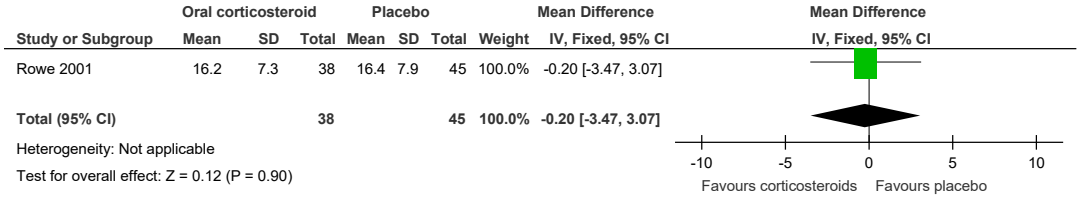


Figure 59: Fatigue: Profile of Mood States - fatigue at 12 weeks (hydrocortisone)

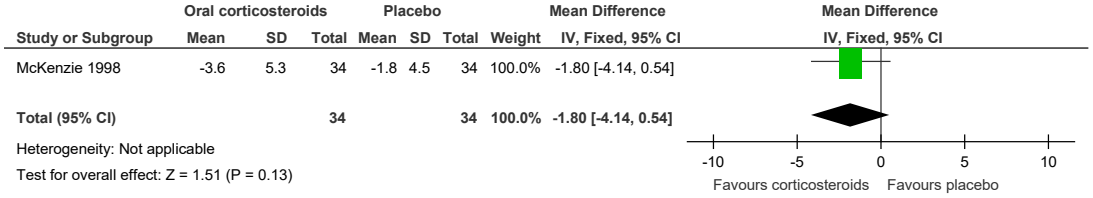


Figure 60: Fatigue: Profile of Mood States – vigour at 11 weeks (fludrocortisone)

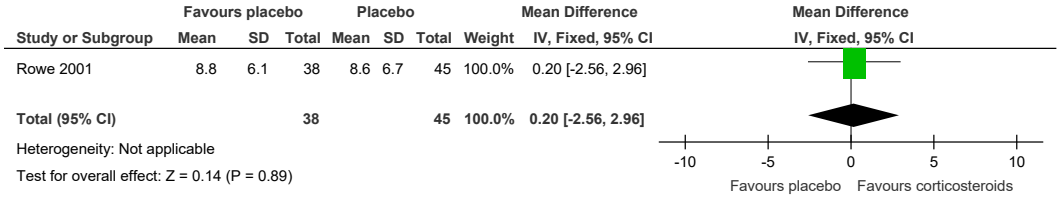


Figure 61: Fatigue: Profile of Mood States - vigour at 12 weeks (hydrocortisone)

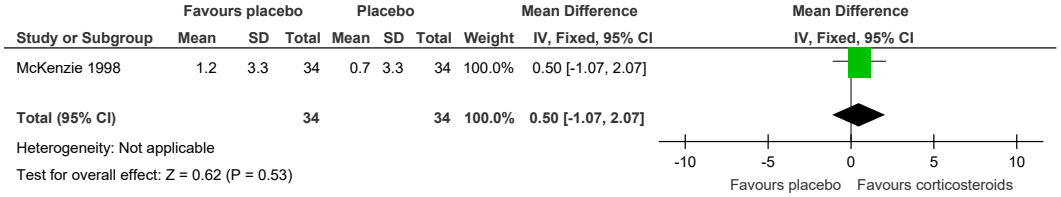


Figure 62: Fatigue: Wood Mental Fatigue Inventory at 11 weeks (fludrocortisone)

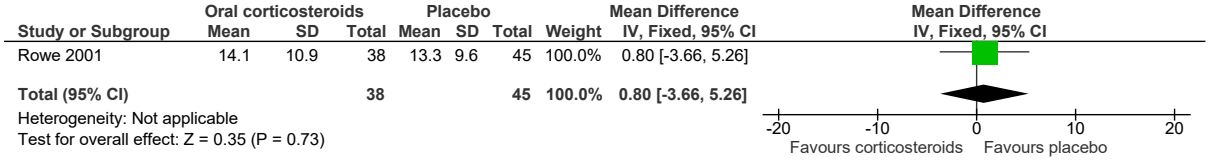


Figure 63: Physical function: SF36 physical function at 11 weeks (fludrocortisone)

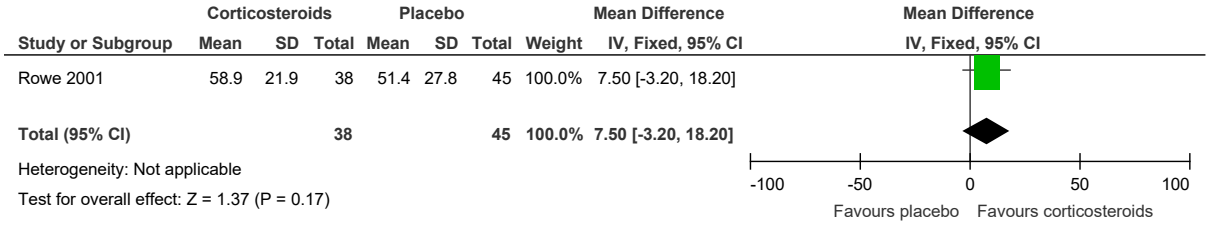


Figure 64: Adverse events: Adverse events leading to study withdrawal at 6 weeks (fludrocortisone)

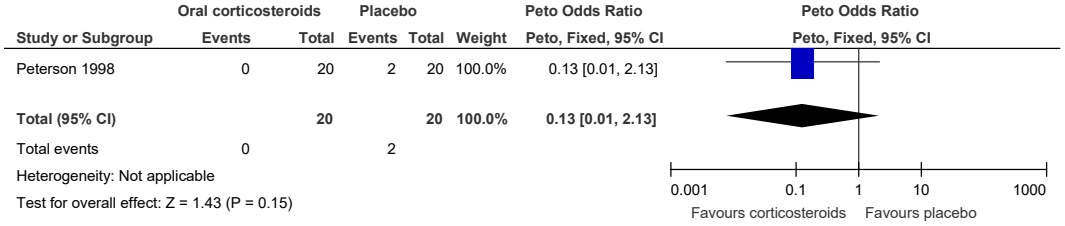


Figure 65: Adverse events: adverse effects / adverse events at 6-11 weeks (fludrocortisone)

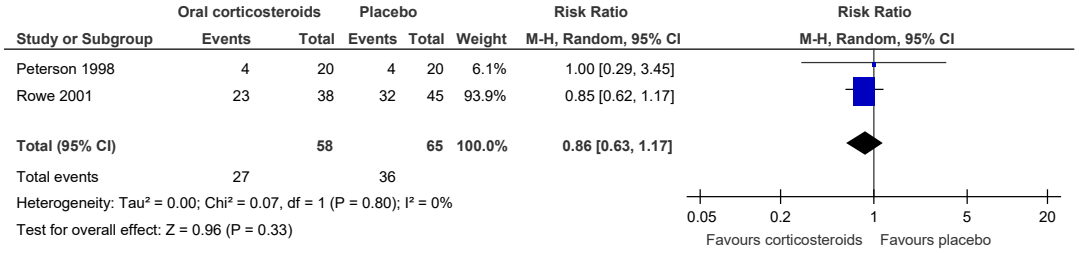


Figure 66: Adverse events: any adverse reaction at 12 weeks (hydrocortisone)

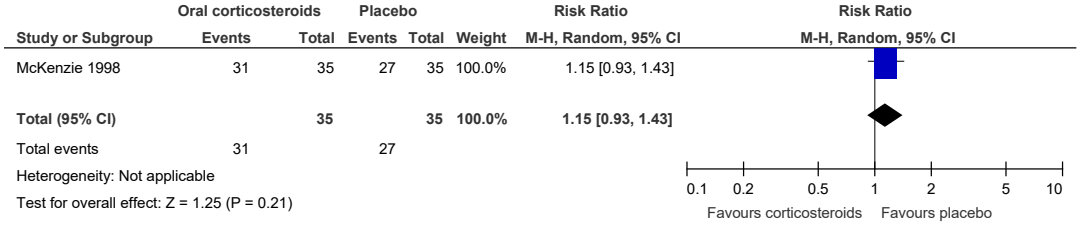


Figure 67: Psychological status: SF36 mental health at 11 weeks (fludrocortisone)

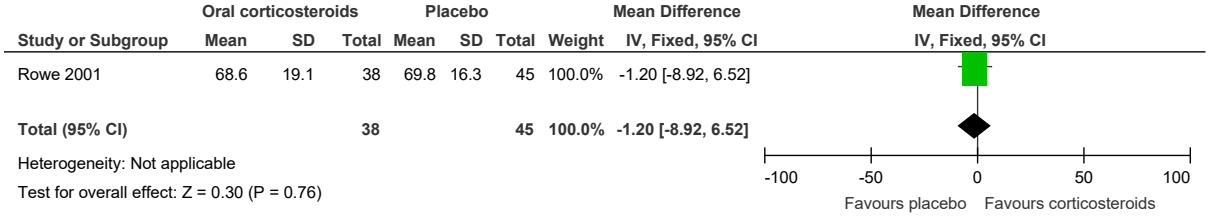


Figure 68: Psychological status: Beck Depression Inventory at 11 weeks (fludrocortisone)

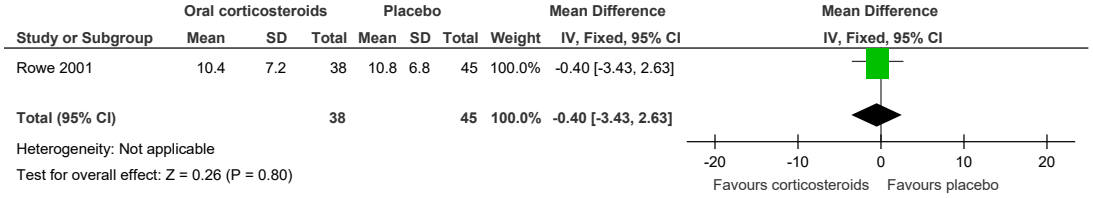


Figure 69: Psychological status: Beck Depression Inventory at 12 weeks (hydrocortisone)

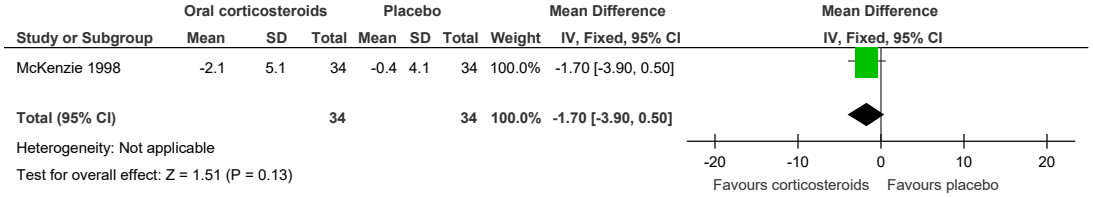


Figure 70: Psychological status: Profile of Mood States - anger, at 12 weeks (hydrocortisone)

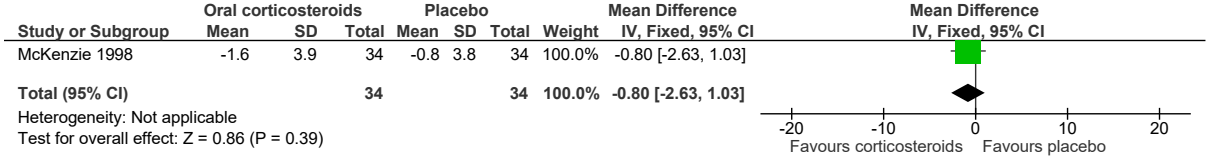


Figure 71: Psychological status: Profile of Mood States - anxiety, at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-0.8	2.5	34	-2.1	3.6	34	100.0%	1.30 [-0.17, 2.77]
Total (95% CI)			**34**			**34**	**100.0%**	**1.30 [-0.17, 2.77]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.73 (P = 0.08)

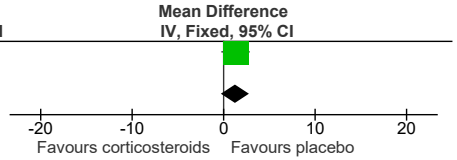


Figure 72: Psychological status: Profile of Mood States - confusion, at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-1.1	3.3	34	-1.4	2.9	34	100.0%	0.30 [-1.18, 1.78]
Total (95% CI)			**34**			**34**	**100.0%**	**0.30 [-1.18, 1.78]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.40 (P = 0.69)

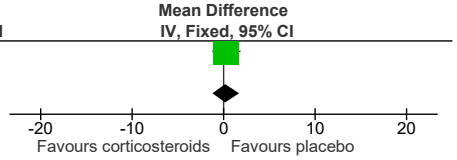


Figure 73: Psychological status: Profile of Mood States - depression, at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-1.6	4.6	34	0	3.8	34	100.0%	-1.60 [-3.61, 0.41]
Total (95% CI)			**34**			**34**	**100.0%**	**-1.60 [-3.61, 0.41]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.56 (P = 0.12)

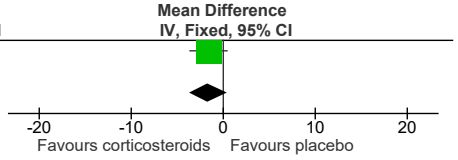


Figure 74: Psychological status: Symptom checklist-90-R general sensitivity index at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-0.1	0.2	34	-0.1	0.2	34	100.0%	0.00 [-0.10, 0.10]
Total (95% CI)			**34**			**34**	**100.0%**	**0.00 [-0.10, 0.10]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)

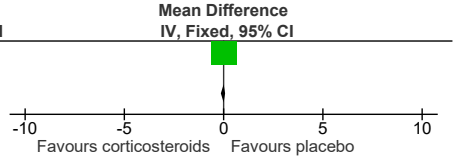


Figure 75: Psychological status: Symptom checklist-90-R positive symptom distress index at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	0	0.3	34	-0.1	0.3	34	100.0%	0.10 [-0.04, 0.24]
Total (95% CI)			**34**			**34**	**100.0%**	**0.10 [-0.04, 0.24]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.37 (P = 0.17)

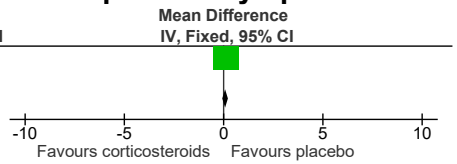


Figure 76: Psychological status: Symptom checklist-90-R positive symptom total at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-2.6	10.8	34	-2.4	11.5	34	100.0%	-0.20 [-5.50, 5.10]
Total (95% CI)			**34**			**34**	**100.0%**	**-0.20 [-5.50, 5.10]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.07 (P = 0.94)

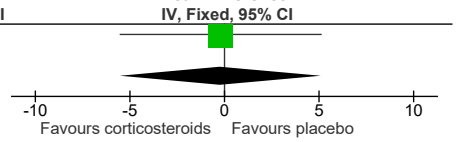


Figure 77: Psychological status: Hamilton Depression Rating Scale at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	-0.8	3.8	32	0.1	2.9	33	100.0%	-0.90 [-2.55, 0.75]
Total (95% CI)			**32**			**33**	**100.0%**	**-0.90 [-2.55, 0.75]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.07 (P = 0.28)

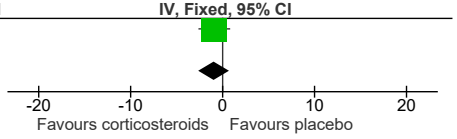


Figure 78: Psychological status: Positive and negative effect scale (PANAS) positive affect at 6 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Peterson 1998	22.7	8.3	20	21.7	6.7	20	100.0%	1.00 [-3.67, 5.67]
Total (95% CI)			**20**			**20**	**100.0%**	**1.00 [-3.67, 5.67]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.42 (P = 0.68)

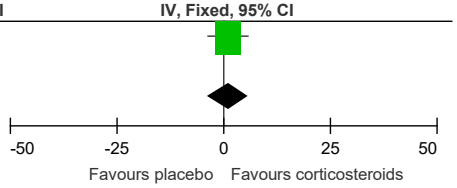


Figure 79: Activity levels: activity scale at 12 weeks (hydrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McKenzie 1998	0.3	1.1	34	0.7	1.4	34	100.0%	-0.40 [-1.00, 0.20]
Total (95% CI)			**34**			**34**	**100.0%**	**-0.40 [-1.00, 0.20]**

Heterogeneity: Not applicable

Test for overall effect: Z = 1.31 (P = 0.19)

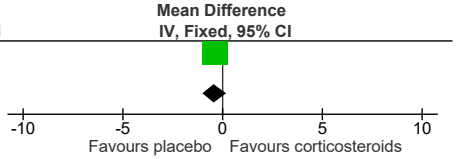


Figure 80: Activity levels: distance before exhausted (ordinal scale) at 6 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Peterson 1990	2.7	1	20	2.7	1.3	20	100.0%	0.00 [-0.72, 0.72]
Total (95% CI)			**20**			**20**	**100.0%**	**0.00 [-0.72, 0.72]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.00 (P = 1.00)

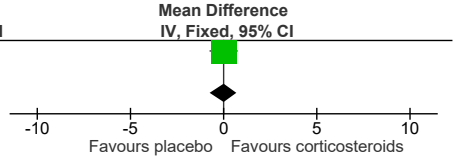


Figure 81: Activity levels: Duke Activity Status Index at 11 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Rowe 2001	9.2	10.6	38	6.7	7.3	45	100.0%	2.50 [-1.49, 6.49]
Total (95% CI)			**38**			**45**	**100.0%**	**2.50 [-1.49, 6.49]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.23 (P = 0.22)

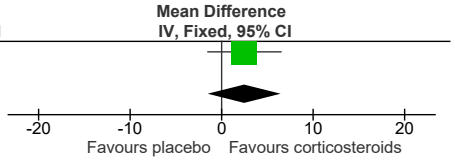


Figure 82: Cognitive function: Reaction time (secs) at 6 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Peterson 1998	0.35	0.07	20	0.36	0.08	20	100.0%	-0.01 [-0.06, 0.04]
Total (95% CI)			**20**			**20**	**100.0%**	**-0.01 [-0.06, 0.04]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.42 (P = 0.67)

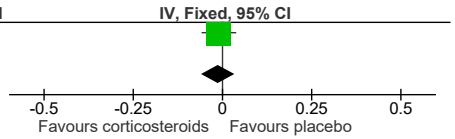


Figure 83: Cognitive function: VAS 0-10 at 6 weeks (fludrocortisone)

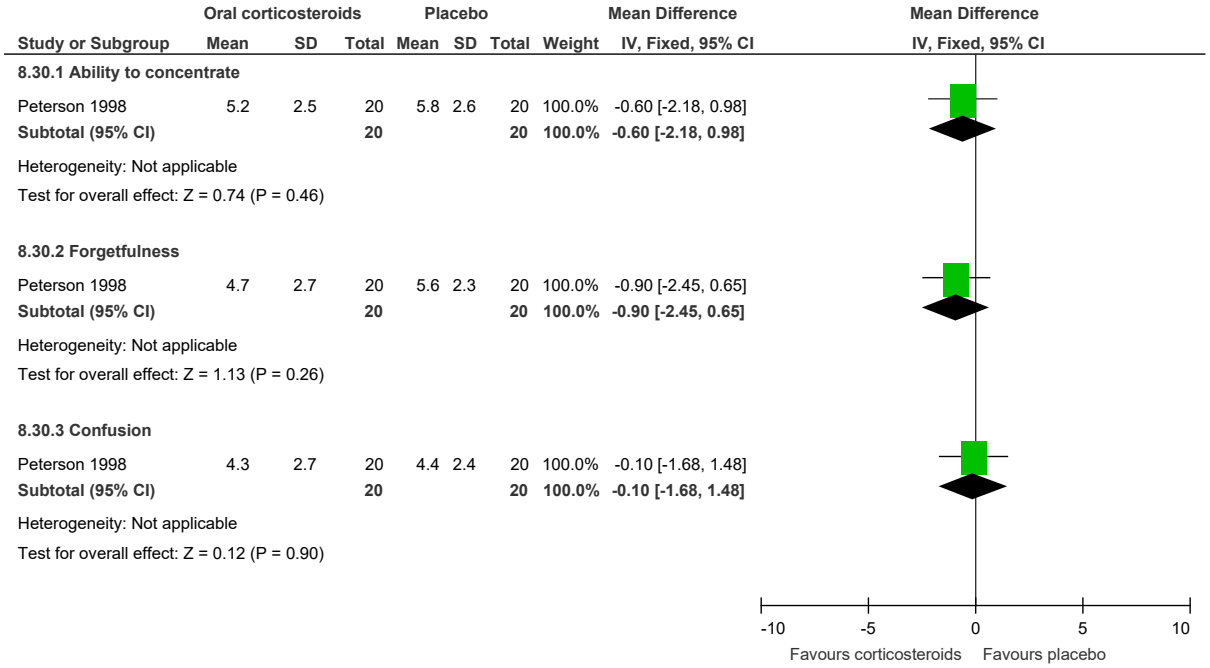


Figure 84: Pain: pain on VAS 0-10 at 6 weeks (fludrocortisone)

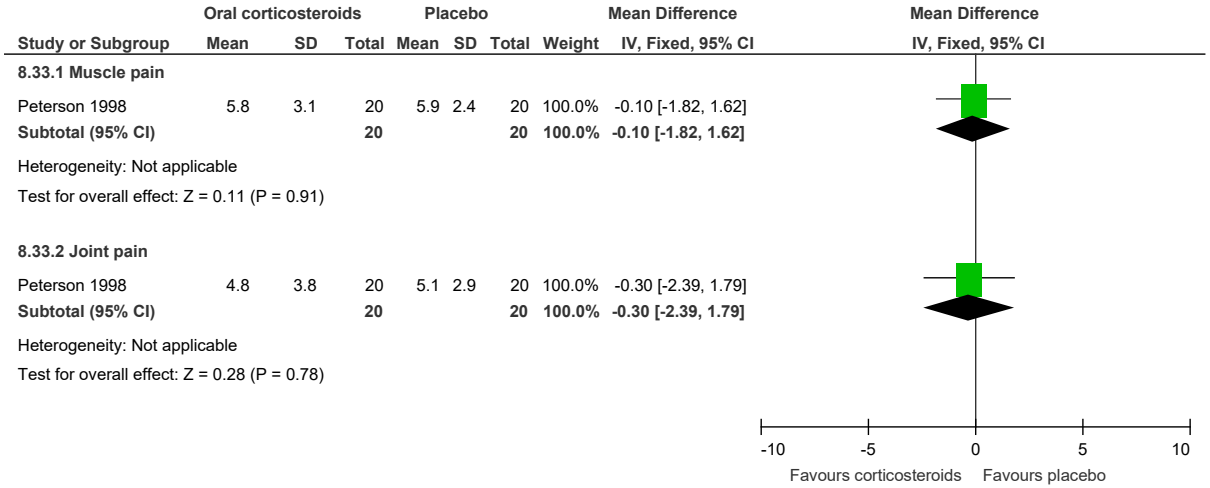


Figure 85: Sleep quality: unrefreshing sleep on VAS 0-10 at 6 weeks (fludrocortisone)

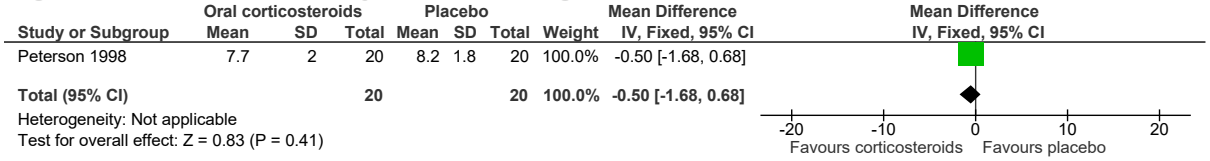


Figure 86: Sleep quality: Functional Outcomes of Sleep Questionnaire at 4-8 weeks (nasal flunisolide)

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
Kakumanu 2003	-3.18	1.7296	100.0%	-3.18 [-6.57, 0.21]
Total (95% CI)			**100.0%**	**-3.18 [-6.57, 0.21]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.84 (P = 0.07)

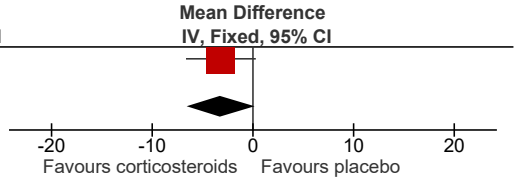


Figure 87: Exercise performance measure: Treadmill time (mins) at 6 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Peterson 1998	22.8	9.2	20	20.2	11.5	20	100.0%	2.60 [-3.85, 9.05]
Total (95% CI)			**20**			**20**	**100.0%**	**2.60 [-3.85, 9.05]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

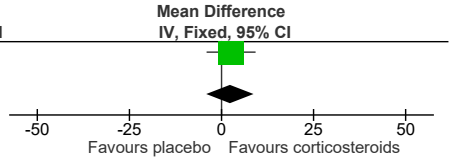


Figure 88: Symptom scales: Wellness scale at 11 weeks (fludrocortisone)

Study or Subgroup	Oral corticosteroids Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Rowe 2001	3.8	11.5	38	2.7	10	45	100.0%	1.10 [-3.58, 5.78]
Total (95% CI)			**38**			**45**	**100.0%**	**1.10 [-3.58, 5.78]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.46 (P = 0.65)

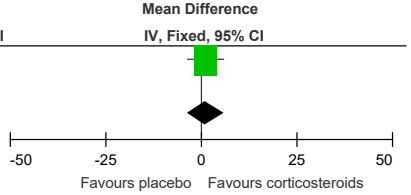


Figure 89: Symptom scales: Wellness scale at 12 weeks (hydrocortisone)

	Oral corticosteroids			Placebo				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
McKenzie 1998	6.3	11.7	30	1.7	8.8	35	100.0%	4.60 [-0.50, 9.70]
Total (95% CI)			**30**			**35**	**100.0%**	**4.60 [-0.50, 9.70]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.77 (P = 0.08)

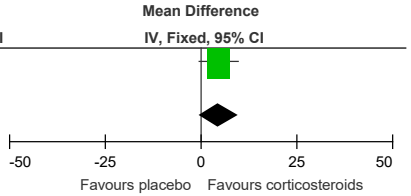


Figure 90: Symptom scales: Sickness Impact Profile at 12 weeks (hydrocortisone)

	Oral corticosteroids			Placebo				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
McKenzie 1998	-2.5	6.4	33	-2.2	6.8	34	100.0%	-0.30 [-3.46, 2.86]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.30 [-3.46, 2.86]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.19 (P = 0.85)

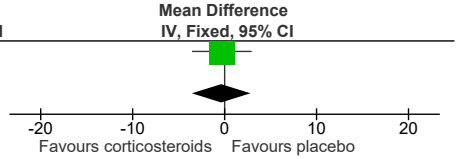
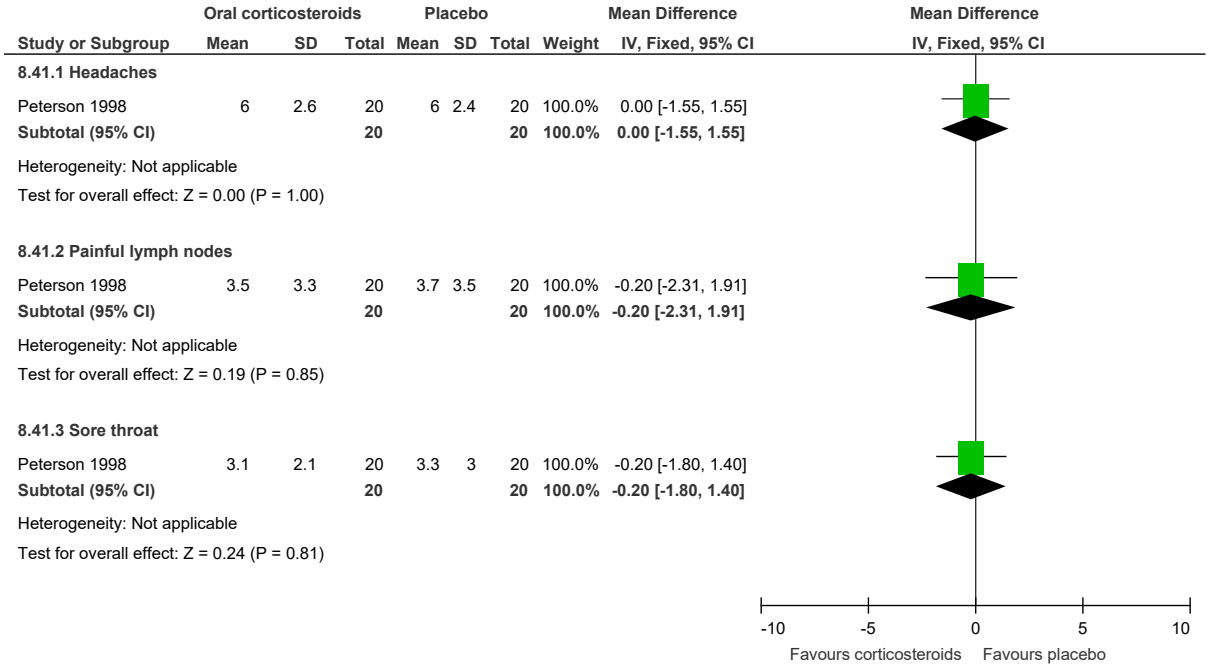


Figure 91: Symptom scales: VAS 0-10 at 6 weeks



E.9 Central antihypertensive drugs (clonidine) versus placebo

Figure 92: Cognitive function: Stockings of Cambridge at 30 minutes

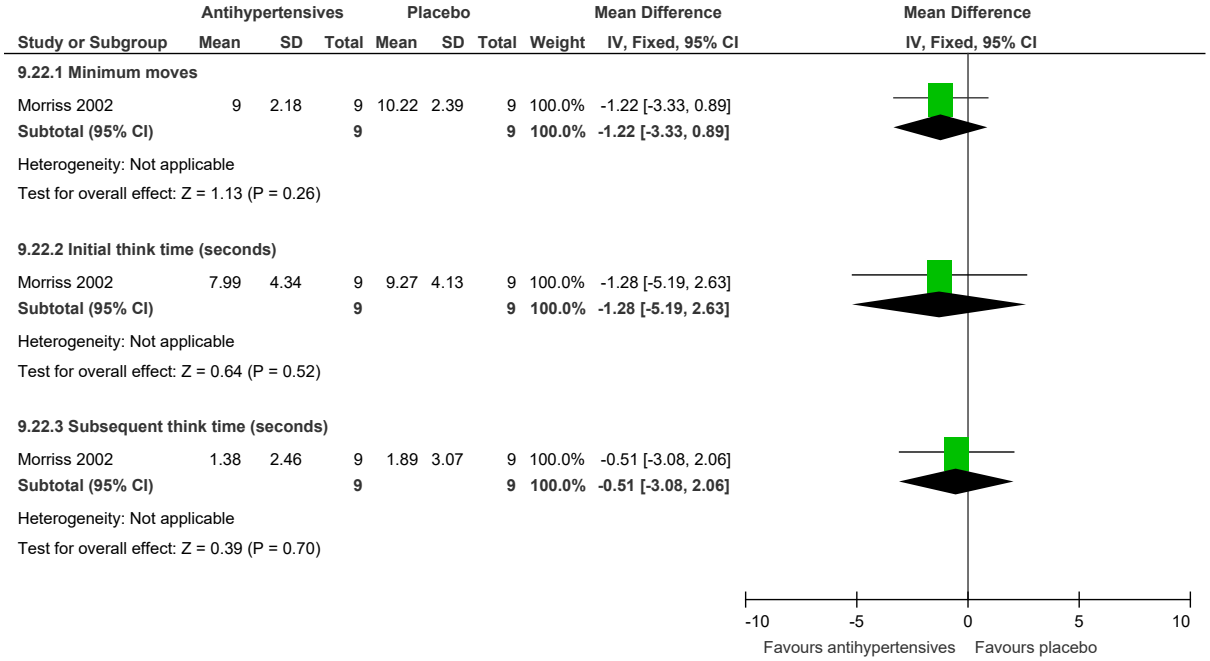


Figure 93: Cognitive function: Rapid Visual Information Processing - reaction time (secs) at 30 minutes

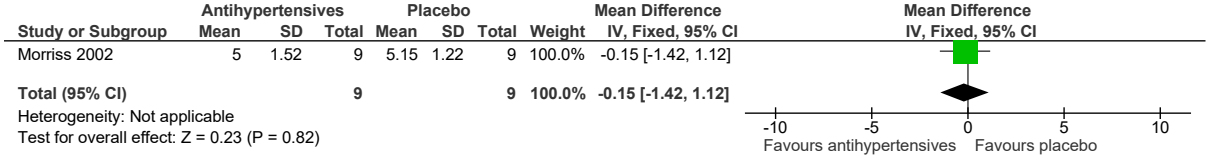


Figure 94: Cognitive function: set sift errors at 30 minutes

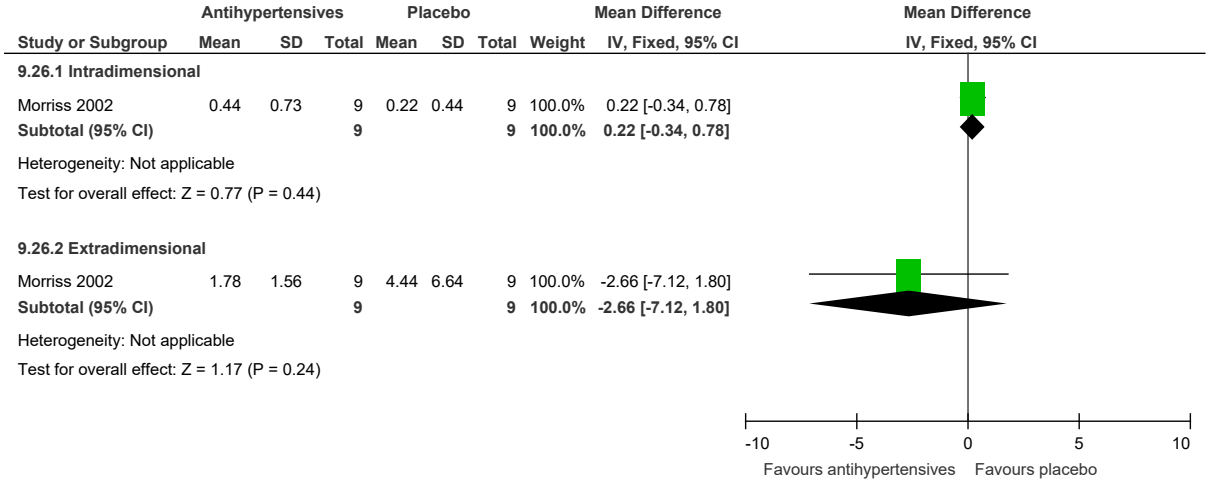


Figure 95: Cognitive function: Spatial working memory at 30 minutes

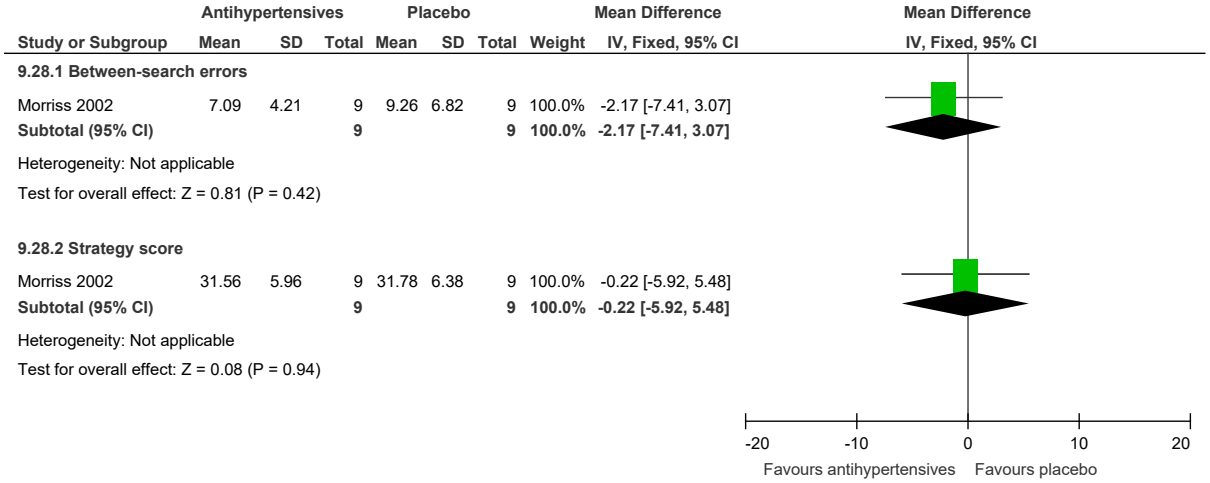


Figure 96: Cognitive function: recognition - number correct at 30 minutes

Study or Subgroup	Antihypertensives Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
9.30.1 Pattern recognition								
Morriss 2002	22.3	1.3	9	21.4	2.2	9	100.0%	0.90 [-0.77, 2.57]
Subtotal (95% CI)			**9**			**9**	**100.0%**	**0.90 [-0.77, 2.57]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.06 (P = 0.29)								
9.30.2 Spatial recognition								
Morriss 2002	15.2	2.9	9	15.3	2.1	9	100.0%	-0.10 [-2.44, 2.24]
Subtotal (95% CI)			**9**			**9**	**100.0%**	**-0.10 [-2.44, 2.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.93)								

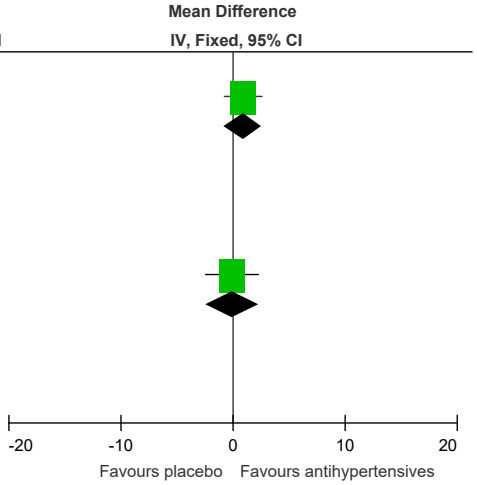


Figure 97: Cognitive function: spatial span - length at 30 minutes

Study or Subgroup	Antihypertensives Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Morriss 2002	6.4	1.26	9	6.1	1.2	9	100.0%	0.30 [-0.84, 1.44]
Total (95% CI)			**9**			**9**	**100.0%**	**0.30 [-0.84, 1.44]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.52 (P = 0.60)								

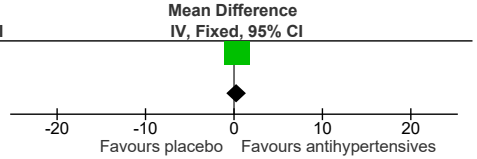


Figure 98: Cognitive function: delayed matching to sample 2 sec delay at 30 minutes

Study or Subgroup	Antihypertensives Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Morriss 2002	6.56	1.69	9	7.78	1.39	9	100.0%	-1.22 [-2.65, 0.21]
Total (95% CI)			**9**			**9**	**100.0%**	**-1.22 [-2.65, 0.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.67 (P = 0.09)								

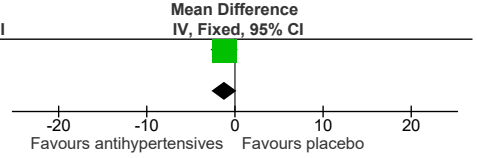
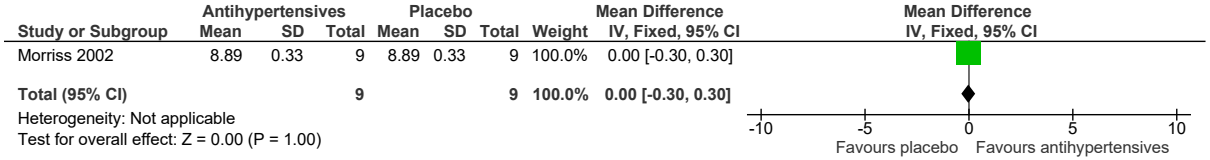


Figure 99: Cognitive function: paired associate learning - sets completed at 30 minutes



E.10 Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo

Figure 100: Quality of Life: SF36 physical total at 4-6 weeks (methylphenidate or dexamphetamine)

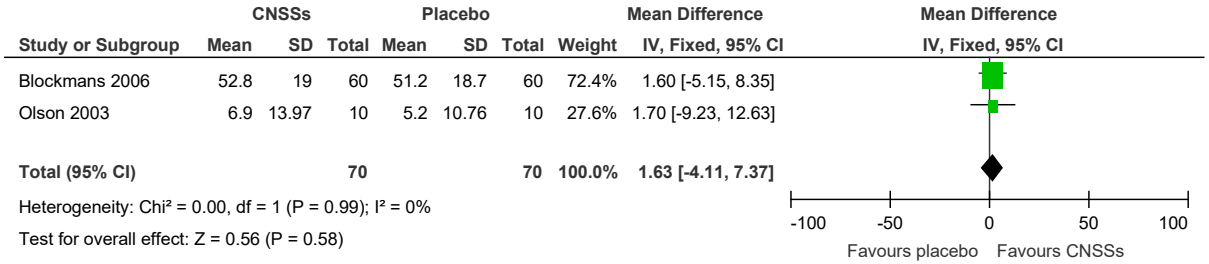


Figure 101: Quality of Life: SF36 mental total at 4-6 weeks (methylphenidate or dexamphetamine)

	CNSSs			Placebo				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Blockmans 2006	51.8	16.4	60	47.3	16.7	60	76.5%	4.50 [-1.42, 10.42]	
Olson 2003	4.2	11.46	10	3.9	12.86	10	23.5%	0.30 [-10.38, 10.98]	
Total (95% CI)			**70**			**70**	**100.0%**	**3.51 [-1.67, 8.69]**	

Heterogeneity: Chi² = 0.45, df = 1 (P = 0.50); I² = 0%

Test for overall effect: Z = 1.33 (P = 0.18)

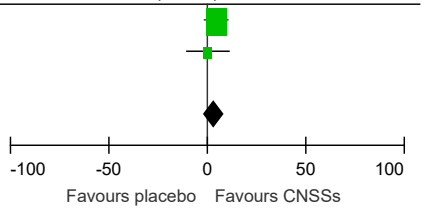


Figure 102: Quality of Life: SF36 at 20 days (modafinil)

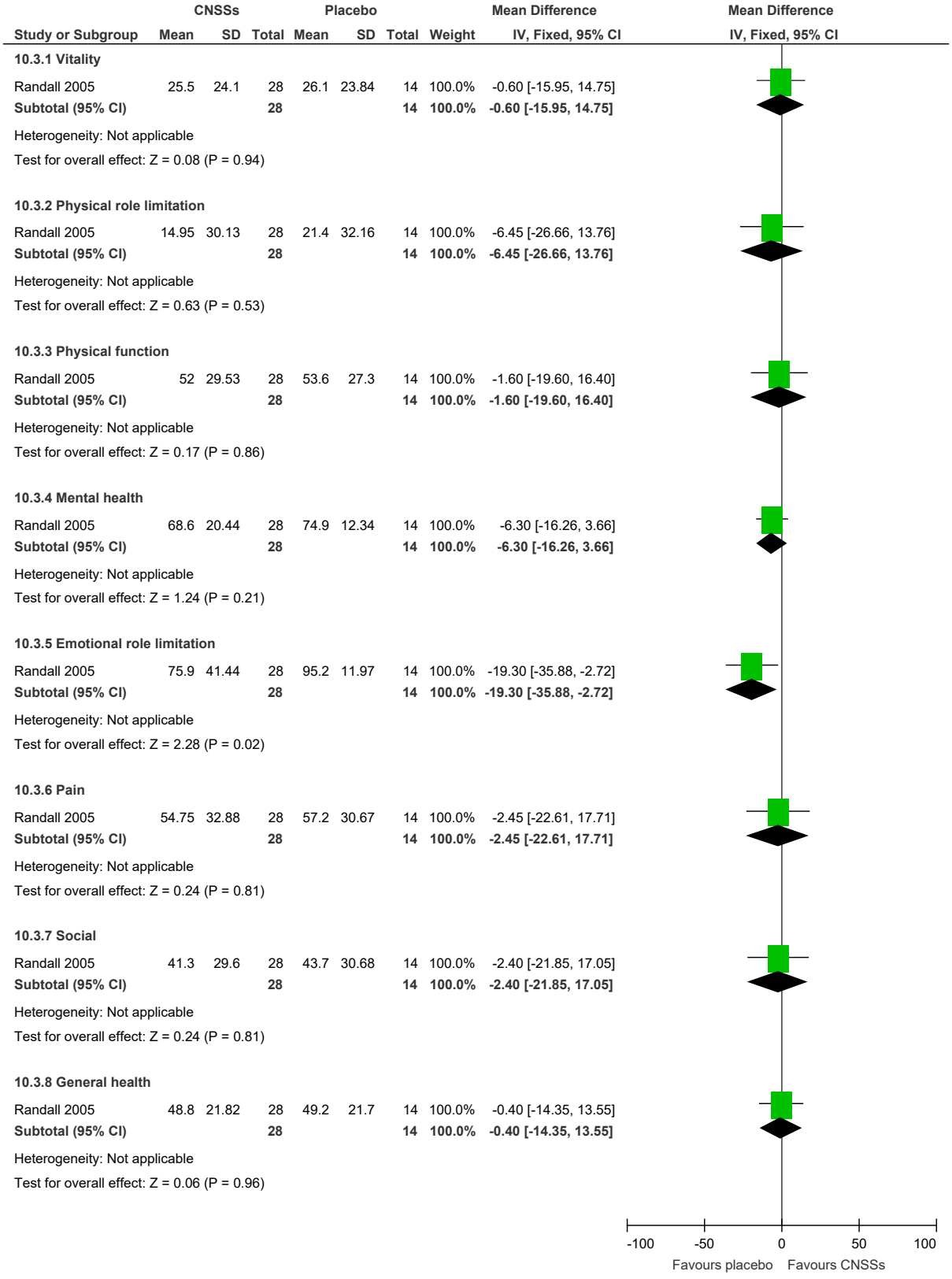


Figure 103: Fatigue: Checklist Individual Strength (CIS) total score at 4-12 weeks (methylphenidate)

Study or Subgroup	CNSSs Mean	CNSSs SD	CNSSs Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
Blockmans 2006	102.8	22.4	60	112.5	11.3	60	60.9%	-9.70 [-16.05, -3.35]
Montoya 2018	-16.9	23.52	63	-13.8	22.15	65	39.1%	-3.10 [-11.02, 4.82]
Total (95% CI)			**123**			**125**	**100.0%**	**-7.12 [-12.07, -2.16]**

Heterogeneity: Chi² = 1.62, df = 1 (P = 0.20); I² = 38%
Test for overall effect: Z = 2.82 (P = 0.005)

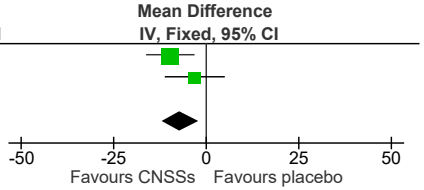


Figure 104: Fatigue: Fatigue Severity Scale at 6 weeks (dexamphetamine or lisdexamphetamine)

Study or Subgroup	CNSSs Mean	CNSSs SD	CNSSs Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Random, 95% CI
Olson 2003	-1.45	1.09	10	-0.03	1.11	10	56.9%	-1.42 [-2.38, -0.46]
Young 2013	-20.92	14.71	13	-5	11.73	11	43.1%	-15.92 [-26.50, -5.34]
Total (95% CI)			**23**			**21**	**100.0%**	**-7.67 [-21.75, 6.40]**

Heterogeneity: Tau² = 90.43; Chi² = 7.15, df = 1 (P = 0.007); I² = 86%
Test for overall effect: Z = 1.07 (P = 0.29)

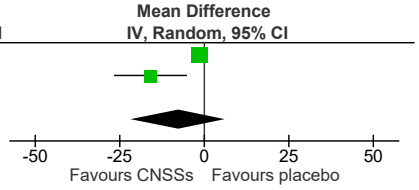


Figure 105: Fatigue: Chalder Physical Fatigue scale at 20 days (modafinil)

Study or Subgroup	CNSSs Mean	CNSSs SD	CNSSs Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
Randall 2005	13.35	5.987	28	13.6	7.85	14	100.0%	-0.25 [-4.92, 4.42]
Total (95% CI)			**28**			**14**	**100.0%**	**-0.25 [-4.92, 4.42]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.10 (P = 0.92)

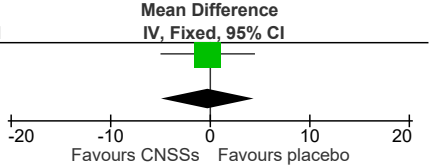


Figure 106: Fatigue: Chalder Mental Fatigue scale at 20 days (modafinil)

Study or Subgroup	CNSSs Mean	CNSSs SD	CNSSs Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
Randall 2005	7.8	3.14	28	7.4	2.99	14	100.0%	0.40 [-1.55, 2.35]
Total (95% CI)			**28**			**14**	**100.0%**	**0.40 [-1.55, 2.35]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.40 (P = 0.69)

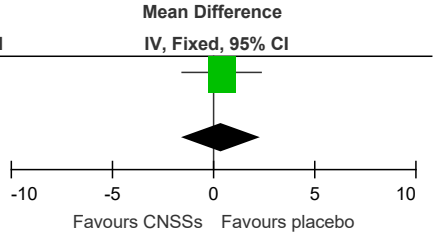


Figure 107: Sleep quality: sleep latency (time taken to fall asleep in mins) at 6 weeks (dexamphetamine)

Study or Subgroup	CNSSs Mean	CNSSs SD	CNSSs Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
Olson 2003	13	5.45	10	11.8	3.77	10	100.0%	1.20 [-2.91, 5.31]
Total (95% CI)			**10**			**10**	**100.0%**	**1.20 [-2.91, 5.31]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.57 (P = 0.57)

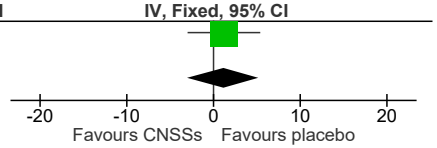


Figure 108: Psychological status: Hospital anxiety and depression scale (HADS) at 4 weeks (methylphenidate)

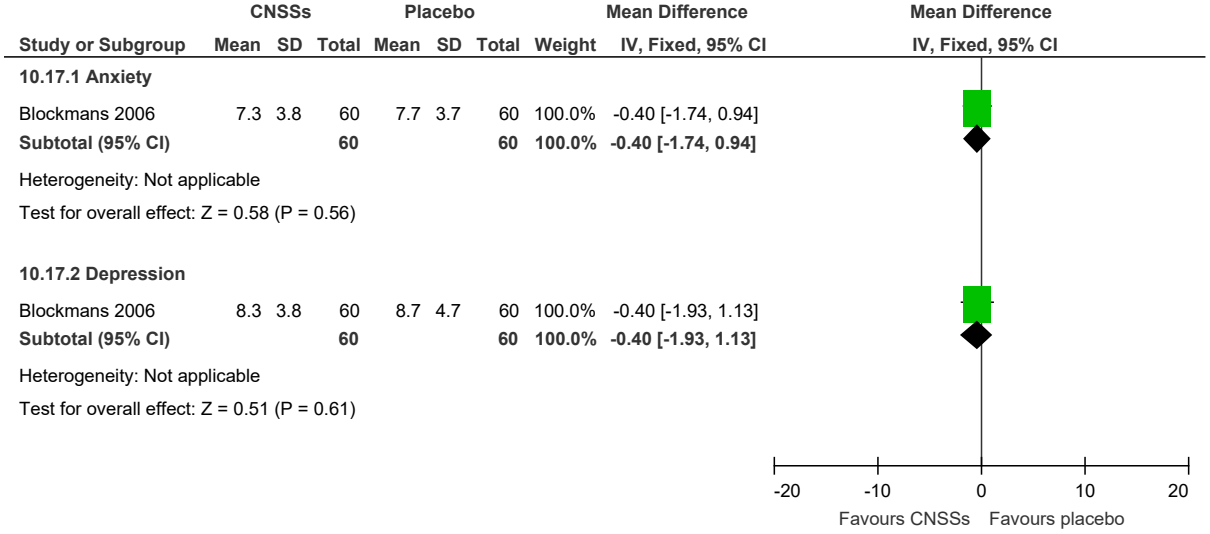


Figure 109: Psychological status: Hamilton Anxiety Scale (improvement) at 6 weeks (lisdexamphetamine)

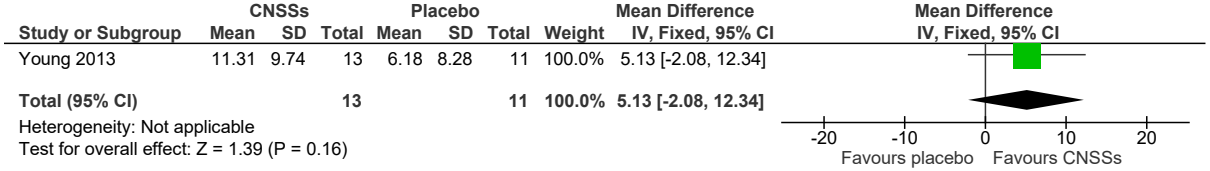


Figure 110: Adverse events: AEs leading to discontinuation (lisdexamphetamine)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Montoya 2018	8	63	3	65	83.8%	2.75 [0.76, 9.90]
Young 2013	2	15	0	11	16.2%	3.75 [0.20, 71.12]
Total (95% CI)		**78**		**76**	**100.0%**	**2.91 [0.90, 9.43]**
Total events	10		3			

Heterogeneity: Chi² = 0.04, df = 1 (P = 0.85); I² = 0%
Test for overall effect: Z = 1.78 (P = 0.07)

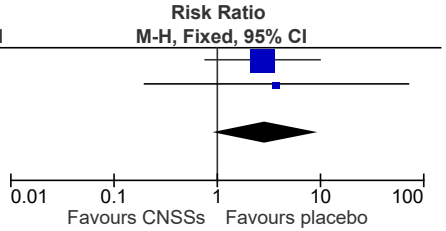


Figure 111: Adverse events: Serious AEs (pyelonephritis) (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Montoya 2018	1	63	0	65	100.0%	7.63 [0.15, 384.58]
Total (95% CI)		**63**		**65**	**100.0%**	**7.63 [0.15, 384.58]**
Total events	1		0			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.02 (P = 0.31)

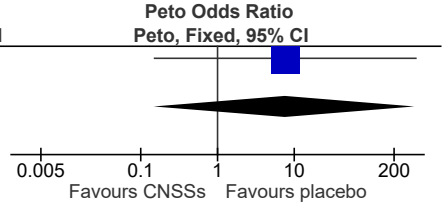


Figure 112: Adverse events: sleepiness (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	21	60	23	60	100.0%	0.91 [0.57, 1.46]
Total (95% CI)		**60**		**60**	**100.0%**	**0.91 [0.57, 1.46]**
Total events	21		23			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.38 (P = 0.71)

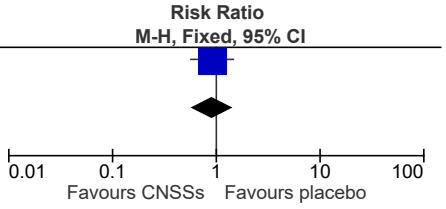


Figure 113: Adverse events: dry mouth (lisdexamphetamine)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	34	60	18	60	96.9%	1.89 [1.21, 2.95]
Young 2013	1	15	0	11	3.1%	2.25 [0.10, 50.54]
Total (95% CI)		**75**		**71**	**100.0%**	**1.90 [1.22, 2.96]**
Total events	35		18			

Heterogeneity: Chi² = 0.01, df = 1 (P = 0.91); I² = 0%
Test for overall effect: Z = 2.84 (P = 0.004)

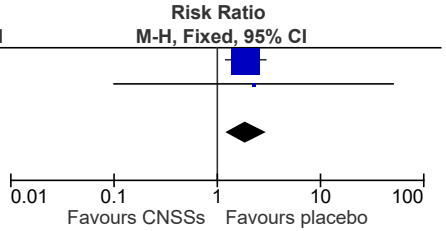


Figure 114: Adverse events: dizziness (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	30	60	38	60	100.0%	0.79 [0.57, 1.08]
Total (95% CI)		**60**		**60**	**100.0%**	**0.79 [0.57, 1.08]**
Total events	30		38			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.46 (P = 0.15)

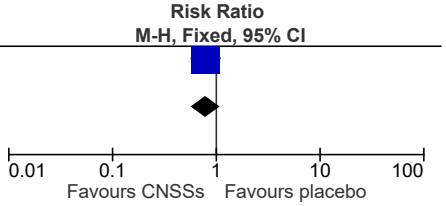


Figure 115: Adverse events: akathisia (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	29	60	34	60	100.0%	0.85 [0.61, 1.20]
Total (95% CI)		**60**		**60**	**100.0%**	**0.85 [0.61, 1.20]**
Total events	29		34			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.91 (P = 0.36)

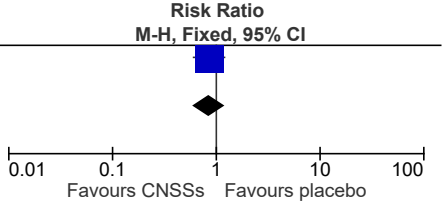


Figure 116: Adverse events: abdominal pain (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	28	60	23	60	100.0%	1.22 [0.80, 1.85]
Total (95% CI)		**60**		**60**	**100.0%**	**1.22 [0.80, 1.85]**
Total events	28		23			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.92 (P = 0.36)

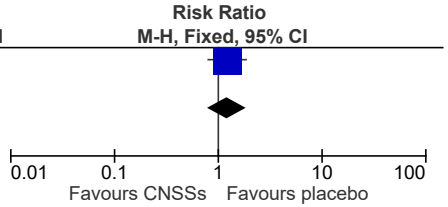


Figure 117: Adverse events: chest pain (methylphenidate)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blockmans 2006	17	60	25	60	100.0%	0.68 [0.41, 1.12]
Total (95% CI)		**60**		**60**	**100.0%**	**0.68 [0.41, 1.12]**
Total events	17		25			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.51 (P = 0.13)

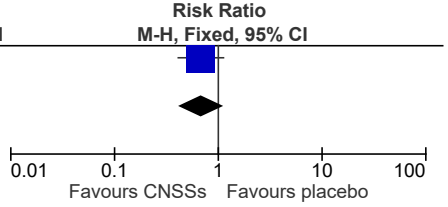


Figure 118: Adverse events: anorexia (lisdexamphetamine)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Olson 2003	5	10	1	10	100.0%	5.00 [0.70, 35.50]
Total (95% CI)		**10**		**10**	**100.0%**	**5.00 [0.70, 35.50]**
Total events	5		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.61 (P = 0.11)

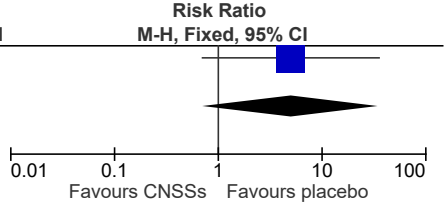


Figure 119: Adverse events: headache (lisdexamphetamine)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Young 2013	2	15	1	11	100.0%	1.47 [0.15, 14.21]
Total (95% CI)		**15**		**11**	**100.0%**	**1.47 [0.15, 14.21]**
Total events	2		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.33 (P = 0.74)

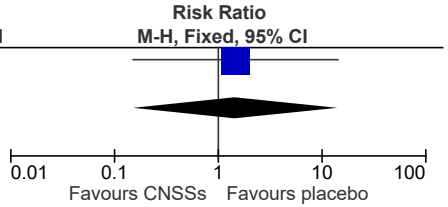


Figure 120: Adverse events: insomnia (lisdexamphetamine)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Young 2013	1	15	0	11	100.0%	5.66 [0.11, 299.01]
Total (95% CI)		**15**		**11**	**100.0%**	**5.66 [0.11, 299.01]**
Total events	1		0			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.86 (P = 0.39)

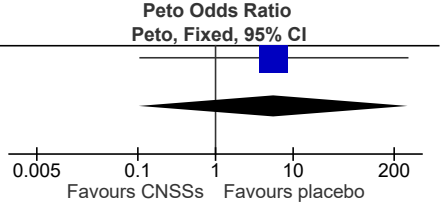


Figure 121: Adverse events (modafinil)

Study or Subgroup	CNSSs Events	CNSSs Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Randall 2005	21	28	8	14	100.0%	1.31 [0.79, 2.17]
Total (95% CI)		**28**		**14**	**100.0%**	**1.31 [0.79, 2.17]**
Total events	21		8			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.06 (P = 0.29)

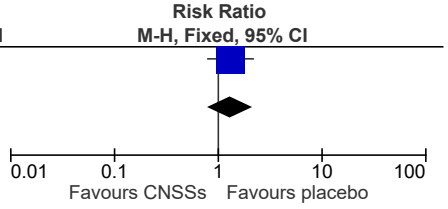


Figure 122: Cognitive function: Behaviour Rating Inventory of Executive Function (BRIEF), global executive composite improvement at 6 weeks (lisdexamphetamine)

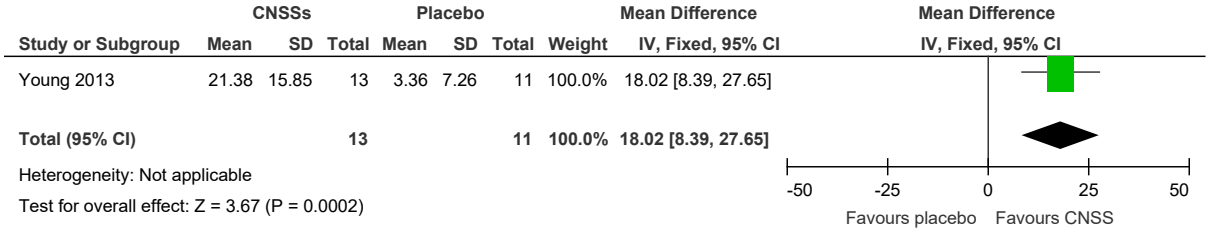


Figure 123: Pain: McGill pain Questionnaire (improvement) at 6 weeks (lisdexamphetamine)

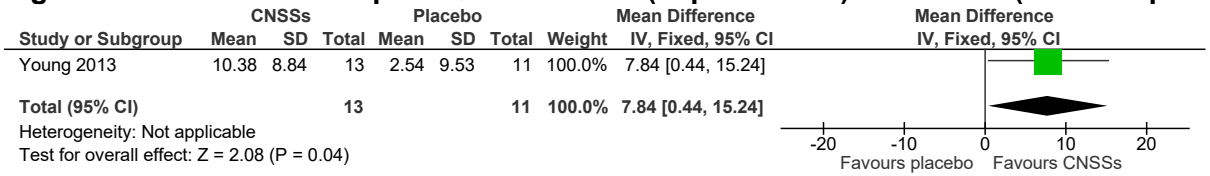
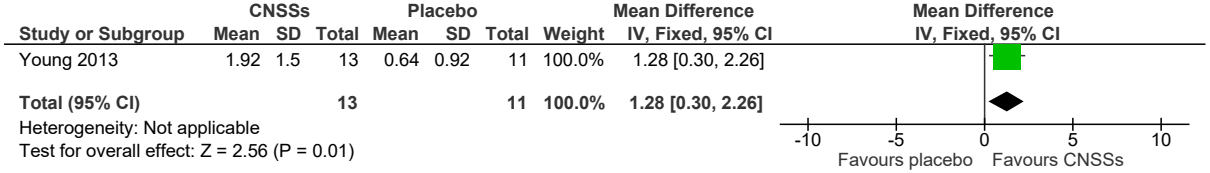


Figure 124: Symptom scales: Clinical Global Improvement – severity (improvement) at 6 weeks (lisdexamphetamine)



E.11 Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo

Figure 125: Fatigue: Multidimensional fatigue inventory (MFI-20) at 9 months (oral)

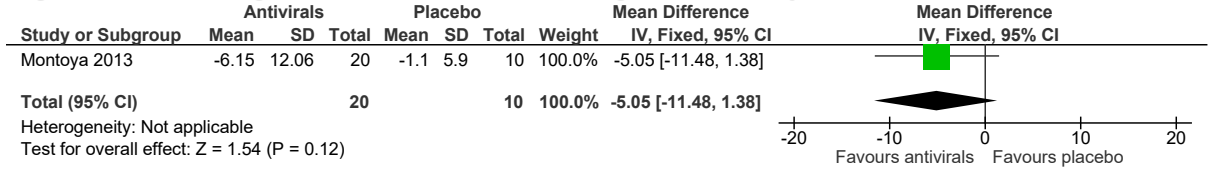


Figure 126: Fatigue: POMS fatigue at 37 days (IV)

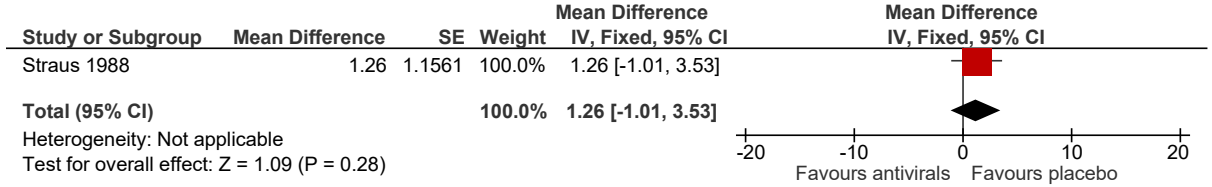


Figure 127: Fatigue: POMS vigour at 37 days (IV)

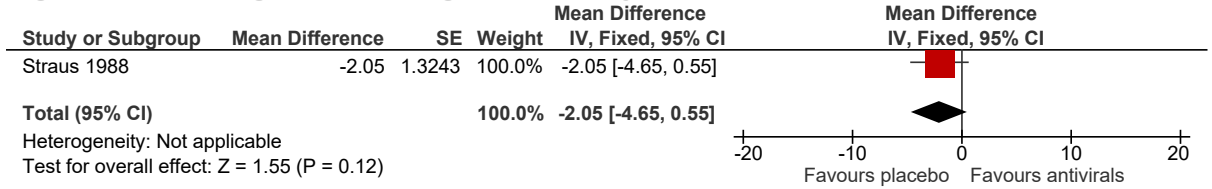


Figure 128: Psychological status: POMS anxiety at 37 days (IV)

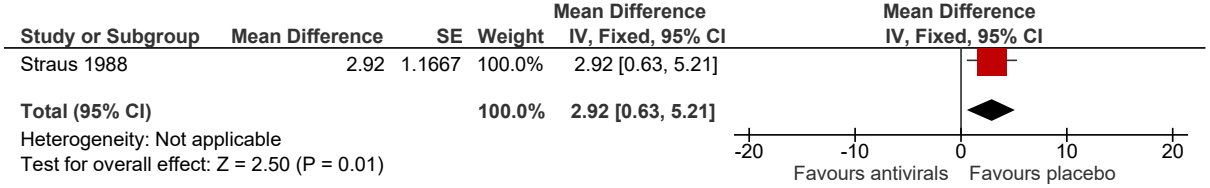


Figure 129: Psychological status: POMS depression at 37 days (IV)

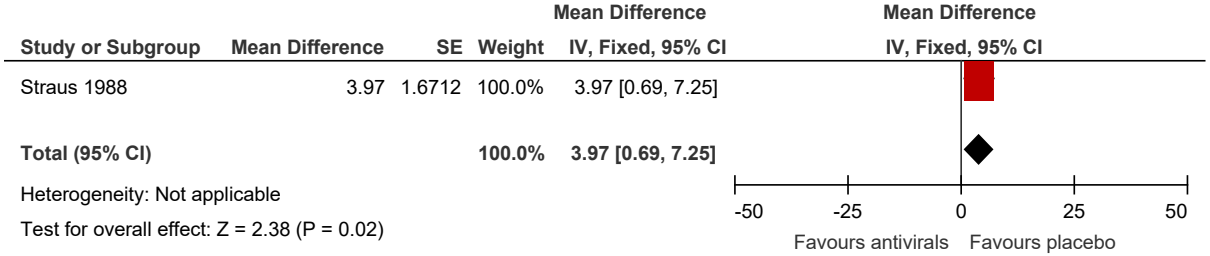


Figure 130: Psychological status: POMS anger at 37 days (IV)

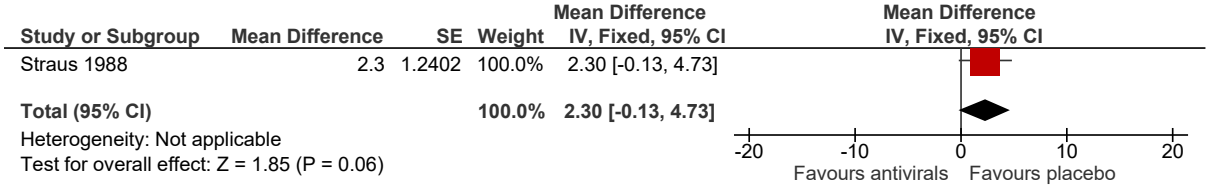


Figure 131: Psychological status: POMS confusion at 37 days (IV)

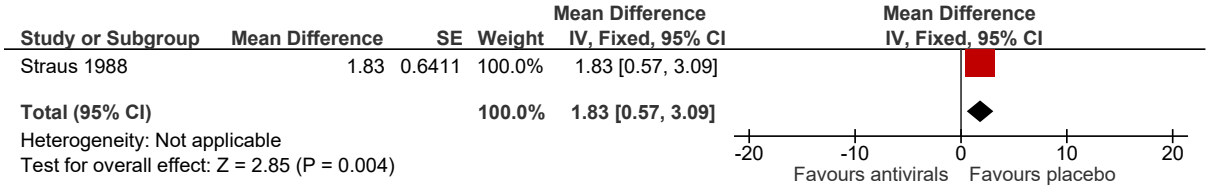


Figure 132: Adverse events: treatment-related adverse events (oral)

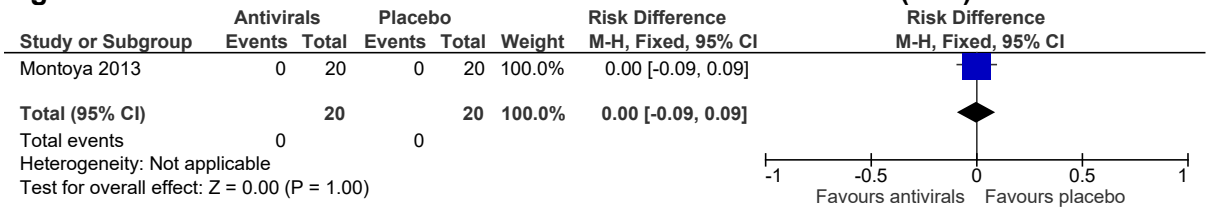


Figure 133: Adverse events: reversible renal failure (IV)

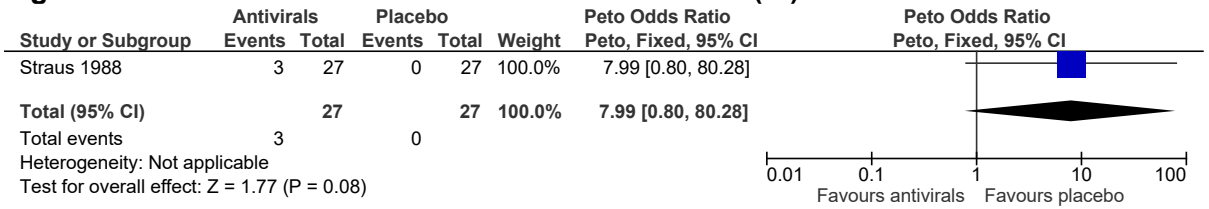


Figure 134: Activity levels: rest (hours/day) at 37 days (IV)

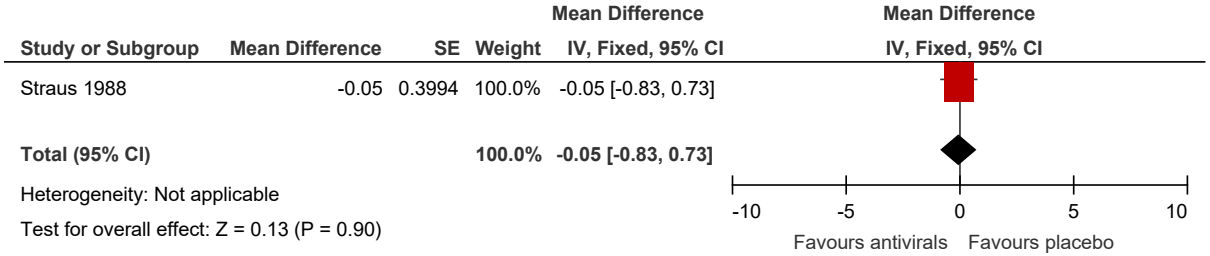
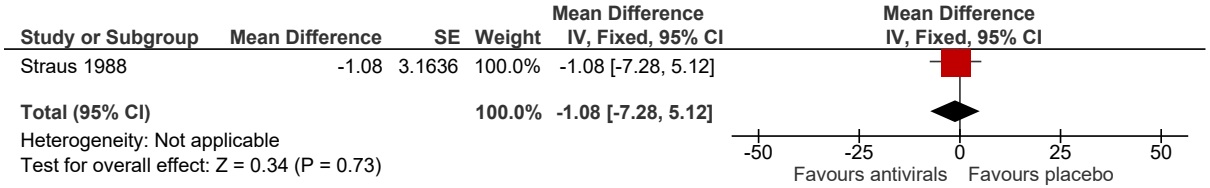


Figure 135: Symptom scales: Wellness score at 37 days (IV)



E.12 5-HT3 antagonists (ondansetron) versus placebo

Figure 136: Fatigue: Checklist Individual Strength fatigue at 12 weeks

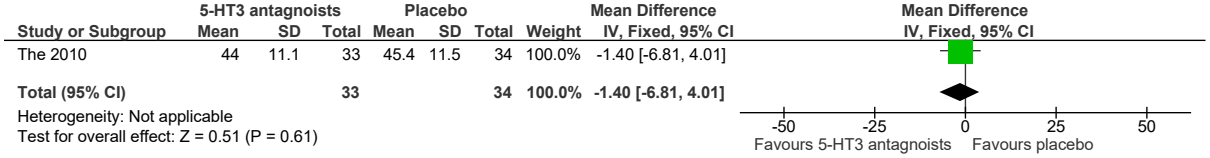


Figure 137: Activity levels: Actometer (objective accelerometer-based method of measuring activity) at 12 weeks

Study or Subgroup	5-HT3 antagnoists Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
The 2010	55	15.5	33	60.6	17.9	34	100.0%	-5.60 [-13.61, 2.41]
Total (95% CI)			**33**			**34**	**100.0%**	**-5.60 [-13.61, 2.41]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.37 (P = 0.17)

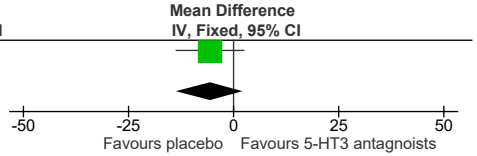


Figure 138: Adverse events: constipation

Study or Subgroup	5-HT3 antagnoists Events	Total	Placebo Events	Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
The 2010	2	33	0	34	100.0%	7.86 [0.48, 128.37]
Total (95% CI)		**33**		**34**	**100.0%**	**7.86 [0.48, 128.37]**
Total events	2		0			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.45 (P = 0.15)

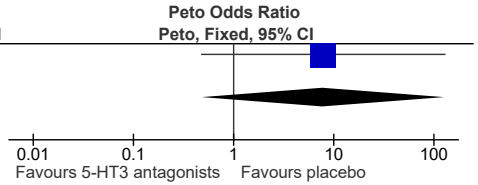


Figure 139: Adverse events: malaise

Study or Subgroup	5-HT3 antagnoists Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
The 2010	3	33	1	34	100.0%	3.09 [0.34, 28.23]
Total (95% CI)		**33**		**34**	**100.0%**	**3.09 [0.34, 28.23]**
Total events	3		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.00 (P = 0.32)

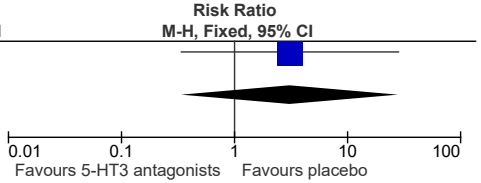
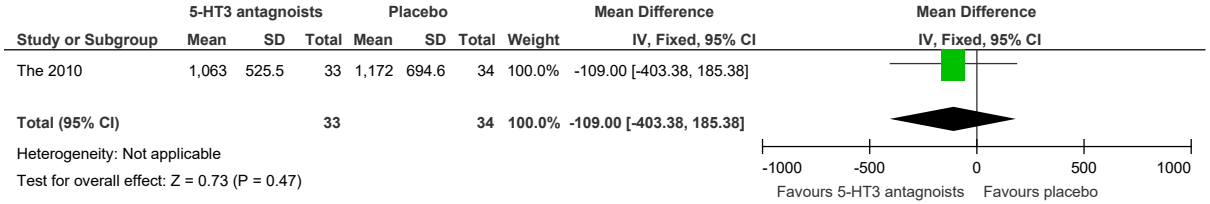


Figure 140: Symptom scales: Sickness Impact Profile (SIP) 8 at 12 weeks



E.13 Galantamine hydrobromide versus placebo

Figure 141: Fatigue: fatigue on VAS at 2 weeks

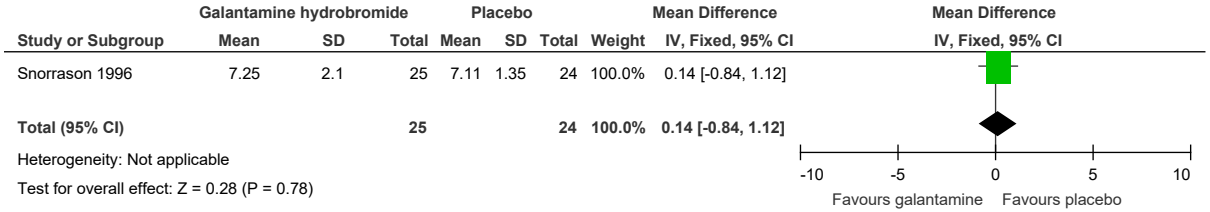


Figure 142: Cognitive function: memory on VAS at 2 weeks

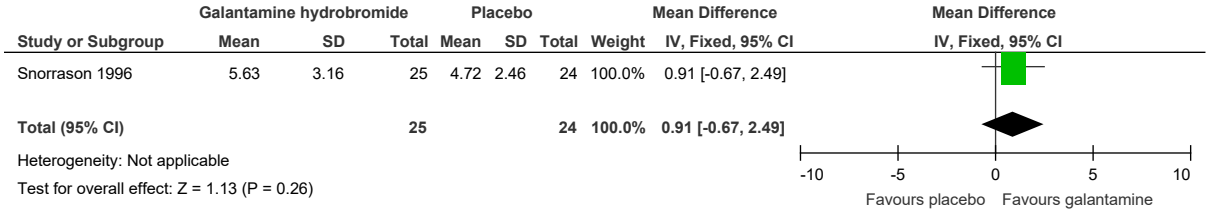


Figure 143: Pain: myalgia on VAS at 2 weeks

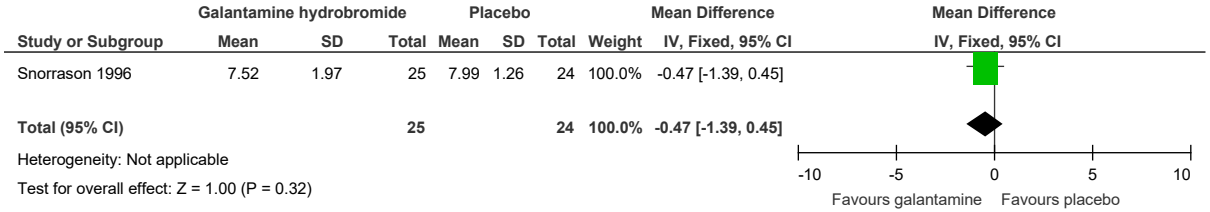


Figure 144: Sleep quality: sleep disturbance on VAS at 2 weeks

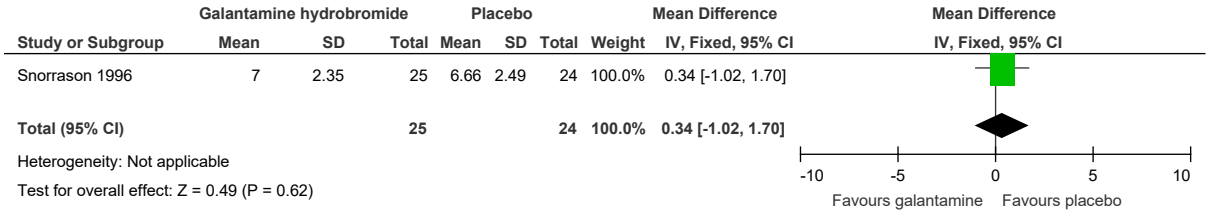


Figure 145: Adverse events: AEs dizziness on VAS at 2 weeks

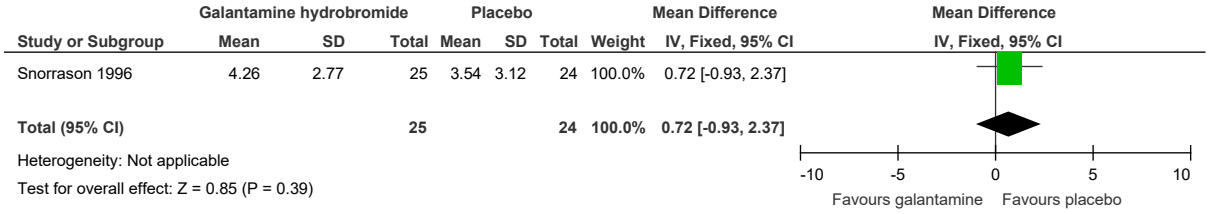


Figure 146: Return to school/work: work capacity/satisfaction on VAS at 2 weeks

Study or Subgroup	Galantamine hydrobromide Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Snorrason 1996	4.92	2.15	25	5.09	1.67	14	100.0%	-0.17 [-1.38, 1.04]
Total (95% CI)			**25**			**14**	**100.0%**	**-0.17 [-1.38, 1.04]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.27 (P = 0.78)

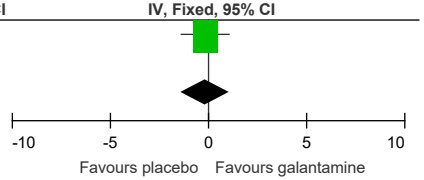
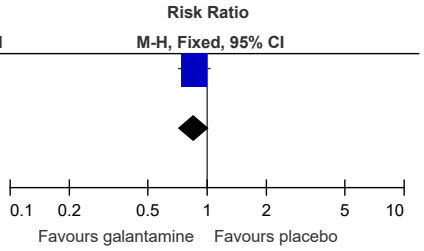


Figure 147: Symptom scales: clinical global impression score, no change or worse at 20 weeks

Study or Subgroup	Galantamine hydrobromide Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Blacker 2004	169	280	47	67	100.0%	0.86 [0.72, 1.03]
Total (95% CI)		**280**		**67**	**100.0%**	**0.86 [0.72, 1.03]**
Total events	169		47			

Heterogeneity: Not applicable

Test for overall effect: Z = 1.61 (P = 0.11)



E.14 Antihistamines (terfenadine) versus placebo

Figure 148: Physical functioning: modified Medical Outcome Study Short Form - physical functioning at 2 months

Study or Subgroup	Antihistamines Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Steinberg 1996	63.1	17.52	14	69.66	18.09	14	100.0%	-6.56 [-19.75, 6.63]
Total (95% CI)			**14**			**14**	**100.0%**	**-6.56 [-19.75, 6.63]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.97 (P = 0.33)

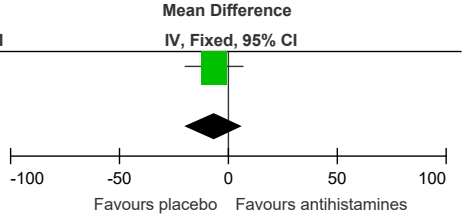
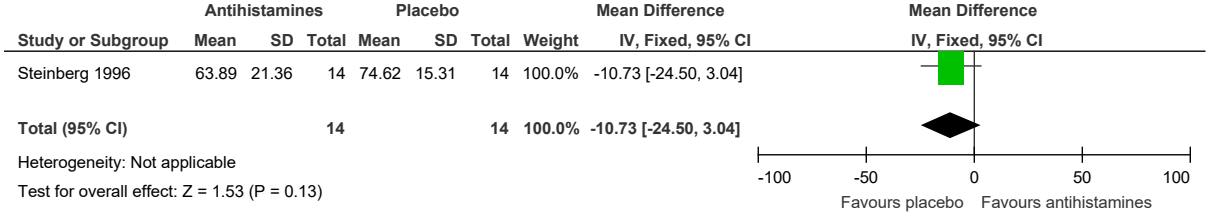


Figure 149: Psychological status: modified Medical Outcome Study Short Form - mental health at 2 months



E.15 Pro-inflammatory cytokine antagonists (anakinra) versus placebo

Figure 150: Mortality: death at 24 weeks

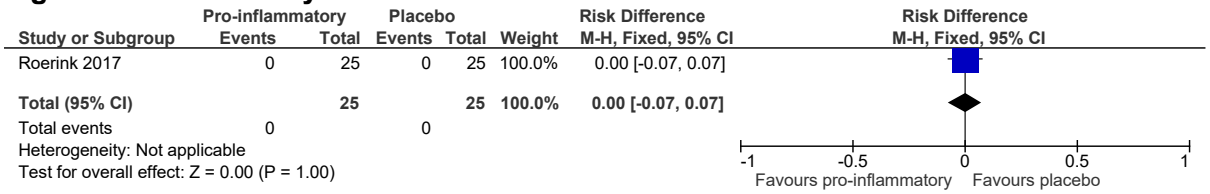


Figure 151: Fatigue: Checklist Individual Strength fatigue at 24 weeks

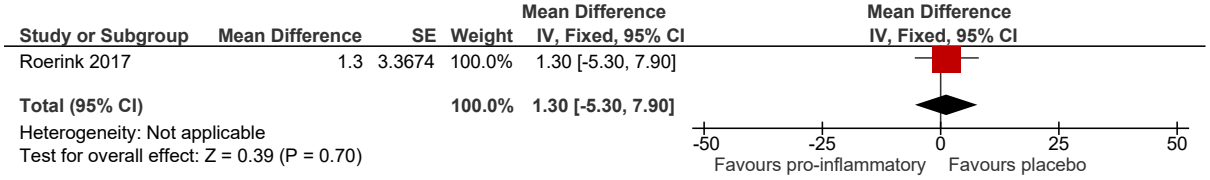


Figure 152: Physical functioning: SF36 physical function at 24 weeks

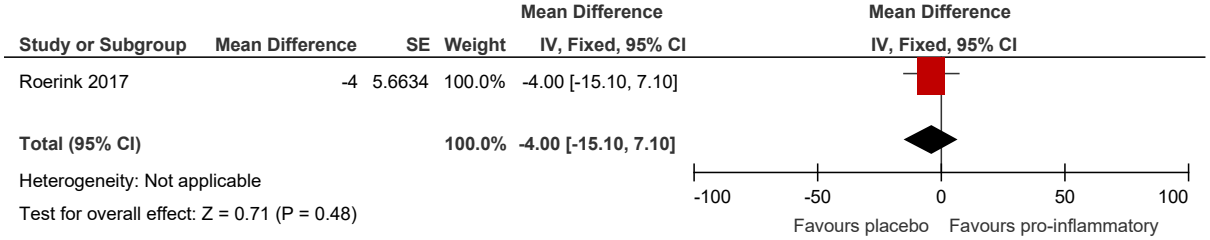


Figure 153: Psychological status: Symptom Checklist 90 at 24 weeks

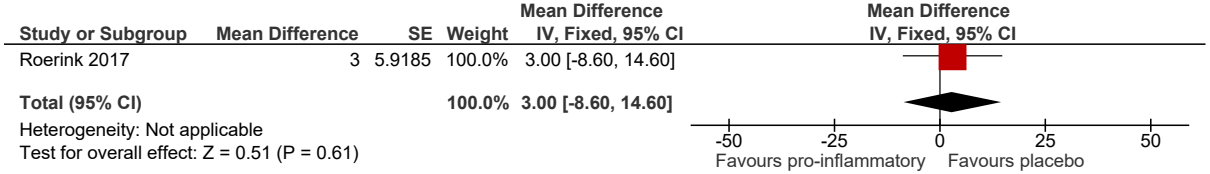


Figure 154: Pain: VAS maximum pain score at 24 weeks

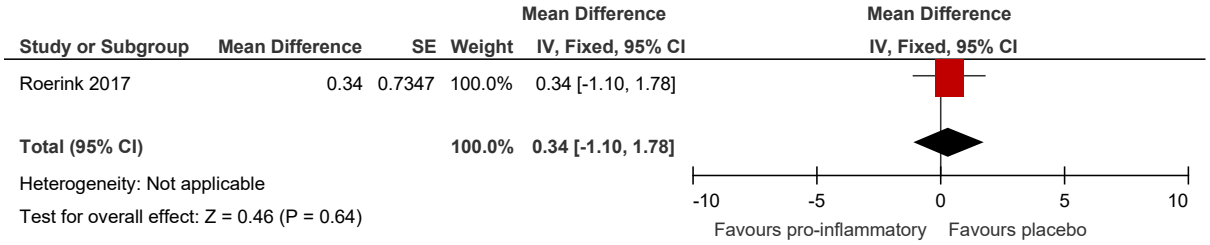


Figure 155: Adverse events

Study or Subgroup	Pro-inflammatory Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Roerink 2017	24	25	14	25	100.0%	1.71 [1.20, 2.45]
Total (95% CI)		**25**		**25**	**100.0%**	**1.71 [1.20, 2.45]**
Total events	24		14			

Heterogeneity: Not applicable

Test for overall effect: Z = 2.96 (P = 0.003)

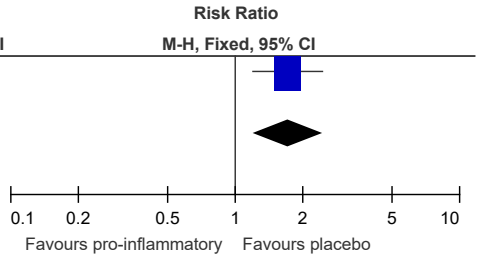


Figure 156: Adverse events: withdrawal due to adverse events

Study or Subgroup	Pro-inflammatory Events	Total	Placebo Events	Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Roerink 2017	1	25	0	25	100.0%	7.39 [0.15, 372.38]
Total (95% CI)		**25**		**25**	**100.0%**	**7.39 [0.15, 372.38]**
Total events	1		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 1.00 (P = 0.32)

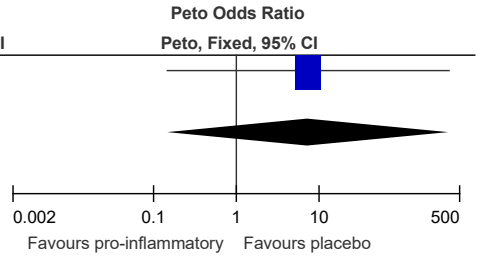
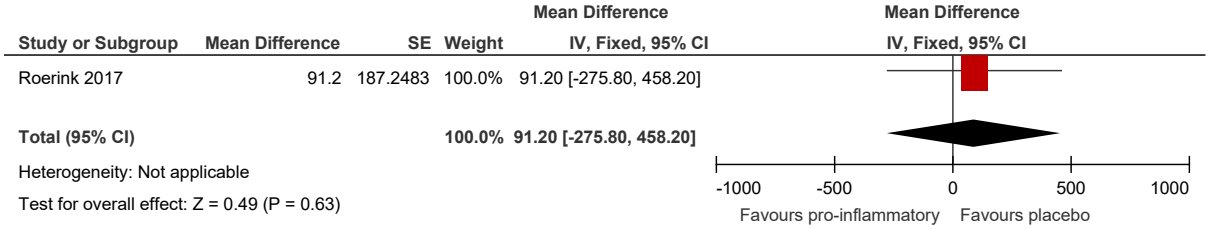


Figure 157: Symptom scales: Sickness Impact Profile at 24 weeks



E.16 Staphylococcus vaccine (Staphypan Berna) versus placebo

Figure 158: Pain: pain on VAS at 32 weeks

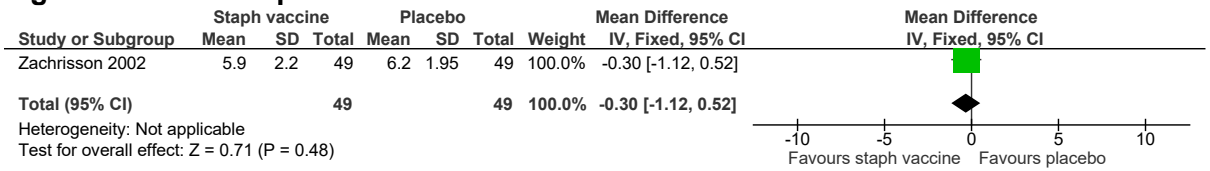


Figure 159: Adverse events

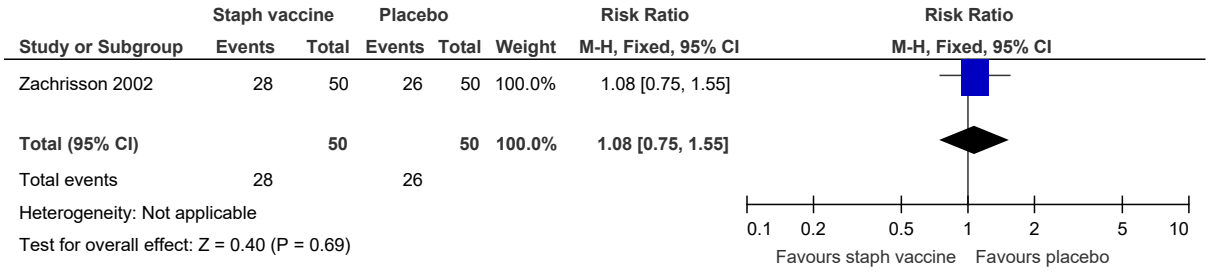
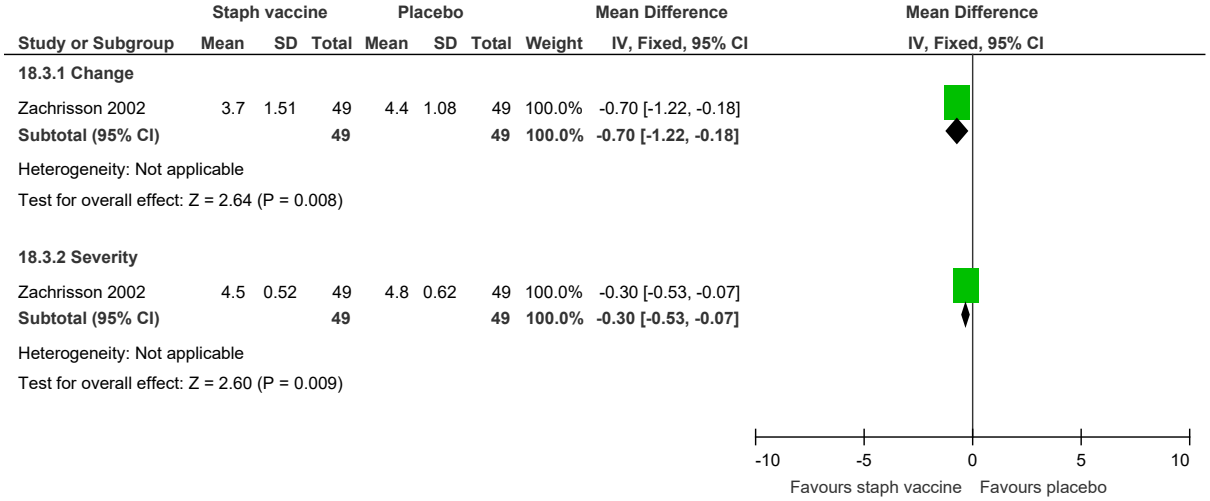


Figure 160: Symptom scales: clinical global impression at 32 weeks



E.17 Children and young people: Central antihypertensive drugs (clonidine) versus placebo

Figure 161: Fatigue: Chalder Fatigue Questionnaire (CFQ) total sum score at 30 weeks

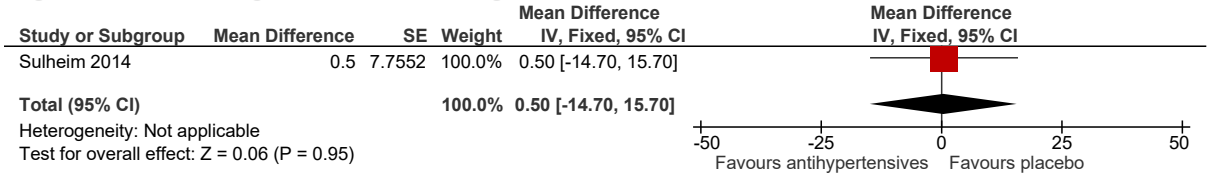


Figure 162: Physical functioning: Fatigue Disability Index (FDI) total sum score at 30 weeks

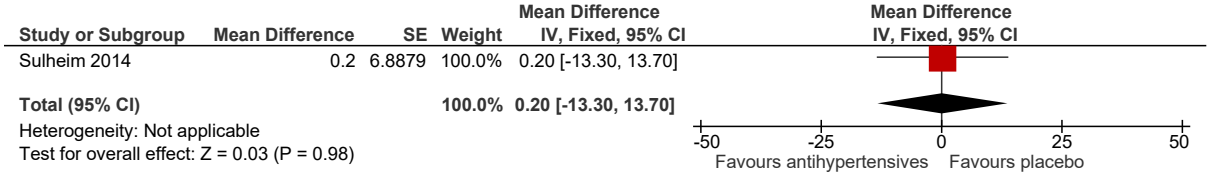


Figure 163: Pain: Brief Pain Inventory average pain score at 30 weeks

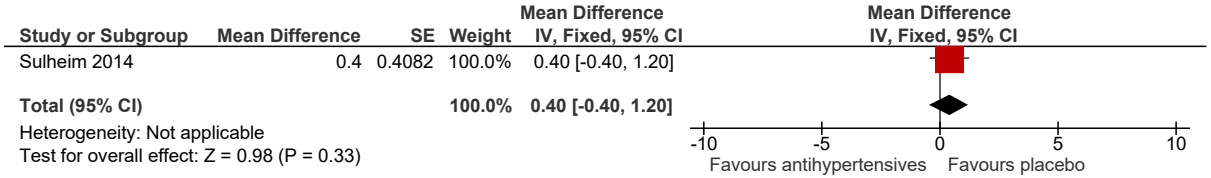


Figure 164: Sleep quality: Karolinska Sleep Questionnaire insomnia score at 30 weeks

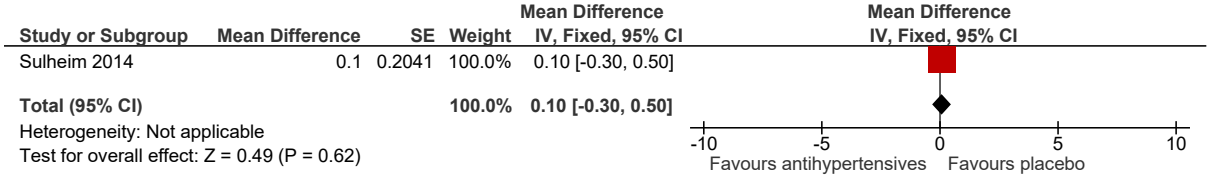


Figure 165: Adverse effects: various self-reported

Study or Subgroup	S/CA drugs Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Sulheim 2014	43	57	33	51	100.0%	1.17 [0.91, 1.50]
Total (95% CI)		**57**		**51**	**100.0%**	**1.17 [0.91, 1.50]**
Total events	43		33			

Heterogeneity: Not applicable

Test for overall effect: Z = 1.20 (P = 0.23)

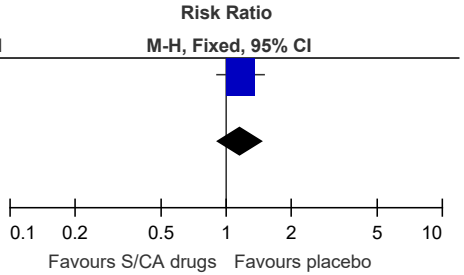


Figure 166: Activity levels: steps per day (accelerometer) at 30 weeks

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
Sulheim 2014	119	466.8453	100.0%	119.00 [-796.00, 1034.00]
Total (95% CI)			**100.0%**	**119.00 [-796.00, 1034.00]**

Heterogeneity: Not applicable

Test for overall effect: Z = 0.25 (P = 0.80)

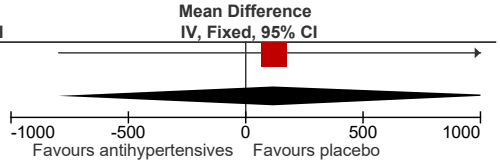


Figure 167: Cognitive function: Digit span backward test total at 30 weeks

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Fixed, 95% CI
Sulheim 2014	-0.5	0.3571	100.0%	-0.50 [-1.20, 0.20]
Total (95% CI)			**100.0%**	**-0.50 [-1.20, 0.20]**

Heterogeneity: Not applicable

Test for overall effect: Z = 1.40 (P = 0.16)

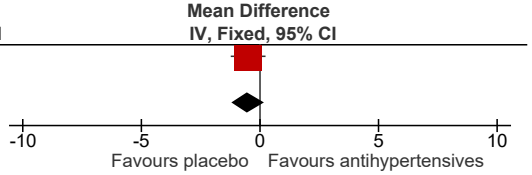
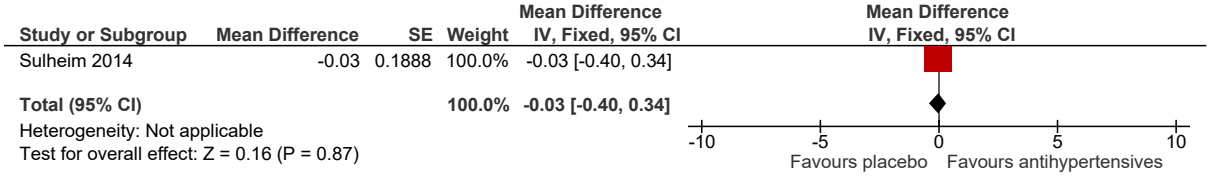


Figure 168: Symptom scales: CFS symptom inventory hypersensitivity score at 30 weeks



Appendix F GRADE and/or GRADE-CERQual tables

Table 22: Clinical evidence profile: Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immunomodulatory drugs (rituximab, rintatolimod, IV immunoglobulin G) versus placebo	Control	Relative (95% CI)	Absolute		
Quality of Life: SF36 physical composite (max % change from baseline) (follow-up 10 months; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	13 (rituximab)	15	-	MD 28 higher (1.56 to 54.44 higher)	⊕⊕OO LOW	CRITICAL
Quality of Life: SF36 mental composite (max % change from baseline) (follow-up 10 months; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	13 (rituximab)	15	-	MD 4 higher (29.52 lower to 37.52 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue/fatigability: Fatigue severity scale (follow up 18 months; range of scores: 9-63; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	77 (rituximab)	74	-	MD 0.07 lower (3.21 lower to 3.07 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Fatigue/fatigability: Fatigue numeric rating scale (follow up 16-20 months; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	77 (rituximab)	74	-	MD 0.06 lower (0.5 lower to 0.39 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Psychological status: Hamilton Depression Scale (follow-up 6 months; range of scores: 0-52; Better indicated by lower values)												
1	randomised trials	very serious[2]	no serious inconsistency	serious[1]	serious[2]	none	23 (IV immunoglobulin G)	26	-	MD 1 lower (3.35 lower to 1.35 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Zung Self-Rating Depression Scale (follow-up 6 months; range of scores: 0-80; Better indicated by lower values)												
1	randomised trials	very serious[2]	no serious inconsistency	serious[1]	serious[2]	none	23 (IV immunoglobulin G)	26	-	MD 1 higher (5.44 lower to 7.44 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: mental health on the Medical Outcome Study Short Form (follow-up 150 days; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	serious[2]	no serious inconsistency	serious[1]	very serious[2]	none	14 (IV immunoglobulin G)	14	-	MD 4.6 lower (16.07 lower to 6.87 higher)	⊕OOO VERY LOW	CRITICAL
Physical functioning: physical functioning on the Medical Outcome Study Short Form/SF36 (follow-up 150 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	serious[1]	very serious[2]	none	14 (IV immunoglobulin G)	14	-	MD 4.2 higher (12.62 lower to 21.02 higher)	⊕OOO VERY LOW	CRITICAL
Physical functioning: physical functioning on the Medical Outcome Study Short Form/SF36 (follow-up 24 months; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	77 (rituximab)	74	-	MD 1.24 higher (7.38 lower to 9.86 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Physical functioning: functional level percentage (follow up 16-20 months; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	77 (rituximab)	74	-	MD 0.68 lower (5.9 lower to 4.54 higher)	⊕⊕⊕O MODERATE	CRITICAL
Adverse events: Serious Adverse Events with possible/probable relation to intervention (follow-up 42 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	1/117 (0.85%) (rintatolimod)	2/117 (1.7%)	RR 0.5 (0.05 to 5.44)	9 fewer per 1000 (from 16 fewer to 76 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: major adverse events (follow-up 21 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	very serious[4]	very serious[2]	none	3/15 (20%) (IV immunoglobulin G)	3/15 (20%)	RR 1 (0.24 to 4.18)	0 fewer per 1000 (from 152 fewer to 636 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: constitutional symptoms (follow-up 3 months)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	56/73 (76.7%) (IV immunoglobulin G)	23/26 (88.5%)	RR 0.87 (0.72 to 1.05)	115 fewer per 1000 (from 248 fewer to 44 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: any serious adverse events with possible/probable relation to intervention (follow up 24 months)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	8/77 (10.4%) (rituximab)	0%	Peto OR 7.82 (1.89 to 32.35)	100 more per 1000 (from 30 more to 180 more)	⊕⊕⊕⊕ HIGH	CRITICAL
Adverse events: any adverse events of at least moderate severity with possible/probable relation to intervention (follow up 24 months)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	26/77 (33.8%) (rituximab)	12/74 (16.2%)	RR 2.08 (1.14 to 3.81)	175 more per 1000 (from 23 more to 456 more)	⊕⊕⊕O MODERATE	CRITICAL
Adverse events: suspected unexpected adverse reactions (follow up 24 months)												

1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	2/77 (2.6%) (rituximab)	1/74 (1.4%)	RR 1.92 (0.18 to 20.75)	12 more per 1000 (from 11 fewer to 267 more)	⊕⊕OO LOW	CRITICAL
Activity levels: mean number of steps per 24 hours (follow up 17-21 months; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	77 (rituximab)	74	-	MD 127 lower (1004 lower to 750 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Exercise performance measure: Treadmill exercise duration in seconds (follow-up 42 weeks; Better indicated by higher values)												
1	randomised trials	serious[3]	no serious inconsistency	serious[1]	serious[2]	none	100 (rintatolimod)	108	-	MD 56 higher (25.94 lower to 137.94 higher)	⊕OOO VERY LOW	CRITICAL
Return to school or work: Resumption of pre-morbid employment status (full-time) (follow-up 6 months)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	no serious imprecision	none	6/23 (26.1%) (IV immunoglobulin G)	0/26 (0%)	Peto OR 10.79 (1.98 to 58.68)	260 more per 1000 (from 80 more to 450 more)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Marked reduction in symptoms and improvement in functional capacity (follow-up 6 months)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	10/23 (43.5%) (IV immunoglobulin G)	3/26 (11.5%)	RR 3.77 (1.18 to 12.04)	320 more per 1000 (from 21 more to 1000 more)	⊕OOO VERY LOW	CRITICAL

[1] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[3] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[4] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature. Further downgraded for outcome indirectness (unclear if major adverse events were treatment-related)

Table 23: Clinical evidence profile: Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antidepressants (duloxetine, fluoxetine, moclobemide) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Quality of Life: SF36 vitality (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (duloxetine)	26	-	MD 3.3 higher (10.3 lower to 16.9 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF-36 physical functioning (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (duloxetine)	26	-	MD 6.8 higher (8.5 lower to 22.1 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF-36 role physical (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (duloxetine)	26	-	MD 11 higher (9 lower to 31 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 mental health (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (duloxetine)	26	-	MD 1.1 lower (11.8 lower to 9.6 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 role emotional (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (duloxetine)	26	-	MD 4.4 higher (24.2 lower to 33 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 bodily pain (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (duloxetine)	26	-	MD 11.4 higher (0.5 lower to 23.3 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 general health (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (duloxetine)	26	-	MD 0 higher (10.8 lower to 10.8 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 social functioning (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (duloxetine)	26	-	MD 0.7 higher (14.7 lower to 16.1 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: 14-item Chalder fatigue scale at 26 weeks (follow-up 26 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[2]	serious[3]	none	35 (fluoxetine)	34	-	MD 0.3 lower (4.06 lower to 3.46 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: MFI-20 general fatigue (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 1 lower (2.8 lower to 0.8 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: MFI-20 physical fatigue (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.9 lower (2.7 lower to 0.9 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: MFI-20 reduced activity (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	27 (duloxetine)	30	-	MD 0 higher (1.8 lower to 1.8 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: MFI-20 reduced motivation (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.8 lower (2.6 lower to 1 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: MFI-20 mental fatigue (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 2.5 lower (4.4 to 0.6 lower)	⊕OOO VERY LOW	CRITICAL
Fatigue: Checklist Individual Strength (CIS) fatigue (follow-up 16 weeks; range of scores: 8-56; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	52 (fluoxetine)	45	-	MD 0.16 lower (0.64 lower to 0.31 higher)	⊕OOO VERY LOW	CRITICAL
Physical functioning: Karnofsky Performance Index (measured in units of standard deviation at baseline) (follow-up 6 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	40 (moclobemide)	37	-	MD 0.28 higher (0.28 lower to 0.84 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Profile of mood states (POMS) fatigue (follow-up 6 weeks; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	40 (moclobemide)	37	-	MD 0.04 lower (0.2 lower to 0.12 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Profile of mood states (POMS) vigour (follow-up 6 weeks; range of scores: 0-32; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	40 (moclobemide)	37	-	MD 0.51 higher (0 to 1.02 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Profile of mood states (POMS) depression (follow-up 6 weeks; range of scores: 0-60; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	40 (moclobemide)	37	-	MD 0.02 higher (0.36 lower to 0.4 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS depression change scores (follow-up 12-26 weeks; range of scores: 0-21; Better indicated by lower values)												
2	randomised trials	very serious[1]	serious[4]	serious[2]	serious[3]	none	62 (fluoxetine or duloxetine)	64	-	MD 0.51 higher (0.72 lower to 1.74 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS anxiety (follow-up 12 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.9 lower (2.4 lower to 0.6 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Beck Depression Inventory (follow-up 16 weeks; range of scores: 0-63; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	52 (fluoxetine)	45	-	MD 0.19 lower (0.35 to 0.02 lower)	⊕OOO VERY LOW	CRITICAL
Pain: Brief Pain Inventory severity (follow-up 12 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.73 lower (1 to 0.46 lower)	⊕OOO VERY LOW	CRITICAL
Pain: Brief Pain Inventory interference (follow-up 12 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	27 (duloxetine)	30	-	MD 0.7 lower (0.96 to 0.44 lower)	⊕OOO VERY LOW	CRITICAL
Adverse events: tremor (follow-up 16 weeks)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	18/45 (40%) (fluoxetine)	13/51 (25.5%)	RR 1.57 (0.87 to 2.83)	145 more per 1000 (from 33 fewer to 466 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: perspiration (follow-up 16 weeks)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	30/45 (66.7%) (fluoxetine)	20/51 (39.2%)	RR 1.7 (1.14 to 2.53)	275 more per 1000 (from 55 more to 600 more)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: VO2 max (mL O2/kg/min) (follow-up 26 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35 (fluoxetine)	34	-	MD 1.1 higher (1.43 lower to 3.63 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Clinical Global Impression of Severity (follow-up 12 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.1 lower (0.3 lower to 0.1 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Clinical Global Impression of Improvement (follow-up 12 weeks; range of scores: 1-7; Better indicated by lower values)												

1	randomised trials	very serious[3]	no serious inconsistency	serious[2]	serious[3]	none	27 (duloxetine)	30	-	MD 0.8 lower (1.7 lower to 0.1 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: CDC symptom inventory (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[2]	serious[3]	none	20 (duloxetine)	26	-	MD 2.7 lower (15.5 lower to 10.1 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Improvement of symptoms (patient-reported) (follow-up 6-16 weeks)												
2	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	32/92 (34.8%) (fluoxetine or moclobemide)	19/94 (20.2%)	RR 1.63 (1.02 to 2.59)	127 more per 1000 (from 4 more to 321 more)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Worsening of symptoms (patient-reported) (follow-up 16 weeks)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	7/45 (15.6%) (fluoxetine)	12/51 (23.5%)	RR 0.66 (0.28 to 1.53)	80 fewer per 1000 (from 169 fewer to 125 more)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[4] Downgraded for inconsistency. I^2=63%

Table 24: Clinical evidence profile: Antidepressants (fluoxetine) versus graded exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antidepressants (fluoxetine) versus graded exercise**	**Control**	**Relative (95% CI)**	**Absolute**		
Fatigue: 14-item Chalder fatigue scale (follow-up 26 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35	34	-	MD 2.7 higher (1.85 lower to 7.25 higher)	⊕OOO VERY LOW	CRITICAL

Psychological status: HADS depression (follow-up 26 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35	34	-	MD 0.5 lower (2.27 lower to 1.27 higher)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: VO2 max (mL O2/kg/min) (follow-up 26 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35	34	-	MD 1.8 lower (4.53 lower to 0.93 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 25: Clinical evidence profile: Antidepressants (fluoxetine) versus combined antidepressants (fluoxetine) & graded exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antidepressants (fluoxetine) versus combined antidepressants (fluoxetine) & graded exercise**	**Control**	**Relative (95% CI)**	**Absolute**		
Fatigue: 14-item Chalder fatigue scale (follow-up 26 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35	33	-	MD 3 higher (1.47 lower to 7.47 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS depression (follow-up 26 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	35	34	-	MD 0.3 higher (1.51 lower to 2.11 higher)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: VO2 max (mL O2/kg/min) (follow-up 26 weeks; Better indicated by higher values)												
1	randomised trials	very serious[2]	no serious inconsistency	serious[2]	serious[3]	none	35	33	-	MD 1 lower (3.41 lower to 1.41 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 26: Clinical evidence profile: Combined antidepressants (fluoxetine) & graded exercise versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined antidepressants (fluoxetine) & graded exercise versus placebo	Control	Relative (95% CI)	Absolute		
Fatigue: 14-item Chalder fatigue scale (follow-up 26 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 3.3 lower (7.71 lower to 1.11 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS depression (follow-up 26 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 0.7 lower (2.28 lower to 0.88 higher)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: VO2 max (mL O2/kg/min) (follow-up 26 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 2.1 higher (0.08 lower to 4.28 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 27: Clinical evidence profile: Combined antidepressants (fluoxetine) & graded exercise versus graded exercise

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined antidepressants (fluoxetine) & graded exercise versus graded exercise	Control	Relative (95% CI)	Absolute		
Fatigue: 14-item Chalder fatigue scale (follow-up 26 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	none	none	33	34	-	MD 0.3 lower (5.41 lower to 4.81 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS depression (follow-up 26 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 0.8 lower (2.52 lower to 0.92 higher)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: VO2 max (mL O2/kg/min) (follow-up 26 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 0.8 lower (3.21 lower to 1.61 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 28: Clinical evidence profile: Antidepressants (fluoxetine) versus antipsychotics (amisulpride)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Antidepressants (fluoxetine) versus antipsychotics (amisulpride)	Control	Relative (95% CI)	Absolute		
Quality of Life: SF12 (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	20	20	-	MD 15.6 lower (18.61 to 12.59 lower)	⊕OOO VERY LOW	CRITICAL
Fatigue: Fatigue Severity Scale (follow-up 12 weeks; range of scores: 9-63; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	20	20	-	MD 12.6 higher (8.26 to 16.94 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS anxiety (follow-up 12 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20	20	-	MD 0.4 higher (0.22 lower to 1.02 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS depression (follow-up 12 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20	20	-	MD 0.1 lower (0.69 lower to 0.49 higher)	⊕OOO VERY LOW	CRITICAL
Pain: pain on VAS (follow-up 12 weeks; range of scores: 0-100; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	20	20	-	MD 12.6 higher (5.8 to 19.4 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: FIBSER global burden (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20	20	-	MD 0.2 lower (0.67 lower to 0.27 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Clinical Global Impression Severity (CGI-S) (follow-up 12 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	20	20	-	MD 1.3 higher (0.75 to 1.85 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 29: Clinical evidence profile: Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Corticosteroids (oral hydrocortisone or fludrocortisone, nasal flunisolide) versus placebo	Control	Relative (95% CI)	Absolute		
Quality of Life: SF36 physical (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 7.6 higher (5.36 lower to 20.56 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 energy or fatigue (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 2.1 higher (7.43 lower to 11.63 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 emotional wellbeing (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 3.8 higher (5.29 lower to 12.89 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 role emotional (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0 higher (14.96 lower to 14.96 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 role physical (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 11.8 lower (29.09 lower to 5.49 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 pain (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.6 lower (15.29 lower to 14.09 higher)	⊕OOO VERY LOW	CRITICAL
Quality of life: SF36 social (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 1.9 higher (11.06 lower to 14.86 higher)	⊕OOO VERY LOW	CRITICAL
Quality of life: SF36 general wellbeing (follow-up 6 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 3.7 lower (12.54 lower to 5.14 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: fatigue on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0 higher (1.1 lower to 1.1 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Chronic Fatigue Syndrome Severity Rating (follow-up 4-8 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	very serious[4]	Serious[3]	none	28 (nasal flunisolide)	28	-	MD 3.17 lower (7.48 lower to 1.14 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Profile of Mood States - fatigue (follow-up 11 weeks; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	no serious imprecision	none	38 (fludrocortisone)	45	-	MD 0.2 lower (3.47 lower to 3.07 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Profile of Mood States - fatigue (follow-up 12 weeks; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	34 (hydrocortisone)	34	-	MD 1.8 lower (4.14 lower to 0.54 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Profile of Mood States - vigour (follow-up 11 weeks; range of scores: 0-32; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	no serious imprecision	none	38 (fludrocortisone)	45	-	MD 0.2 higher (2.56 lower to 2.96 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Profile of Mood States - vigour (follow-up 12 weeks; range of scores: 0-32; Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	34 (hydrocortisone)	34	-	MD 0.5 higher (1.07 lower to 2.07 higher)	⊕⊕OO LOW	CRITICAL
Fatigue: Wood Mental Fatigue Inventory (follow-up 11 weeks; range of scores: 0-36; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 0.8 higher (3.66 lower to 5.26 higher)	⊕OOO VERY LOW	CRITICAL
Physical function: SF36 physical function (follow-up 11 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 7.5 higher (3.2 lower to 18.2 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: SF36 mental health (follow-up 11 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 1.2 lower (8.92 lower to 6.52 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: adverse events leading to study withdrawal (follow-up 6 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	0/20 (0%) (fludrocortisone)	2/20 (10%)	Peto OR 0.13 (0.01 to 2.13)	100 fewer per 1000 (from 250 fewer to 50 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: adverse effects / adverse events (follow-up 6-11 weeks)												
2	randomised trials	very serious[1]	no serious inconsistency	serious[5]	no serious imprecision	none	27/58 (46.6%) (fludrocortisone)	36/65 (55.4%)	RR 0.86 (0.63 to 1.17)	78 fewer per 1000 (from 205 fewer to 94 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: any adverse reaction (follow-up 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	31/35 (88.6%) (hydrocortisone)	27/35 (77.1%)	RR 1.15 (0.93 to 1.43)	116 more per 1000 (from 54 fewer to 332 more)	⊕OOO VERY LOW	CRITICAL
Psychological status: Beck Depression Inventory (follow-up 11weeks; range of scores: 0-63; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 0.4 lower (3.43 lower to 2.63 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Beck Depression Inventory (follow-up 12 weeks; range of scores: 0-63; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	34 (hydrocortisone)	34	-	MD 1.7 lower (3.9 lower to 0.5 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Profile of Mood States - anger (follow-up 12 weeks; range of scores: 0-48; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	34 (hydrocortisone)	34	-	MD 0.8 lower (2.63 lower to 1.03 higher)	⊕⊕OO LOW	CRITICAL
Psychological status: Profile of Mood States - anxiety (follow-up 12 weeks; range of scores: 0-36; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	34 (hydrocortisone)	34	-	MD 1.3 higher (0.17 lower to 2.77 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Profile of Mood States - confusion (follow-up 12 weeks; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	34 (hydrocortisone)	34	-	MD 0.3 higher (1.18 lower to 1.78 higher)	⊕⊕OO LOW	CRITICAL
Psychological status: Profile of Mood States - depression (follow-up 12 weeks; range of scores: 0-60; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[1]	none	34 (hydrocortisone)	34	-	MD 1.6 lower (3.61 lower to 0.41 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Symptom checklist-90-R general sensitivity index (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	34 (hydrocortisone)	34	-	MD 0 higher (0.1 lower to 0.1 higher)	⊕⊕OO LOW	CRITICAL
Psychological status: Symptom checklist-90-R positive symptom distress index (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	very serious[2]	serious[3]	none	34 (hydrocortisone)	34	-	MD 0.1 higher (0.04 lower to 0.24 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Symptom checklist-90-R positive symptom total (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	34 (hydrocortisone)	34	-	MD 0.2 lower (5.5 lower to 5.1 higher)	⊕⊕OO LOW	CRITICAL
Psychological status: Hamilton Depression Rating Scale (follow-up 12 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	32 (hydrocortisone)	33	-	MD 0.9 lower (2.55 lower to 0.75 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: Positive and negative effect scale (PANAS) positive affect (follow-up 6 weeks; range of scores: 10-50; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 1 higher (3.67 lower to 5.67 higher)	⊕OOO VERY LOW	CRITICAL
Activity levels: activity scale (follow-up 12 weeks; range of scores: not reported; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	34 (hydrocortisone)	34	-	MD 0.4 lower (1 lower to 0.2 higher)	⊕OOO VERY LOW	CRITICAL
Activity levels: distance before exhausted (ordinal scale) (follow-up 6 weeks; range of scores: 1-5; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0 higher (0.72 lower to 0.72 higher)	⊕OOO VERY LOW	CRITICAL
Activity levels: Duke Activity Status Index (follow-up 11 weeks; range of scores: 0-58.2; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 2.5 higher (1.49 lower to 6.49 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Reaction time (secs) (follow-up 6 weeks; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.01 lower (0.06 lower to 0.04 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: inability to concentrate on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 0.6 lower (2.18 lower to 0.98 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: forgetfulness on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 0.9 lower (2.45 lower to 0.65 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: confusion on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.1 lower (1.68 lower to 1.48 higher)	⊕OOO VERY LOW	CRITICAL
Pain: muscle pain on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.1 lower (1.82 lower to 1.62 higher)	⊕OOO VERY LOW	CRITICAL
Pain: joint pain on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.3 lower (2.39 lower to 1.79 higher)	⊕OOO VERY LOW	CRITICAL
Sleep quality: unrefreshing sleep on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 0.5 lower (1.68 lower to 0.68 higher)	⊕OOO VERY LOW	CRITICAL
[NASAL] Sleep quality: Functional Outcomes of Sleep Questionnaire (follow-up 4-8 weeks; range of scores: not reported; Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	very serious[4]	serious[3]	none	28 (nasal flunisolide)	28	-	MD 0.89 higher (0.99 lower to 2.77 higher)	⊕OOO VERY LOW	CRITICAL
[NASAL] Sleep quality: Epworth Sleepiness Scale (follow-up 4-8 weeks; range of scores: 0-24; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	very serious[2]	serious[3]	none	28 (nasal flunisolide)	28	-	MD 3.18 lower (6.57 lower to 0.21 higher)	⊕OOO VERY LOW	CRITICAL
Exercise performance measure: Treadmill time (mins) (follow-up 6 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 2.6 higher (3.85 lower to 9.05 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Wellness scale (follow-up 11 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	very serious[5]	serious[3]	none	38 (fludrocortisone)	45	-	MD 1.1 higher (3.58 lower to 5.78 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Wellness scale (follow-up 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	30 (hydrocortisone)	35	-	MD 4.6higher (0.5 lower to 9.7 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Sickness Impact Profile (follow-up 12 weeks; range of scores: 0-68; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	33 (hydrocortisone)	34	-	MD 0.3 lower (3.46 lower to 2.86 higher)	⊕⊕OO LOW	CRITICAL
Symptom scales: headaches on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0 higher (1.55 lower to 1.55 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: painful lymph nodes on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	20 (fludrocortisone)	20	-	MD 0.2 lower (2.31 lower to 1.91 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: sore throat on VAS (follow-up 6 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20 (fludrocortisone)	20	-	MD 0.2 lower (1.8 lower to 1.4 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature (original analysis); percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details].
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[4] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. 2) Additionally downgraded due to all participants having rhinitis (Kakumanu 2003)
[5] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments): 1) downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature [original analysis]; percentage of participants with PEM unclear [PEM reanalysis – see Appendix G for additional details]. 2) Additionally downgraded due the majority of evidence coming from a study where all participants had neurally-mediated hypotension (Rowe 2001)

Table 30: Clinical evidence profile: Central antihypertensive drugs (clonidine) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Central antihypertensive drugs (clonidine) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Cognitive function: Stockings of Cambridge - minimum moves (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 1.22 lower (3.33 lower to 0.89 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Stockings of Cambridge - initial think time (secs) (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 1.28 lower (5.19 lower to 2.63 higher)	⊕OOO VERY LOW	CRITICAL

Cognitive function: Stockings of Cambridge - subsequent thinking time (secs) (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.51 lower (3.08 lower to 2.06 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Rapid Visual Information Processing - reaction time (secs) (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.15 lower (1.42 lower to 1.12 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Intradimensional (IDS) set sift errors (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.22 higher (0.34 lower to 0.78 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Extradimensional (EDS) set shift errors (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 2.66 lower (7.12 lower to 1.8 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Spatial working memory: between-search errors (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 2.17 lower (7.41 lower to 3.07 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Spatial working memory: strategy score (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.22 lower (5.92 lower to 5.48 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: pattern recognition - number correct (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 0.9 higher (0.77 lower to 2.57 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: spatial recognition - number correct (follow-up 30 minutes; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.1 lower (2.44 lower to 2.24 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: spatial span - length (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0.3 higher (0.84 lower to 1.44 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: delayed matching to sample 2 sec delay (follow-up 30 minutes; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	9	9	-	MD 1.22 lower (2.65 lower to 0.21 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: paired associate learning - sets completed (follow-up 30 minutes; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	9	9	-	MD 0 higher (0.3 lower to 0.3 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 31: Clinical evidence profile: Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Central nervous system stimulants (methylphenidate, modafinil, dexamphetamine, lisdexamphetamine) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Quality of Life: SF36 physical total (follow-up 4-6 weeks; range of scores: 0-100; Better indicated by higher values)												

2	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	70 (methylphenidate or dexamphetamine)	70	-	MD 1.63 higher (4.11 lower to 7.37 higher)	⊕⊕OO LOW	CRITICAL
Quality of Life: SF36 mental total (follow-up 4-6 weeks; range of scores: 0-100; Better indicated by higher values)												
2	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	70 (methylphenidate or dexamphetamine)	70	-	MD 3.51 higher (1.67 lower to 8.69 higher)	⊕⊕OO LOW	CRITICAL
Quality of Life: SF36 vitality (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 0.6 lower (15.95 lower to 14.75 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 physical role limitation (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	28 (modafinil)	14	-	MD 6.45 lower (26.66 lower to 13.76 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 physical function (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 1.6 lower (19.6 lower to 16.4 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 mental health (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 6.3 lower (16.26 lower to 3.66 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 emotional role limitation (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	28 (modafinil)	14	-	MD 19.3 lower (35.88 to 2.72 lower)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 pain (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 2.45 lower (22.61 lower to 17.71 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 social (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 2.4 lower (21.85 lower to 17.05 higher)	⊕OOO VERY LOW	CRITICAL
Quality of Life: SF36 general health (follow-up 20 days; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 0.4 lower (14.35 lower to 13.55 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Checklist Individual Strength (CIS) total score (follow-up 4-12 weeks; range of scores: 20-140; Better indicated by lower values)												
2	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	123 (methylphenidate)	125	-	MD 7.12 lower (12.07 to 2.16 lower)	⊕⊕OO LOW	CRITICAL
Fatigue: Fatigue Severity Scale (follow-up 6 weeks; range of scores: 9-63; Better indicated by lower values)												
2	randomised trials	serious[3]	very serious[4]	serious[1]	very serious[2]	none	23 (dexamphetamine or lisdexamphetamine)	21	-	MD 7.67 lower (21.75 lower to 6.4 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Chalder Physical Fatigue scale (follow-up 20 days; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 0.25 lower (4.92 lower to 4.42 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: Chalder Mental Fatigue scale (follow-up 20 days; range of scores: 0-12; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	28 (modafinil)	14	-	MD 0.4 higher (1.55 lower to 2.35 higher)	⊕OOO VERY LOW	CRITICAL
Sleep quality: sleep latency (time taken to fall asleep in mins) (follow-up 6 weeks; Better indicated by lower values)												

1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	10 (dexamphetamine)	10	-	MD 1.2 higher (2.91 lower to 5.31 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: HADS anxiety (follow-up 4 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	no serious imprecision	none	60 (methylphenidate)	60	-	MD 0.4 lower (1.74 lower to 0.94 higher)	⊕⊕⊕O MODERATE	CRITICAL
Psychological status: HADS depression (follow-up 4 weeks; range of scores: 0-21; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	no serious imprecision	none	60 (methylphenidate)	60	-	MD 0.4 lower (1.93 lower to 1.13 higher)	⊕⊕⊕O MODERATE	CRITICAL
Psychological status: Hamilton Anxiety Scale improvement (follow-up 6 weeks; range of scores: 0-56; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	13 (lisdexamphetamine)	11	-	MD 5.13 higher (2.08 lower to 12.34 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: AEs leading to discontinuation (follow-up 6-12 weeks)												
2	randomised trials	serious[3]	no serious inconsistency	serious[1]	serious[2]	none	10/78 (12.8%) (methylphenidate or lisdexamphetamine)	3/76 (3.9%)	RR 2.91 (0.9 to 9.43)	75 more per 1000 (from 4 fewer to 333 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: Serious AEs (pyelonephritis) (follow-up 12 weeks)												
1	randomised trials	serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	1/63 (1.6%) (methylphenidate)	0/65 (0%)	Peto OR 7.63 (0.15 to 384.58)	20 more per 1000 (from 30 fewer to 60 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: sleepiness (follow-up 4 weeks)												

1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	21/60 (35%) (methylphenidate)	23/60 (38.3%)	RR 0.91 (0.57 to 1.46)	34 fewer per 1000 (from 165 fewer to 176 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: dry mouth (follow-up 4-6 weeks)												
2	randomised trials	serious[3]	no serious inconsistency	serious[1]	serious[2]	none	35/75 (46.7%) (methylphenidate or lisdexamphetamine)	18/71 (25.4%)	RR 1.9 (1.22 to 2.96)	228 more per 1000 (from 56 more to 497 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: dizziness (follow-up 4 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	30/60 (50%) (methylphenidate)	38/60 (63.3%)	RR 0.79 (0.57 to 1.08)	133 fewer per 1000 (from 272 fewer to 51 more)	⊕⊕OO LOW	CRITICAL
Adverse events: akathisia (follow-up 4 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	29/60 (48.3%) (methylphenidate)	34/60 (56.7%)	RR 0.85 (0.61 to 1.2)	85 fewer per 1000 (from 221 fewer to 113 more)	⊕⊕OO LOW	CRITICAL
Adverse events: abdominal pain (follow-up 4 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	28/60 (46.7%) (methylphenidate)	23/60 (38.3%)	RR 1.22 (0.8 to 1.85)	84 more per 1000 (from 77 fewer to 326 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: chest pain (follow-up 4 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	17/60 (28.3%) (methylphenidate)	25/60 (41.7%)	RR 0.68 (0.41 to 1.12)	133 fewer per 1000 (from 246 fewer to 50 more)	⊕⊕OO LOW	CRITICAL
Adverse events: anorexia (follow-up 6 weeks)												

1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	5/10 (50%) (dexamphetamine)	1/10 (10%)	RR 5 (0.7 to 35.5)	400 more per 1000 (from 30 fewer to 1000 more)	⊕⊕OO LOW	CRITICAL
Adverse events: headache (follow-up 6 weeks)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	2/15 (13.3%) (lisdexamphetamine)	1/11 (9.1%)	RR 1.47 (0.15 to 14.21)	43 more per 1000 (from 77 fewer to 1000 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: insomnia (follow-up 6 weeks)												
1	randomised trials	very serious[2]	no serious inconsistency	serious[1]	very serious[2]	none	1/15 (6.7%) (lisdexamphetamine)	0/11 (0%)	Peto OR 5.66 (0.11 to 299.01)	70 more per 1000 (from 120 fewer to 250 more)	⊕OOO VERY LOW	CRITICAL
Adverse events (follow-up 20 days)												
1	randomised trials	serious[3]	no serious inconsistency	serious[1]	very serious[2]	none	21/28 (75%) (modafinil)	8/14 (57.1%)	RR 1.31 (0.79 to 2.17)	177 more per 1000 (from 120 fewer to 669 more)	⊕OOO VERY LOW	CRITICAL
Cognitive function: Behaviour Rating Inventory of Executive Function (BRIEF), improvement in global executive composite (follow-up 6 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	no serious imprecision	none	13 (lisdexamphetamine)	11	-	MD 18.02 higher (8.39 to 27.65 higher)	⊕OOO VERY LOW	CRITICAL
Pain: McGill pain Questionnaire improvement (follow-up 6 weeks; range of scores: 0-78; Better indicated by lower values)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	13 (lisdexamphetamine)	11	-	MD 7.84 higher (0.44 to 15.24 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Clinical Global Improvement - severity (follow-up 6 weeks; range of scores: 1-7; Better indicated by lower values)												

1	randomised trials	very serious[3]	no serious inconsistency	serious[1]	serious[2]	none	13 (lisdexamphetamine)	11	-	MD 1.28 higher (0.3 to 2.26 higher)	⊕OOO VERY LOW	CRITICAL

[1] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[3] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[4] Heterogeneity, I2=86%, p=0.05, unexplained by subgroup analysis.

Table 32: Clinical evidence profile: Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antiviral drugs (IV acyclovir or oral valganciclovir) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Fatigue: Multidimensional fatigue inventory (MFI-20) (follow-up 9 months; range of scores; 20-100; Better indicated by lower values)												
1 [original analysis]	randomised trials	very serious[1]	no serious inconsistency	very serious[2]	serious[3]	none	20 (oral valganciclovir)	10	-	MD 5.05 lower (11.48 lower to 1.38 higher)	⊕OOO VERY LOW	CRITICAL
1 [PEM reanalysis]	randomised trials	very serious[1]	no serious inconsistency	serious[4]	serious[3]	none						
Fatigue: POMS fatigue (follow-up 37 days; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 1.26 higher (1.01 lower to 3.53 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue: POMS vigour (follow-up 37 days; range of scores: 0-32; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 2.05 lower (4.65 lower to 0.55 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: POMS anxiety (follow-up 37 days; range of scores: 0-36; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 2.92 higher (0.63 to 5.21 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: POMS depression (follow-up 37 days; range of scores: 0-60; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 3.97 higher (0.69 to 7.25 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: POMS anger (follow-up 37 days; range of scores: 0-48; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 2.3 higher (0.13 lower to 4.73 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: POMS confusion (follow-up 37 days; range of scores: 0-28; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 1.83 higher (0.57 to 3.09 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: treatment-related adverse events (follow-up 9 months)												
1 [original analysis]	randomised trials	very serious[1]	no serious inconsistency	very serious[2]	no serious imprecision	none	0/20 (0%) (oral valganciclovir)	0/10 (0%)	RD 0.00 (-0.14 to 0.14)	0 more per 1000 (from 140 fewer to 140 more)	⊕OOO VERY LOW	CRITICAL
1 [PEM reanalysis]	randomised trials	very serious[1]	no serious inconsistency	serious[4]	no serious imprecision	none						
Adverse events: reversible renal failure (follow-up 37 days)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	3/27 (11.1%) (IV acyclovir)	0/27 (0%)	Peto OR 7.99 (0.8 to 80.28)	11 more per 1000 (from 20 fewer to 240 more)	⊕OOO VERY LOW	CRITICAL
Activity levels: rest (hours/day) (follow-up 37 days; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 0.05 lower (0.83 lower to 0.73 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Wellness score (follow-up 37 days; range of scores: not reported; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	27 (IV acyclovir)	27	-	MD 1.08 lower (7.28 lower to 5.12 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature (one increment). Montoya 2013 was additionally downgraded due to population having suspected viral onset and requirement to have elevated antibody titres. [original analysis]
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[4] The majority of the evidence included an indirect population (downgraded by one increment): requirement for suspected viral onset and elevated viral antibody tiers (Montoya 2013). [PEM reanalysis]

Table 33: Clinical evidence profile: 5-HT3 antagonists (ondansetron) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**5-HT3 antagonists (ondansetron) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Fatigue: CIS fatigue (follow-up 12 weeks; range of scores: 8-56; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 1.4 lower (6.81 lower to 4.01 higher)	⊕OOO VERY LOW	CRITICAL
Activity levels: Actometer (objective accelerometer-based method of measuring activity) (follow-up 12 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 5.6 lower (13.61 lower to 2.41 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: constipation (follow-up 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	2/33 (6.1%)	0/34 (0%)	Peto OR 7.86 (0.48 to 128.37)	60 more per 1000 (from 40 fewer to 160 more)	⊕OOO VERY LOW	CRITICAL
Adverse events: malaise (follow-up 12 weeks)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	3/33 (9.1%)	1/34 (2.9%)	RR 3.09 (0.34 to 28.23)	61 more per 1000 (from 19 fewer to 801 more)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Sickness Impact Profile (SIP) 8 (follow-up 12 weeks; range of scores 0-5799; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	33	34	-	MD 109 lower (403.38 lower to 185.38 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 34: Clinical evidence profile: Galantamine hydrobromide versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Galantamine hydrobromide versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Fatigue: fatigue on VAS (follow-up 2 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	25	24	-	MD 0.14 higher (0.84 lower to 1.12 higher)	⊕OOO VERY LOW	CRITICAL
Cognitive function: memory on VAS (follow-up 2 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	25	24	-	MD 0.91 higher (0.67 lower to 2.49 higher)		CRITICAL
Pain: myalgia on VAS (follow-up 2 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	25	24	-	MD 0.47 lower (1.39 lower to 0.45 higher)	⊕OOO VERY LOW	CRITICAL
Sleep quality: sleep disturbance on VAS (follow-up 2 weeks; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	25	24	-	MD 0.34 higher (1.02 lower to 1.7 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: AEs dizziness on VAS (follow-up 2 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	25	24	-	MD 0.72 higher (0.93 lower to 2.37 higher)	⊕OOO VERY LOW	CRITICAL
Return to work: work capacity/satisfaction on VAS (follow-up 2 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	25	14	-	MD 0.17 lower (1.38 lower to 1.04 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: clinical global impression score, no change or worse (follow-up 20 weeks)												
1	randomised trials	very serious[3]	no serious inconsistency	serious[2]	serious[3]	none	169/280 (60.4%)	47/67 (70.1%)	RR 0.86 (0.72 to 1.03)	98 fewer per 1000 (from 196 fewer to 21 more)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 35: Clinical evidence profile: Antihistamines (terfenadine) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antihistamines versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Physical functioning: modified Medical Outcome Study Short Form - physical functioning (follow-up 2 months; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	14	14	-	MD 6.56 lower (19.75 lower to 6.63 higher)	⊕OOO VERY LOW	CRITICAL
Psychological status: modified Medical Outcome Study Short Form - mental health (follow-up 2 months; range of scores: 0-100; Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	14	14	-	MD 10.73 lower (24.5 lower to 3.04 higher)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 36: Clinical evidence profile: Pro-inflammatory cytokine antagonists (anakinra) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Pro-inflammatory cytokine antagonists (anakinra) versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Mortality (follow-up 24 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	no serious imprecision	none	0/25 (0%)	0/25 (0%)	RD 0.00 (-0.07 to 0.07)	0 more per 1000 (from 70 fewer to 70 more)	⊕⊕⊕O MODERATE	CRITICAL
Fatigue: CIS fatigue (follow-up 24 weeks; range of scores: 8-56; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	25	25	-	MD 1.3 higher (5.3 lower to 7.9 higher)	⊕OOO VERY LOW	CRITICAL
Physical functioning: SF36 physical function (follow-up 24 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	25	25	-	MD 4 lower (15.1 lower to 7.1 higher)	⊕⊕OO LOW	CRITICAL
Psychological status: Symptom Checklist 90 (follow-up 24 weeks; range of scores: 90-450; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	no serious imprecision[2]	none	25	25	-	MD 3 higher (8.6 lower to 14.6 higher)	⊕⊕⊕O MODERATE	CRITICAL

Pain: VAS maximum pain score (follow-up 24 weeks; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	25	25	-	MD 0.34 higher (1.1 lower to 1.78 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events (follow-up 24 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	no serious imprecision	none	24/25 (96%)	14/25 (56%)	RR 1.71 (1.2 to 2.45)	398 more per 1000 (from 112 more to 812 more)	⊕⊕⊕O MODERATE	CRITICAL
Adverse events: withdrawal due to adverse events (follow-up 24 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	none	1/25 (4%)	0/25 (0%)	Peto OR 7.39 (0.15 to 372.38)	40 more per 1000 (from 60 fewer to 140 more)	⊕OOO VERY LOW	CRITICAL
Symptom scales: Sickness Impact Profile (follow-up 24 weeks; range of scores: 0-5799; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious[1]	serious[2]	none	25	25	-	MD 91.2 higher (275.8 lower to 458.2 higher)	⊕⊕OO LOW	CRITICAL

[1] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 37: Clinical evidence profile: Staphylococcus vaccine (Staphypan Berna) versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Staphylococcus vaccine versus placebo**	**Control**	**Relative (95% CI)**	**Absolute**		
Pain: pain on VAS (follow-up 32 weeks; range of scores: unclear; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	very serious[1]	no serious imprecision	none	49	49	-	MD 0.3 lower (1.12 lower to 0.52 higher)	⊕⊕OO LOW	CRITICAL
Adverse events (follow-up 32 weeks)												

1	randomised trials	no serious risk of bias	no serious inconsistency	very serious[1]	very serious[2]	none	28/50 (56%)	26/50 (52%)	RR 1.08 (0.75 to 1.55)	42 more per 1000 (from 130 fewer to 286 more)	⊕OOO VERY LOW	CRITICAL
Symptom scales: clinical global impression of change (follow-up 32 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	very serious[1]	serious[2]	none	49	49	-	MD 0.7 lower (1.22 to 0.18 lower)	⊕OOO VERY LOW	CRITICAL
Symptom scales: clinical global impression of severity (follow-up 32 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	very serious[1]	serious[2]	none	49	49	-	MD 0.3 lower (0.53 to 0.07 lower)	⊕OOO VERY LOW	CRITICAL

[1] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature. Zachrisson 2002 was downgraded twice due to population also meeting diagnostic criteria for fibromyalgia.
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 38: Clinical evidence profile: Central antihypertensive drugs (clonidine) versus placebo (children and young people)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Children and young people: Sympathomimetic/central antihypertensive drugs versus placebo**		**Relative (95% CI)**	**Absolute**		
Fatigue: Chalder Fatigue Questionnaire (CFQ) total sum score (follow-up 30 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	54	49	-	MD 0.5 higher (14.7 lower to 15.7 higher)	⊕OOO VERY LOW	CRITICAL
Physical functioning: Fatigue Disability Index (FDI) total sum score (follow-up 30 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	none	54	49	-	MD 0.2 higher (13.3 lower to 13.7 higher)	⊕OOO VERY LOW	CRITICAL
Pain: BPI average pain score (follow-up 30 weeks; range of scores: 0-10; Better indicated by lower values)												

1	randomised trials	serious[1]	no serious inconsistency	serious[4]	serious[3]	none	54	49	-	MD 0.4 higher (0.4 lower to 1.2 higher)	⊕OOO VERY LOW	CRITICAL
Sleep quality: KSQ insomnia score (follow-up 30 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	54	49	-	MD 0.1 higher (0.3 lower to 0.5 higher)	⊕OOO VERY LOW	CRITICAL
Adverse effects: various self-reported (follow-up 9 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	very serious[2,4]	serious[3]	none	43/57 (75.4%)	33/51 (64.7%)	RR 1.17 (0.91 to 1.5)	110 more per 1000 (from 58 fewer to 324 more)	⊕OOO VERY LOW	CRITICAL
Activity levels: steps per day (accelerometer) (follow-up 30 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	54	49	-	MD 119 higher (796 lower to 1034 higher)	⊕⊕OO LOW	CRITICAL
Cognitive function: Digit span backward test total (follow-up 30 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	none	54	49	-	MD 0.5 lower (1.2 lower to 0.2 higher)	⊕OOO VERY LOW	CRITICAL
Symptom scales: CFS symptom inventory hypersensitivity score (follow-up 30 weeks; range of scores: not reported; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	54	49	-	MD 0.03 lower (0.4 lower to 0.34 higher)	⊕⊕OO LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment) or a very indirect population (downgraded by two increments). Populations were downgraded if the ME/CFS diagnostic criteria used did not include PEM as a compulsory feature
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[4] Outcome indirectness: Some adverse effects are poorly defined, e.g. "unwellness" and "other"

Appendix G PEM reanalysis

G.1 PEM reanalysis – Methods and rationale

After considering the stakeholder comments the committee agreed to revisit the evidence for the intervention reviews further scrutinising the information on PEM reported in the trials and the application of indirectness in the evidence. In the original analysis studies were downgraded for indirectness if the diagnostic criteria used in the study did not have PEM as a compulsory feature.

We looked for any published information on the percentage of participants with PEM in the included trials, or subgroup analyses in study participants with PEM. The papers for all included studies were reviewed again, as well as any published supplements. The excluded studies list was also re-examined to ensure any relevant information relating to PEM in the included studies were not missed. Unpublished data was not accepted for this analysis.

The committee agreed that studies using criteria without PEM as a compulsory feature (e.g. 1994 CDC criteria, Oxford criteria) should not be downgraded if a high proportion of study participants had PEM and this was adequately described. In order to not downgrade the following criteria must be met:

- ≥95% of study participants are reported to have PEM (or a subgroup analysis where ≥95% participants are reported to have PEM)

AND

- If another term is used other than PEM (e.g. post-exertional fatigue) there must be a clear description that indicates all of the following:
    - Symptom worsening that follows minimal physical or mental activity that was previously tolerated
    - Symptom worsening is typically delayed (12-24 hours after the activity)
    - The impact is prolonged

OR

- ≥95% of study participants meet diagnostic criteria where PEM is compulsory (e.g. IOM 2015, NICE 2007, Carruthers 2003/Canadian criteria).

Only new or changed results are reported below. Results for outcomes where applying the above criteria did not change the indirectness rating are not shown. See original results in Section 1.1.4, Appendix E (forest plots), Appendix F (GRADE tables), and the committee's discussion and interpretation of the evidence in Section 1.2 of this report. Additionally, results reported below have also been added throughout the report, alongside the original results.

G.2 PEM reanalysis – Summary of results

Immunomodulatory drugs

No new information on PEM identified that required re-analysis.

Antidepressants

No new information on PEM identified that required re-analysis.

Corticosteroids

No new information on PEM identified that required re-analysis.

Antihypertensive drugs

No new information on PEM identified that required re-analysis.

CNS stimulants

No new information on PEM identified that required re-analysis.

Antiviral drugs

For the comparison of oral valganciclovir vs placebo (1 study – Montoya 2013), the percentage of participants with PEM was reported (96.7% of participants had PEM). This study had previously been downgraded as the 1994 CDC criteria was used which does not have PEM as a compulsory feature.

Changing the indirectness rating resulted in a change from very serious indirectness to serious indirectness (the study remained downgraded due to other concerns about the population). The overall quality of the evidence remained very low for all outcomes.

Amended GRADE tables are presented below in section G.3. Also see original results in Section 1.1.4, Appendix E (forest plots), Appendix F (GRADE tables), and the committee's discussion and interpretation of the evidence in Section 1.2 of this report.

For the remaining study in this category (IV acyclovir vs placebo) there was no new information on PEM identified that required re-analysis.

Other drugs

No new information on PEM identified that required re-analysis.

G.3 PEM reanalysis – Reporting of PEM

Table 39: Summary PEM reporting in each study

Study	Intervention	Reporting of PEM	Effect on indirectness rating
Immunomodulatory drugs			
Fluge 2011[22]	Rituximab vs placebo	A composite continuous outcome which included fatigue, post-exertional exhaustion, need from rest and daily functioning was reported, but the percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Fluge 2019[23]	Rituximab vs placebo	All participants met the Canadian criteria (Carruthers 2003) which has PEM as a compulsory feature.	Remain not downgraded (all participants met the Canadian criteria)
Lloyd 1990[44]	IVIG vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Peterson 1990[65]	IVIG vs placebo	96.4% of participants reported to have “post-exertional fatigue (prolonged). This did not meet the committee’s accepted definition of PEM.	Remain downgraded (unclear if symptom described in study is PEM)
Strayer 2012[88]	IVIG vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Vollmer-Conna 1997[99]	Rintatolimod vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Antidepressants			

Study	Intervention	Reporting of PEM	Effect on indirectness rating
Arnold 2015[4]	Duloxetine (SNRI) vs placebo	A composite continuous outcome which included eight symptoms, including postexertional fatigue was reported, but the percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear).
Hickie 2000[34]	Moclobemide (MAOI) vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Pardini 2011[63]	Fluoxetine (SSRI) vs amisulpride (atypical antipsychotic)	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Vercoulen 1996[97]	Fluoxetine (SSRI) vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Wearden 1998[103]	Fluoxetine (SSRI) + GET vs Drug placebo + GET vs Fluoxetine (SSRI) + exercise control vs Drug placebo + exercise control	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Corticosteroids			
Kakumanu 2003[37]	Nasal flunisolide vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Mckenzie 1998[47]	Hydrocortisone vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Peterson 1998[64]	Fludrocortisone vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Rowe 2001[77]	Fludrocortisone vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)

Study	Intervention	Reporting of PEM	Effect on indirectness rating
Antihypertensive drugs			
Morriss 2002[56]	IV clonidine vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Sulheim 2014[91] [CYP]	Oral clonidine vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Central nervous system stimulants			
Blockmans 2006[9]	Methylphenidate vs placebo	PEM was reported as a continuous outcome, but the percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Montoya 2018[52]	Methylphenidate vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Olson 2003[61]	Dexamphetamine vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Young 2013[107]	Lis-dexamphetamine vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Randall 2005[68]	Modafinil vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Antiviral drugs			
Montoya 2013[53]	Oral valganciclovir vs placebo	96.7% of participants had PEM (95% of participants in valganciclovir arm and 100% in placebo arm.	Un-downgraded (the indirectness rating changes to serious from very serious as the study meets the 95% threshold for PEM, but remains downgraded for other population concerns)

Study	Intervention	Reporting of PEM	Effect on indirectness rating
Straus 1988[86]	IV acyclovir vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
5-HT3 antagonists			
The 2010[93]	Ondansetron vs placebo	The mean number of CDC symptoms were reported, but the percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Galantamine hydrobromide			
Blacker 2004[8]	Galantamine hydrobromide vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Snorrason 1996[79]	Galantamine hydrobromide vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Antihistamines			
Steinberg 1996[84]	Terfenadine vs placebo	82.1% of participants had “post-exertional fatigue (prolonged)”.	Remain downgraded (unclear if symptom described in study is PEM and <95%)
Pro-inflammatory cytokine antagonists			
Roerink 2017[70]	Anakinra vs placebo	The mean number of CDC symptoms were reported, but the percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)
Staphylococcus vaccine			
Zachrisson 2002[109]	Staphylococcus toxoid vaccine vs placebo	The percentage of participants with PEM was not reported.	Remain downgraded (percentage of participants with PEM unclear)

G.4 PEM reanalysis – GRADE tables

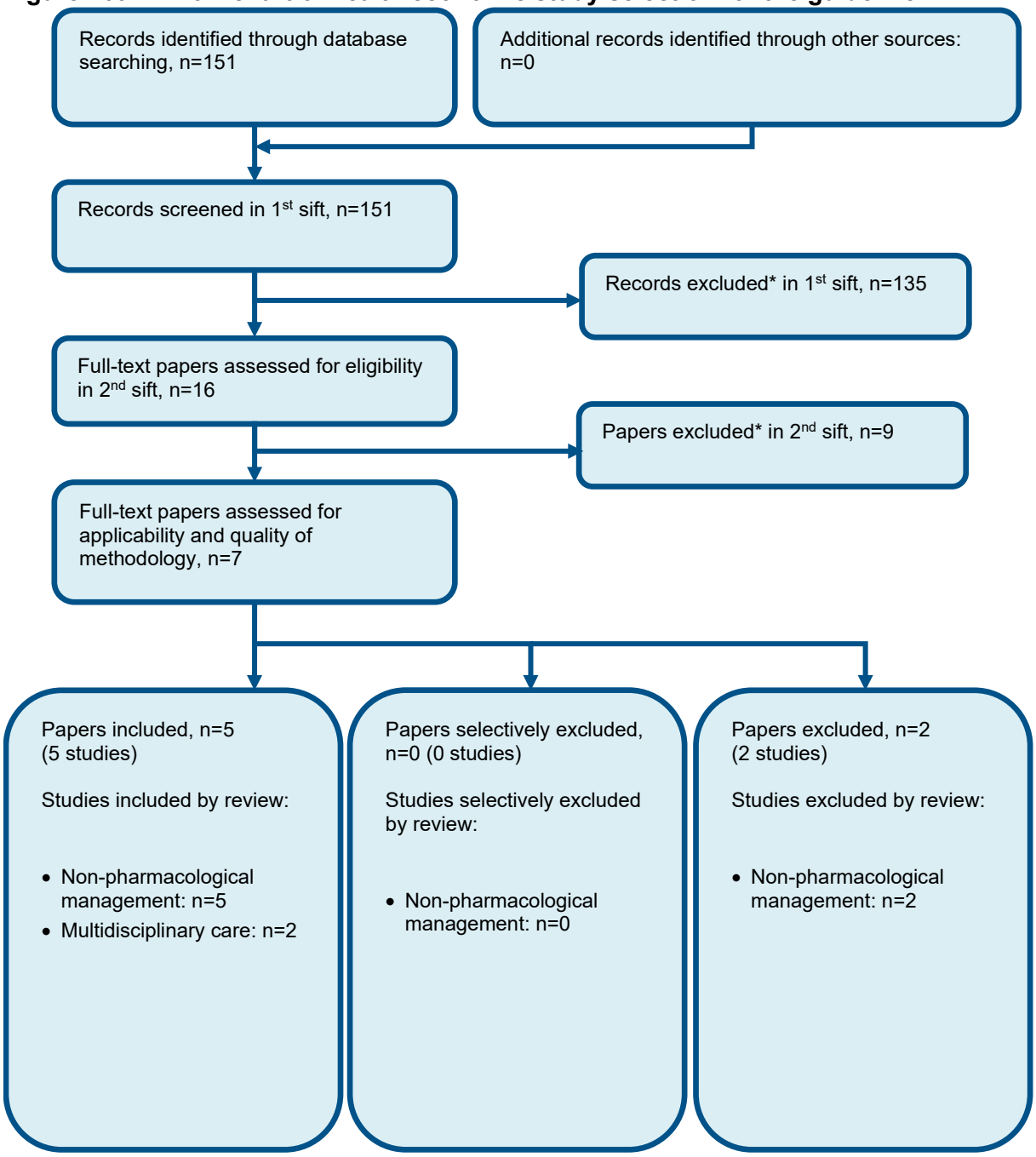
Antiviral drugs

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Antiviral drugs (oral valganciclovir) versus placebo	Control	Relative (95% CI)	Absolute		
Fatigue: Multidimensional fatigue inventory (MFI-20) (follow-up 9 months; range of scores; 20-100; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	none	20	10	-	MD 5.05 lower (11.48 lower to 1.38 higher)	⊕OOO VERY LOW	CRITICAL
Adverse events: treatment-related adverse events (follow-up 9 months)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	no serious imprecision	none	0/20 (0%)	0/10 (0%)	RD 0.00 (-0.14 to 0.14)	0 more per 1000 (from 140 fewer to 140 more)	⊕OOO VERY LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] The majority of the evidence included an indirect population (downgraded by one increment): requirement for suspected viral onset and elevated viral antibody tiers (Montoya 2013). [PEM reanalysis]
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Appendix H Economic evidence study selection

Figure 169: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language

NB. Two papers were included in both the non-pharma and the multidisciplinary care reviews, in parallel with the review of clinical effectiveness.

Appendix I Economic evidence tables

No economic evaluations were found.

Appendix J Excluded studies

Table 40: Studies excluded from the clinical review

Study	Exclusion reason
Agger 2017[1]	Incorrect population; multiorgan bodily distress syndrome
Andersson 1998[2]	Incorrect study design (non-randomised)
Anonymous 2004[25]	News article
Anonymous 2019[76]	Editorial
Anonymous 2019[3]	Summary article (not full study)
Anonymous 2020[81]	Abstract only
Ascough 2020[5]	Systematic review with different PICO
Baschetti 1999[6]	Letter
Behan 1995[7]	Conference proceedings
Blockmans 2003[10]	Crossover, no washout period
Brook 1993[11]	Correspondence
Brostoff 2000[12]	On National Registry only - paper cannot be found
Cleare 1998[13]	Conference proceeding
Cleare 1999[14]	Crossover, no washout period
Cleare 2000[17]	No relevant outcomes
Cleare 2001[15]	No relevant outcomes
Cleare 2004[16]	No relevant outcomes
Collatz 2016[18]	Systematic review with different PICO
De Vinci 1996[19]	No relevant outcomes
Diaz-Mitoma 2003[20]	Only 14/16 patients had CFS; no useable outcomes
Fagermoen 2015[21]	No relevant outcomes
Fluge 2015[24]	Non-randomised follow up to Fluge 2011
Glazachev 2017[26]	Article not in English

Study	Exclusion reason
Gracious 1991[27]	Case study
Guo 2015[28]	Article not in English
Hall 2016[29]	Secondary analysis of an included trial
Hartz 2003[30]	Incorrect study design (non-randomised)
Hermans 2018[31]	Study not relevant: one-off pain treatment; outcomes not relevant
Hickie 1992[33]	Incorrect study design (non-randomised)
Hickie 1999[32]	Incorrect study design (non-randomised)
Houdenhove 2010[35]	Narrative review
Kakumanu 2001[36]	Abstract only
Kreijkamp-Kaspers 2011[38]	Incorrect study design (non-randomised study and review)
Lane 1999[39]	Unobtainable - record on national research registry.
Lerner 2001[42]	Correspondence
Lerner 2002[40]	Incorrect study design (non-randomised); incorrect comparison (comparing two patient groups not treatments)
Lerner 2007[41]	No useable outcomes
Liu 2010[43]	Article not in English
Lloyd 1993[45]	No useable outcomes
Main 2000[46]	Unobtainable - only exists on national register or research so may not be published
Mckenzie 2000[48]	No relevant outcomes
Meeus 2013[49]	Incorrect intervention (one-off pain treatment)
Mehta 1995[50]	Incorrect study design (non-randomised; single case study with cross-over)
Mitchell 2000[51]	Unobtainable. Only reported on trial database
Moorkens 1998[54]	No useable outcomes
Morriss 1996[55]	Citation only
Natelson 1996[58]	Analysis technique unclear and not possible to interpret
Natelson 1998[57]	No useable outcomes
Nilsson 2018[60]	Incorrect intervention (drug not licensed)
Pae 2009[62]	Systematic review is not relevant to review question or unclear PICO

Study	Exclusion reason
Plioplys 1997[66]	Incorrect study design (non-randomised)
Raijmakers 2019[67]	Incorrect population (Q fever fatigue syndrome)
Rekeland 2020[69]	Incorrect study design (non-randomised)
Roerink 2015[71]	Study protocol
Roerink 2017[72]	Incorrect study design (RCT results reported elsewhere)
Roerink 2018[73]	No relevant outcomes
Rowe 1997[74]	No relevant extractable outcomes
Rowe 1999[75]	Not all participants were randomised
See 1996[78]	Crossover, no washout period
Spath 2000[80]	No useable outcomes
Springer 1992[82]	Abstract only
Staud 2017[83]	Incorrect intervention (one-off pain treatment); no relevant outcomes
Stouch 2010[85]	No relevant outcomes
Strayer 1994[87]	No relevant extractable outcomes
Stubhaug 2008[89]	Incorrect population (neurasthenia population; not all participants met CFS criteria)
Suhadolnik 1994[90]	No relevant outcomes
Teitelbaum 1999[92]	Incorrect population (principally FMS; not all patients had co-existing CFS)
The 2014[94]	Intervention details unclear
Tiev 1999[95]	Article not in English
Vedhara 1997[96]	No useable outcomes
Vercoulen 1997[98]	Article not in English
Vorob'eva 2017[101]	Article not in English
Vorob'eva 2019[100]	Incorrect population; no diagnosis of ME/CFS
Wearden 1996[102]	Citation only
Williams 2002[104]	No usable outcome data - results reported as medians (IQR)
Wilson 2019[105]	Citation only
Young 2010[106]	On clinical trials website only - otherwise unavailable
Zachrisson 2004[108]	No relevant outcomes

Appendix K MIDs for continuous outcomes

Table 41: MID for continuous outcomes (0.5 x SD): Immunomodulatory drugs versus placebo

Outcomes	MID
Quality of Life: SF36 physical composite (max % change from baseline) Scale from: 0 to 100	8.5
Quality of Life: SF36 mental composite (max % change from baseline) Scale from: 0 to 100	16
Fatigue/fatigability: Fatigue severity scale Scale from: 9 to 63.	2.5
Fatigue/fatigability: numeric rating scale Scale from: 0 to 10.	0.71
Psychological status: Hamilton Depression Scale Scale from: 0 to 52.	1.55
Psychological status: Zung Self-Rating Depression Scale Scale from: 0 to 80.	4.75
Psychological status: mental health on the Medical Outcome Study Short Form Scale from: 0 to 100	7.63
Physical functioning: physical functioning on the Medical Outcome Study Short Form Scale from: 0 to 100	10.5
Physical functioning: physical functioning on the SF36 Scale from: 0 to 100	13.56
Physical functioning: functional level percentage	5.08
Activity levels: mean number of steps per 24 hours	1036.5
Exercise performance measure: Treadmill exercise duration in seconds	122.98

Table 42: MID for continuous outcomes (0.5 x SD): Antidepressants versus placebo

Outcomes	MID
Quality of Life: SF36 vitality Scale from: 0 to 100.	7.5
Quality of Life: SF-36 physical functioning Scale from: 0 to 100.	9.9
Quality of Life: SF-36 role physical Scale from: 0 to 100.	15.43
Quality of Life: SF36 mental health Scale from: 0 to 100.	8.56
Quality of Life: SF36 role emotional Scale from: 0 to 100.	19.23
Quality of Life: SF36 bodily pain Scale from: 0 to 100.	8.33
Quality of Life: SF36 general health Scale from: 0 to 100.	9.2
Quality of Life: SF36 social functioning Scale from: 0 to 100.	10.93
Fatigue: 14-item Chalder fatigue scale Scale from: not reported	3.89
Fatigue: MFI-20 general fatigue Scale from: not reported	0.98
Fatigue: MFI-20 physical fatigue Scale from: not reported	1.43
Fatigue: MFI-20 reduced activity Scale from: not reported	1.78
Fatigue: MFI-20 reduced motivation Scale from: not reported	1.68
Fatigue: MFI-20 mental fatigue Scale from: not reported	1.55

Outcomes	MID
Fatigue: Checklist Individual Strength (CIS) fatigue Scale from: not reported	0.21
Physical functioning: Karnofsky Performance Index Scale from: not reported	0.65
Psychological status: Profile of mood states (POMS) fatigue Scale from: not reported	0.15
Psychological status: Profile of mood states (POMS) vigour Scale from: not reported	0.55
Psychological status: Profile of mood states (POMS) depression Scale from: not reported	0.35
Psychological status: HADS depression Scale from: 0 to 21.	1.44
Psychological status: HADS depression Scale from: 0 to 21.	1.95
Psychological status: HADS anxiety Scale from: 0 to 21.	1.63
Psychological status: Beck Depression Inventory Scale from: 0 to 63.	1.13
Pain: Brief Pain Inventory severity Scale from: 0 to 10.	0.98
Pain: Brief Pain Inventory interference Scale from: 0 to 10.	1.13
Exercise performance measure: VO2 max (mL O2/kg/min)	2.44
Symptom scales: Clinical Global Impression of Severity Scale from: not reported	0.19
Symptom scales: Clinical Global Impression of Improvement Scale from: not reported	0.87
Symptom scales: CDC symptom inventory Scale from: not reported	12.08

Table 43: MID for continuous outcomes (0.5 x SD): Antidepressants versus graded exercise

Outcomes	MID
Fatigue: 14-item Chalder fatigue scale Scale from: not reported	5.45
Psychological status: HADS depression Scale from: 0 to 21	1.75
Exercise performance measure: VO2 max (mL O2/kg/min)	2.87

Table 44: MID for continuous outcomes (0.5 x SD): Antidepressants versus combined antidepressants and graded exercise

Outcomes	MID
Fatigue: 14-item Chalder fatigue scale Scale from: not reported	5.22
Psychological status: HADS depression Scale from: 0 to 21	1.84
Exercise performance measure: VO2 max (mL O2/kg/min)	2.12

Table 45: MID for continuous outcomes (0.5 x SD): Combined antidepressants and graded exercise versus placebo

Outcomes	MID
Fatigue: 14-item Chalder fatigue scale Scale from: not reported	3.89
Psychological status: HADS depression Scale from: 0 to 21	1.44
Exercise performance measure: VO2 max (mL O2/kg/min)	2.44

Table 46: MID for continuous outcomes (0.5 x SD): Combined antidepressants and graded exercise versus graded exercise

Outcomes	MID
Fatigue: 14-item Chalder fatigue scale Scale from: not reported	5.45
Psychological status: HADS depression Scale from: 0 to 21	1.75
Exercise performance measure: VO2 max (mL O2/kg/min)	2.87

Table 47: MID for continuous outcomes (0.5 x SD): Antidepressants versus antipsychotics

Outcomes	MID
Quality of Life: SF12 Scale from: 0 to 100.	3.53
Fatigue: Fatigue Severity Scale Scale from: 9 to 63.	3.18
Psychological status: HADS anxiety Scale from: 0 to 21.	0.43
Psychological status: HADS depression Scale from: 0 to 21.	0.43
Pain: pain on VAS Scale from: 0 to 100.	4.75
Adverse events: FIBSER global burden Scale from: not reported.	0.35
Symptom scales: Clinical Global Impression Severity (CGI-S) Scale from: 1 to 7.	0.38

Table 48: MID for continuous outcomes (0.5 x SD): Corticosteroids versus placebo

Outcomes	MID
Quality of Life: SF36 physical Scale from: 0 to 100.	11.48
Quality of Life: SF36 energy or fatigue Scale from: 0 to 100.	7.78
Quality of Life: SF36 emotional wellbeing Scale from: 0 to 100.	6.43
Quality of Life: SF36 role emotional Scale from: 0 to 100.	12.23
Quality of Life: SF36 role physical Scale from: 0 to 100.	14.8
Quality of Life: SF36 pain Scale from: 0 to 100.	13.7
Quality of life: SF36 social Scale from: 0 to 100.	12.55
Quality of life: SF36 general wellbeing Scale from: 0 to 100.	6.23
Fatigue: fatigue on VAS Scale from: 0 to 10.	0.8
Fatigue: Profile of Mood States – fatigue Scale from: 0 to 28	2.75 (hydrocortisone) 3.95 (fludrocortisone)
Fatigue: Profile of Mood States – vigour Scale from: 0 to 32	2.15 (hydrocortisone) 3.35 (fludrocortisone)
Fatigue: Wood Mental Fatigue Inventory Scale from: 0 to 36.	4.48
Physical function: SF36 physical function Scale from: 0 to 100.	11.3

Outcomes	MID
Psychological status: SF36 mental health Scale from: 0 to 100.	8.6
Psychological status: Beck Depression Inventory Scale from: 0 to 63.	1.95 (hydrocortisone) 3.4 (fludrocortisone)
Psychological status: Profile of Mood States – anger Scale from: 0 to 48.	2.68
Psychological status: Profile of Mood States – anxiety Scale from: 0 to 36.	2.58
Psychological status: Profile of Mood States – confusion Scale from: 0 to 28.	2.35
Psychological status: Profile of Mood States – depression Scale from: 0 to 60.	2.38
Psychological status: Symptom checklist-90-R general sensitivity index Scale from: not reported	0.15
Psychological status: Symptom checklist-90-R positive symptom distress index Scale from: not reported	0.23
Psychological status: Symptom checklist-90-R positive symptom total Scale from: not reported	6.38
Psychological status: Hamilton Depression Rating Scale Scale from: not reported	1.73
Psychological status: Positive and negative effect scale (PANAS) positive affect Scale from: 10 to 50.	3.08
Activity levels: activity scale Scale from: not reported	0.98
Activity levels: distance before exhausted (ordinal scale) Scale from: 1 to 5.	0.58
Activity levels: Duke Activity Status Index	3.88

Outcomes	MID
Scale from: 0 to 58.2	
Cognitive function: Reaction time (secs)	0.03
Cognitive function: inability to concentrate on VAS Scale from: 0 to 10.	1.3
Cognitive function: forgetfulness on VAS Scale from: 0 to 10.	1.23
Cognitive function: confusion on VAS Scale from: 0 to 10.	1.43
Pain: muscle pain on VAS Scale from: 0 to 10.	1.4
Pain: joint pain on VAS Scale from: 0 to 10.	1.78
Sleep quality: unrefreshing sleep on VAS Scale from: 0 to 10.	0.85
Sleep quality: Functional Outcomes of Sleep Questionnaire Scale from: not reported	1.8
Sleep quality: Epworth Sleepiness Scale Scale from: 0 to 24.	3.24
Exercise performance measure: Treadmill time (mins)	5.73
Symptom scales: Wellness scale Scale from: 0 to 100.	7.8 (hydrocortisone) 5.0 (fludrocortisone)
Symptom scales: Sickness Impact Profile Scale from: 0 to 68.	4.18
Symptom scales: headaches on VAS Scale from: 0 to 10.	1.25
Symptom scales: painful lymph nodes on VAS Scale from: 0 to 10.	1.78

Outcomes	MID
Symptom scales: sore throat on VAS Scale from: 0 to 10.	1.38
Symptom scales: Chronic Fatigue Syndrome Severity Rating Scale from: not reported	4.12

Table 49: MID for continuous outcomes (0.5 x SD): Antihypertensive drugs versus placebo

Outcomes	MID
Cognitive function: Stockings of Cambridge - minimum moves	1.2
Cognitive function: Stockings of Cambridge - initial think time (secs)	2.07
Cognitive function: Stockings of Cambridge - subsequent thinking time (secs)	1.54
Cognitive function: Rapid Visual Information Processing - reaction time (secs)	0.61
Cognitive function: Intradimensional (IDS) set sift errors	0.22
Cognitive function: Extradimensional (EDS) set shift errors	3.32
Cognitive function: Spatial working memory: between-search errors	3.41
Cognitive function: Spatial working memory: strategy score	3.19
Cognitive function: pattern recognition - number correct	1.1
Cognitive function: spatial recognition - number correct	1.05
Cognitive function: spatial span - length	0.6
Cognitive function: delayed matching to sample 2 sec delay	0.7
Cognitive function: paired associate learning - sets completed	0.17

Table 50: MID for continuous outcomes (0.5 x SD): Central nervous system stimulants versus placebo

Outcomes	MID
Quality of Life: SF36 physical total Scale from: 0 to 100.	5.42
Quality of Life: SF36 mental total Scale from: 0 to 100.	6.53
Quality of Life: SF36 vitality Scale from: 0 to 100.	2.05
Quality of Life: SF36 physical role limitation Scale from: 0 to 100.	4.5
Quality of Life: SF36 physical function Scale from: 0 to 100.	3.85
Quality of Life: SF36 mental health Scale from: 0 to 100.	1.85
Quality of Life: SF36 emotional role limitation Scale from: 0 to 100.	3.8
Quality of Life: SF36 pain Scale from: 0 to 100.	4.4
Quality of life: SF36 social Scale from: 0 to 100.	4
Quality of life: SF36 general health Scale from: 0 to 100.	2.5
Fatigue: Checklist Individual Strength (CIS) total score Scale from: 20 to 140.	9.24
Fatigue: Fatigue Severity Scale Scale from: 9 to 63.	2.24
Fatigue: Chalder Physical Fatigue scale Scale from: 0 to 21.	0.6

Outcomes	MID
Fatigue: Chalder Mental Fatigue scale Scale from: 0 to 12.	0.35
Sleep quality: sleep latency (time taken to fall asleep in mins)	3.99
Psychological status: HADS anxiety Scale from: 0 to 21.	2.2
Psychological status: HADS depression Scale from: 0 to 21.	1.95
Psychological status: Hamilton Anxiety Scale	4.14
Cognitive function: Behaviour Rating Inventory of Executive Function (BRIEF), global executive composite Scale from: not reported.	3.63
Pain: McGill pain Questionnaire Scale from: 0 to 78	4.77
Symptom scales: Clinical global improvement Scale from: not reported.	0.46

Table 51: MID for continuous outcomes (0.5 x SD): Antiviral drugs versus placebo

Outcomes	MID
Fatigue: Multidimensional fatigue inventory (MFI-20) Scale from: 20 to 100.	7.15
Fatigue: POMS fatigue Scale from: 0 to 28.	2.13

Outcomes	MID
Fatigue: POMS vigour Scale from: 0 to 32.	2.44
Psychological status: POMS anxiety Scale from: 0 to 36.	2.15
Psychological status: POMS depression Scale from: 0 to 60.	3.07
Psychological status: POMS anger Scale from: 0 to 48.	2.28
Psychological status: POMS confusion Scale from: 0 to 28.	1.18
Activity levels: rest (hours/day)	0.74
Symptom scales: Wellness score Scale from: not reported.	5.81

Table 52: MID for continuous outcomes (0.5 x SD): 5-HT3 antagonists versus placebo

Outcomes	MID
Fatigue: CIS fatigue Scale from: 20 to 140.	5.75
Activity levels: Actometer (objective accelerometer-based method of measuring activity)	8.95
Symptom scales: Sickness Impact Profile (SIP) Scale from: 0 to 5799.	347.3

Table 53: MID for continuous outcomes (0.5 x SD): Galantamine hydrobromide versus placebo

Outcomes	No of Participants (studies) Follow up
Fatigue: fatigue on VAS Scale from:	0.68
Cognitive function: memory on VAS Scale from:	1.23
Pain: myalgia on VAS Scale from:	0.63
Sleep quality: sleep disturbance on VAS Scale from:	1.25
Adverse events: AEs dizziness on VAS Scale from:	1.56
Return to school/work: work capacity/satisfaction on VAS Scale from:	0.84

Table 54: MID for continuous outcomes (0.5 x SD): Antihistamines versus placebo

Outcomes	MID
Physical functioning: modified Medical Outcome Study Short Form Scale from: 0 to 100.	7.87
Psychological status: modified Medical Outcome Study Short Form - mental health Scale from: 0 to 100.	7.46

Table 55: MID for continuous outcomes (0.5 x SD): Pro-inflammatory cytokine antagonists versus placebo

Outcomes	MID
Fatigue: Checklist Individual Strength fatigue Scale from: 8 to 56.	2
Physical functioning: SF36 physical function Scale from: 0 to 100.	10.5
Psychological status: Symptom Checklist 90 Scale from: 90 to 450.	15.25
Pain: VAS maximum pain score Scale from: 0 to 10.	1
Symptom scales: Sickness Impact Profile Scale from: 0 to 5799.	326.25

Table 56: MID for continuous outcomes (0.5 x SD): Staphylococcus vaccine versus placebo

Outcomes	MID
Pain: pain on VAS Scale from: unclear.	0.98
Symptom scales: clinical global impression of change Scale from: 1 to 7.	0.54
Symptom scales: clinical global impression of severity Scale from: 1 to 7.	0.29

Table 57: MID for continuous outcomes (0.5 x SD): Children and young people: Sympathomimetic/central antihypertensive drugs versus placebo

Outcomes	MID
Fatigue: Chalder Fatigue Questionnaire (CFQ) Scale from: not reported.	3.13
Physical functioning: Fatigue Disability Index (FDI) Scale from: not reported.	4.6
Pain: Brief Pain Inventory average pain score Scale from: 0 to 10.	1.05
Sleep quality: Karolinska Sleep Questionnaire insomnia score Scale from: not reported	0.48
Activity levels: steps per day (accelerometer)	1200.25
Cognitive function: Digit span backward test total Scale from: not reported.	0.98
Symptom scales: CFS symptom inventory hypersensitivity score Scale from: not reported.	0.63

References

1. Agger JL, Schroder A, Gormsen LK, Jensen JS, Jensen TS, Fink PK. Imipramine versus placebo for multiple functional somatic syndromes (STreSS-3): a double-blind, randomised study. The Lancet Psychiatry. 2017; 4(5):378-388

2. Andersson M, Bagby JR, Dyrehag L, Gottfries C. Effects of staphylococcus toxoid vaccine on pain and fatigue in patients with fibromyalgia/chronic fatigue syndrome. European Journal of Pain (London, England). 1998; 2(2):133-142

3. Anonymous. Rituximab for Patients With Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. Annals of Internal Medicine. 2019; 170(9):1-27

4. Arnold LM, Blom TJ, Welge JA, Mariutto E, Heller A. A randomized, placebo-controlled, double-blinded trial of duloxetine in the treatment of general fatigue in patients with chronic fatigue syndrome. Psychosomatics. 2015; 56(3):242-253

5. Ascough C, King H, Serafimova T, Beasant L, Jackson S, Baldock L et al. Interventions to treat pain in paediatric CFS/ME: a systematic review. BMJ Paediatrics Open. 2020; 4(1):e000617

6. Baschetti R. Low-dose hydrocortisone for treatment of chronic fatigue syndrome: A randomized controlled trial (letter). JAMA. 1999; 281(20):1887

7. Behan PO, Hannifah H. 5-HT reuptake inhibitors in CFS. EOS Rivista di Immunologia ed Immunofarmacologia. 1995; 15(1-2):66-69

8. Blacker CV, Greenwood DT, Wesnes KA, Wilson R, Woodward C, Howe I et al. Effect of galantamine hydrobromide in chronic fatigue syndrome: A randomized controlled trial. JAMA. 2004; 292(10):1195-1204

9. Blockmans D, Persoons P, Van Houdenhove B, Bobbaers H. Does methylphenidate reduce the symptoms of chronic fatigue syndrome? American Journal of Medicine. 2006; 119(2):167.e123-130

10. Blockmans D, Persoons P, Van Houdenhove B, Lejeune M, Bobbaers H. Combination therapy with hydrocortisone and fludrocortisone does not improve symptoms in chronic fatigue syndrome: A randomized, placebo-controlled, double-blind, crossover study. American Journal of Medicine. 2003; 114(9):736-741

11. Brook MG, Bannister BA, Weir WR. Interferon-alpha therapy for patients with chronic fatigue syndrome. Journal of Infectious Diseases. 1993; 168(3):791-792

12. Brostoff J. A phase II, randomised, placebo controlled study to assess the safety and efficacy of anti cholineskinase drugs in patients with a diagnosis of chronic fatigue syndrome. 2000. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-00418089/full Last accessed: 30/10/2019.

13. Cleare AJ, Heap E, Malhi G, Wessely S, O'Keane V, Miell J. Randomised, double-blind, placebo-controlled crossover trial of low-dose hydrocortisone in chronic fatigue syndrome. XXIst Collegium Internationale Neuro Psychopharmacologicum, Glasgow, Scotland 12-16 July, 1998. 1998;

14. Cleare AJ, Heap E, Malhi GS, Wessely S, O'Keane V, Miell J. Low-dose hydrocortisone in chronic fatigue syndrome: A randomised crossover trial. Lancet. 1999; 353(9151):455-458

15. Cleare AJ, O'Keane V, Miell J. Plasma leptin in chronic fatigue syndrome and a placebo-controlled study of the effects of low-dose hydrocortisone on leptin secretion. Clinical Endocrinology. 2001; 55(1):113-119

16. Cleare AJ, O'Keane V, Miell JP. Levels of DHEA and DHEAS and responses to CRH stimulation and hydrocortisone treatment in chronic fatigue syndrome. Psychoneuroendocrinology. 2004; 29(6):724-732

17. Cleare AJ, Sookdeo SS, Jones J, O'Keane V, Miell JP. Integrity of the growth hormone/insulin-like growth factor system is maintained in patients with chronic fatigue syndrome. Journal of Clinical Endocrinology and Metabolism. 2000; 85(4):1433-1439

18. Collatz A, Johnston SC, Staines DR, Marshall-Gradisnik SM. A systematic review of drug therapies for chronic fatigue syndrome/myalgic encephalomyelitis. Clinical Therapeutics. 2016; 38(6):1263-1271.e1269

19. De Vinci C, Levine PH, Pizza G, Fudenberg HH, Orens P, Pearson G et al. Lessons from a pilot study of transfer factor in chronic fatigue syndrome. Biotherapy. 1996; 9(1-3):87-90

20. Diaz-Mitoma F, Turgonyi E, Kumar A, Lim W, Larocque L, Hyde BM. Clinical improvement in chronic fatigue syndrome is associated with enhanced natural killer cell-mediated cytotoxicity: The results of a pilot study with Isoprinosine. Journal of Chronic Fatigue Syndrome. 2003; 11(2):71-93

21. Fagermoen E, Sulheim D, Winger A, Andersen AM, Gjerstad J, Godang K et al. Effects of low-dose clonidine on cardiovascular and autonomic variables in adolescents with chronic fatigue: A randomized controlled trial. BMC Pediatrics. 2015; 15:117

22. Fluge O, Bruland O, Risa K, Storstein A, Kristoffersen EK, Sapkota D et al. Benefit from B-lymphocyte depletion using the anti-CD20 antibody rituximab in chronic fatigue syndrome. A double-blind and placebo-controlled study. PloS One. 2011; 6(10):e26358

23. Fluge O, Rekeland IG, Lien K, Thurmer H, Borchgrevink PC, Schafer C et al. B-Lymphocyte depletion in patients with myalgic encephalomyelitis/chronic fatigue syndrome: A randomized, double-blind, placebo-controlled trial. Annals of Internal Medicine. 2019; 170(9):585-593

24. Fluge O, Risa K, Lunde S, Alme K, Rekeland IG, Sapkota D et al. B-Lymphocyte Depletion in Myalgic Encephalopathy/ Chronic Fatigue Syndrome. An Open-Label Phase II Study with Rituximab Maintenance Treatment. PloS One. 2015; 10(7):e0129898

25. Galantamine of no benefit in chronic fatigue syndrome. Pharmaceutical Journal. 2004; 273(7317):375-

26. Glazachev OS, Dudnik E N, Zagaynaya EE. Pharmacological treatment of patients with chronic fatigue syndrome. Zhurnal Nevrologii i Psikhiatrii Imeni SS Korsakova. 2017; 117(4):40-44

27. Gracious B, Wisner KL. Nortriptyline in chronic fatigue syndrome: A double blind, placebo-controlled single case study. Biological Psychiatry. 1991; 30(4):405-408

28. Guo FC, Guo YJ. Clinical research of combined traditional chinese and western medicine in the treatment of chronic fatigue syndrome. China Journal of Chinese Medicine. 2015; 30(1):133-135

29. Hall KT, Kossowsky J, Oberlander TF, Kaptchuk TJ, Saul JP, Wyller VB et al. Genetic variation in catechol-O-methyltransferase modifies effects of clonidine treatment in chronic fatigue syndrome. Pharmacogenomics Journal. 2016; 16(5):454-460

30. Hartz AJ, Bentler SE, Brake KA, Kelly MW. The effectiveness of citalopram for idiopathic chronic fatigue. Journal of Clinical Psychiatry. 2003; 64(8):927-935

31. Hermans L, Nijs J, Calders P, De Clerck L, Moorkens G, Hans G et al. Influence of morphine and naloxone on pain modulation in rheumatoid arthritis, chronic fatigue syndrome/fibromyalgia, and controls: a double-blind, randomized, placebo-controlled, cross-over study. Pain Practice. 2018; 18(4):418-430

32. Hickie I. Nefazodone for patients with chronic fatigue syndrome. Australian and New Zealand Journal of Psychiatry. 1999; 33(2):278-280

33. Hickie I, Lloyd A, Wakefield D. Immunological and psychological dysfunction in patients receiving immunotherapy for chronic fatigue syndrome. Australian and New Zealand Journal of Psychiatry. 1992; 26(2):249-256

34. Hickie IB, Wilson AJ, Wright JM, Bennett BK, Wakefield D, Lloyd AR. A randomized, double-blind placebo-controlled trial of moclobemide in patients with chronic fatigue syndrome. Journal of Clinical Psychiatry. 2000; 61(9):643-648

35. Houdenhove BV, Pae CU, Luyten P. Chronic fatigue syndrome: Is there a role for non-antidepressant pharmacotherapy? Expert Opinion on Pharmacotherapy. 2010; 11(2):215-223

36. Kakumanu S, Mende C, Lehman E, Yeageer M, Craig T. The effect of topical nasal corticosteroids in patients with chronic fatigue syndrome and rhinitis. Journal of Allergy and Clinical Immunology. 2001; 107(2):S153

37. Kakumanu SS, Mende CN, Lehman EB, Hughes K, Craig TJ. Effect of topical nasal corticosteroids on patients with chronic fatigue syndrome and rhinitis. Journal of the American Osteopathic Association. 2003; 103(9):423-427

38. Kreijkamp-Kaspers S, Brenu EW, Marshall S, Staines D, Van Driel ML. Treating chronic fatigue syndrome - a study into the scientific evidence for pharmacological treatments. Australian Family Physician. 2011; 40(11):907-912

39. Lane RJ. A randomised, placebo controlled study to assess safety and efficacy of galantamine hydrobromide in chronic fatigue syndrome. National Research Register. 1999;

40. Lerner AM, Beqaj SH, Deeter RG, Dworkin HJ, Zervos M, Chang CH et al. A six-month trial of valacyclovir in the Epstein-Barr virus subset of chronic fatigue syndrome: Improvement in left ventricular function. Drugs of Today. 2002; 38(8):549-561

41. Lerner AM, Beqaj SH, Deeter RG, Fitzgerald JT. Valacyclovir treatment in Epstein-Barr virus subset chronic fatigue syndrome: Thirty-six months follow-up. In Vivo. 2007; 21(5):707-713

42. Lerner AM, Zervos M, Chang CH, Beqaj S, Goldstein J, O'Neill W et al. A small, randomized, placebo-controlled trial of the use of antiviral therapy for patients with chronic fatigue syndrome. Clinical Infectious Diseases. 2001; 32(11):1657-1658

43. Liu CZ, Lei B. Effect of tuina on oxygen free radicals metabolism in patients with chronic fatigue syndrome. Chinese Acupuncture & Moxibustion. 2010; 30(11):946-948

44. Lloyd A, Hickie I, Wakefield D, Boughton C, Dwyer J. A double-blind, placebo-controlled trial of intravenous immunoglobulin therapy in patients with chronic fatigue syndrome. American Journal of Medicine. 1990; 89(5):561-568

45. Lloyd AR, Hickie I, Brockman A, Hickie C, Wilson A, Dwyer J et al. Immunologic and psychologic therapy for patients with chronic fatigue syndrome: A double-blind, placebo-controlled trial. American Journal of Medicine. 1993; 94(2):197-203

46. Main J. A phase II randomised placebo controlled study to assess the safety and efficacy of galantamine hydrobromide 25mg tid and 10mg tid taken for a period of 16 wks in patients with a diagnosis of chronic fatigue syndrome (MREC). National Research Register. 2000;

47. McKenzie R, O'Fallon A, Dale J, Demitrack M, Sharma G, Deloria M et al. Low-dose hydrocortisone for treatment of chronic fatigue syndrome: A randomized controlled trial. JAMA. 1998; 280(12):1061-1066

48. McKenzie R, Reynolds JC, O'Fallon A, Dale J, Deloria M, Blackwelder W et al. Decreased bone mineral density during low dose glucocorticoid administration in a randomized, placebo controlled trial. Journal of Rheumatology. 2000; 27(9):2222-2226

49. Meeus M, Ickmans K, Struyf F, Hermans L, Van Noesel K, Oderkerk J et al. Does acetaminophen activate endogenous pain inhibition in chronic fatigue syndrome/fibromyalgia and rheumatoid arthritis? A double-blind randomized controlled cross-over trial. Pain Physician. 2013; 16(2):E61-70

50. Mehta VK, Blume GB. A randomized trial of fluoxetine in a patient with persistent fatigue. Journal of the American Board of Family Practice. 1995; 8(3):230-232

51. Mitchell AJ. A phase II randomised, placebo-controlled study to assess the safety and efficacy of galantamine hydrobromide 25mg TID , 5mg TID, 75mg TID and 10mg TID taken for a period of 16 weeks in patients with a diagnosis of chronic fatigue syndrome (CFS). National Research Register. 2000;

52. Montoya JG, Anderson JN, Adolphs DL, Bateman L, Klimas N, Levine SM et al. KPAX002 as a treatment for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): A prospective, randomized trial. International Journal of Clinical and Experimental Medicine. 2018; 11(3):2890-2900

53. Montoya JG, Kogelnik AM, Bhangoo M, Lunn MR, Flamand L, Merrihew LE et al. Randomized clinical trial to evaluate the efficacy and safety of valganciclovir in a subset of patients with chronic fatigue syndrome. Journal of Medical Virology. 2013; 85(12):2101-2109

54. Moorkens G, Wynants H, Abs R. Effect of growth hormone treatment in patients with chronic fatigue syndrome: A preliminary study. Growth Hormone and IGF Research. 1998; 8(Suppl B):131-133

55. Morriss R, Wearden A, Mullis R, Strickland P, Appleby L, Campbell I et al. A double-blind placebo controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome (CFS). 8th Congress of the Association of European Psychiatrists London, UK 7-12 July. 1996;

56. Morriss RK, Robson MJ, Deakin JF. Neuropsychological performance and noradrenaline function in chronic fatigue syndrome under conditions of high arousal. Psychopharmacology. 2002; 163(2):166-173

57. Natelson BH, Cheu J, Hill N, Bergen M, Korn L, Denny T et al. Single-blind, placebo phase-in trial of two escalating doses of selegiline in the chronic fatigue syndrome. Neuropsychobiology. 1998; 37(3):150-154

58. Natelson BH, Cheu J, Pareja J, Ellis SP, Policastro T, Findley TW. Randomized, double blind, controlled placebo-phase in trial of low dose phenelzine in the chronic fatigue syndrome. Psychopharmacology. 1996; 124(3):226-230

59. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [Updated 2018]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

60. Nilsson MKL, Zachrisson O, Gottfries CG, Matousek M, Peilot B, Forsmark S et al. A randomised controlled trial of the monoaminergic stabiliser (-)-OSU6162 in treatment of myalgic encephalomyelitis/chronic fatigue syndrome. Acta Neuropsychiatrica. 2018; 30(3):148-157

61. Olson LG, Ambrogetti A, Sutherland DC. A pilot randomized controlled trial of dexamphetamine in patients with chronic fatigue syndrome. Psychosomatics. 2003; 44(1):38-43

62. Pae CU, Marks DM, Patkar AA, Masand PS, Luyten P, Serretti A. Pharmacological treatment of chronic fatigue syndrome: Focusing on the role of antidepressants. Expert Opinion on Pharmacotherapy. 2009; 10(10):1561-1570

63. Pardini M, Guida S, Primavera A, Krueger F, Cocito L, Gialloreti LE. Amisulpride vs. fluoxetine treatment of chronic fatigue syndrome: A pilot study. European Neuropsychopharmacology. 2011; 21(3):282-286

64. Peterson PK, Pheley A, Schroeppel J, Schenck C, Marshall P, Kind A et al. A preliminary placebo-controlled crossover trial of fludrocortisone for chronic fatigue syndrome. Archives of Internal Medicine. 1998; 158(8):908-914

65. Peterson PK, Shepard J, Macres M, Schenck C, Crosson J, Rechtman D et al. A controlled trial of intravenous immunoglobulin G in chronic fatigue syndrome. American Journal of Medicine. 1990; 89(5):554-560

66. Plioplys AV, Plioplys S. Amantadine and L-carnitine treatment of Chronic Fatigue Syndrome. Neuropsychobiology. 1997; 35(1):16-23

67. Raijmakers RPH, Keijmel SP, Breukers EMC, Bleijenberg G, van der Meer JWM, Bleeker-Rovers CP et al. Long-term effect of cognitive behavioural therapy and doxycycline treatment for patients with Q fever fatigue syndrome: One-year follow-up of the Qure study. Journal of Psychosomatic Research. 2019; 116:62-67

68. Randall DC, Cafferty FH, Shneerson JM, Smith IE, Llewelyn MB, File SE. Chronic treatment with modafinil may not be beneficial in patients with chronic fatigue syndrome. Journal of Psychopharmacology. 2005; 19(6):647-660

69. Rekeland IG, Fosså A, Lande A, Ktoridou-Valen I, Sørland K, Holsen M et al. Intravenous cyclophosphamide in myalgic encephalomyelitis/chronic fatigue syndrome. An open-label phase II study. Frontiers in Medicine. 2020; 7:162

70. Roerink ME, Bredie SJH, Heijnen M, Dinarello CA, Knoop H, Van der Meer JWM. Cytokine inhibition in patients with chronic fatigue syndrome: A randomized trial. Annals of Internal Medicine. 2017; 166(8):557-564

71. Roerink ME, Knoop H, Bredie SJ, Heijnen M, Joosten LA, Netea MG et al. Cytokine inhibition in chronic fatigue syndrome patients: Study protocol for a randomized controlled trial. Trials. 2015; 16:439

72. Roerink ME, Knoop H, Bronkhorst EM, Mouthaan HA, Hawinkels LJAC, Joosten LAB et al. Cytokine signatures in chronic fatigue syndrome patients: A case control study and the effect of anakinra treatment. Journal of Translational Medicine. 2017; 15:267

73. Roerink ME, Roerink S, Skoluda N, van der Schaaf ME, Hermus A, van der Meer JWM et al. Hair and salivary cortisol in a cohort of women with chronic fatigue syndrome. Hormones and Behavior. 2018; 103:1-6

74. Rowe KS. Double-blind randomized controlled trial to assess the efficacy of intravenous gammaglobulin for the management of chronic fatigue syndrome in adolescents. Journal of Psychiatric Research. 1997; 31(1):133-147

75. Rowe KS. Five-year follow-up of young people with chronic fatigue syndrome following the double blind randomised controlled intravenous gammaglobulin trial. Journal of Chronic Fatigue Syndrome. 1999; 5(3-4):97-107

76. Rowe PC. Myalgic encephalomyelitis/chronic fatigue syndrome: trial fails to confirm earlier observations of rituximab's effectiveness. Annals of Internal Medicine. 2019; 170(9):656-657

77. Rowe PC, Calkins H, DeBusk K, McKenzie R, Anand R, Sharma G et al. Fludrocortisone acetate to treat neurally mediated hypotension in chronic fatigue syndrome: A randomized controlled trial. JAMA. 2001; 285(1):52-59

78. See DM, Tilles JG. Alpha-Interferon treatment of patients with chronic fatigue syndrome. Immunological Investigations. 1996; 25(1-2):153-164

79. Snorrason E, Geirsson A, Stefansson K. Trial of a selective acetylcholinesterase inhibitor, galanthamine hydrobromide, in the treatment of chronic fatigue syndrome. Journal of Chronic Fatigue Syndrome. 1996; 2(2-3):35-54

80. Spath M, Welzel D, Farber L. Treatment of chronic fatigue syndrome with 5-HT3 receptor antagonists--preliminary results. Scandinavian Journal of Rheumatology. 2000; 29(113):72-77

81. Spironolactone as treatment for chronic fatigue syndrome in patients with positive epstein bar virus serology. Internal Medicine Journal. 2020; 50(S1):19-19

82. Springer RE, Wray BB, Latham JE. A double blind, placebo controlled study of intravenous gammaglobulin (GG) in the treatment of chronic fatigue syndrome (CFS). Annals of Allergy. 1992; 68:76

83. Staud R, Kizer T, Robinson ME. Muscle injections with lidocaine improve resting fatigue and pain in patients with chronic fatigue syndrome. Journal of Pain Research. 2017; 10:1477-1486

84. Steinberg P, McNutt BE, Marshall P, Schenck C, Lurie N, Pheley A et al. Double-blind placebo-controlled study of the efficacy of oral terfenadine in the treatment of chronic fatigue syndrome. Journal of Allergy and Clinical Immunology. 1996; 97(1 Pt 1):119-126

85. Stouch BC, Strayer D, Carter W. Cardiac toxicity in chronic fatigue syndrome: Results from a randomized 40-week multicenter double-blind placebo control trial of rintatolimod. Journal of Applied Research. 2010; 10(3):80-87

86. Straus SE, Dale JK, Tobi M, Lawley T, Preble O, Blaese RM et al. Acyclovir treatment of the chronic fatigue syndrome. Lack of efficacy in a placebo-controlled trial. New England Journal of Medicine. 1988; 319(26):1692-1698

87. Strayer DR, Carter WA, Brodsky I, Cheney P, Peterson D, Salvato P et al. A controlled clinical trial with a specifically configured RNA drug, poly(I).poly(C12U), in chronic fatigue syndrome. Clinical Infectious Diseases. 1994; 18(Suppl 1):S88-95

88. Strayer DR, Carter WA, Stouch BC, Stevens SR, Bateman L, Cimoch PJ et al. A double-blind, placebo-controlled, randomized, clinical trial of the TLR-3 agonist rintatolimod in severe cases of chronic fatigue syndrome. PloS One. 2012; 7(3):e31334

89. Stubhaug B, Lie SA, Ursin H, Eriksen HR. Cognitive-behavioural therapy v. mirtazapine for chronic fatigue and neurasthenia: Randomised placebo-controlled trial. British Journal of Psychiatry. 2008; 192(3):217-223

90. Suhadolnik RJ, Reichenbach NL, Hitzges P, Adelson ME, Peterson DL, Cheney P et al. Changes in the 2-5A synthetase/RNase L antiviral pathway in a controlled clinical trial with poly(I)-poly(C12U) in chronic fatigue syndrome. In Vivo. 1994; 8(4):599-604

91. Sulheim D, Fagermoen E, Winger A, Andersen AM, Godang K, Müller F et al. Disease mechanisms and clonidine treatment in adolescent chronic fatigue syndrome: A combined cross-sectional and randomized clinical trial. JAMA pediatrics. 2014; 168(4):351-360

92. Teitelbaum JE, Bird B, Greenfield RM, Weiss A, Muenz L, Gould L. Effective treatment for fibromyalgia (FMS) and chronic fatigue syndrome (CFIDS) - A placebo-controlled study. Journal of the chronic fatigue syndrome. 1999; 5(3-4):116-117

93. The GK, Bleijenberg G, Buitelaar JK, van der Meer JW. The effect of ondansetron, a 5-HT3 receptor antagonist, in chronic fatigue syndrome: A randomized controlled trial. Journal of Clinical Psychiatry. 2010; 71(5):528-533

94. The GK, Verkes RJ, Fekkes D, Bleijenberg G, van der Meer JW, Buitelaar JK. Tryptophan depletion in chronic fatigue syndrome, a pilot cross-over study. BMC Research Notes. 2014; 7:650

95. Tiev KP, Cabane J, Imbert JC. Treatment of chronic postinfectious fatigue: Randomized double-blind study of two doses of sulbutiamine (400-600 mg/day) versus placebo. La Revue de Medecine Interne. 1999; 20(10):912-918

96. Vedhara K, Llewelyn MB, Fox JD, Jones M, Jones R, Clements GB et al. Consequences of live poliovirus vaccine administration in chronic fatigue syndrome. Journal of Neuroimmunology. 1997; 75(1-2):183-195

97. Vercoulen JHMM, Swanink CMA, Zitman FG, Vreden SGS, Hoofs MPE, Fennis JFM et al. Randomised, double-blind, placebo-controlled study of fluoxetine in chronic fatigue syndrome. Lancet. 1996; 347(9005):858-861

98. Vercoulen JHMM, Zitman FG, Fennis JFM, Galama JMD, Van-Der-Meer JWM, Bleijenberg G. No effect of fluoxetine in chronic fatigue syndrome; randomized double-blind placebo-controlled trial. Nederlands Tijdschrift voor Geneeskunde. 1997; 141(31):1531-1535

99. Vollmer-Conna U, Hickie I, Hadzi-Pavlovic D, Tymms K, Wakefield D, Dwyer J et al. Intravenous immunoglobulin is ineffective in the treatment of patients with chronic fatigue syndrome. American Journal of Medicine. 1997; 103(1):38-43

100. Vorob'eva OV, Rusaya VV. Efficacy and safety of noofen in the treatment of chronic fatigue syndrome in patients with cerebrovascular failure. Neuroscience and Behavioral Physiology. 2019; 49(2):246-251

101. Vorob'Eva OV, Rusaya VV. Efficacy and safety of noophen in the treatment of chronic fatigue syndrome in patients with cerebrovascular insufficiency. Zhurnal nevrologii i psihiatrii imeni SS Korsakova. 2017; 117(11):31-36

102. Wearden A, Morriss R, Mullis R, Strickland P, Pearson D, Appleby L et al. A double-blind, placebo-controlled treatment trial of fluoxetine and a graded exercise programme for chronic fatigue syndrome. Anglo Portuguese consultation liaison psychiatry conference, Lisbon. 1996;

103. Wearden AJ, Morriss RK, Mullis R, Strickland PL, Pearson DJ, Appleby L et al. Randomised, double-blind, placebo-controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome. British Journal of Psychiatry. 1998; 172:485-490

104. Williams G, Waterhouse J, Mugarza J, Minors D, Hayden K. Therapy of circadian rhythm disorders in chronic fatigue syndrome: No symptomatic improvement with melatonin or phototherapy. European Journal of Clinical Investigation. 2002; 32(11):831-837

105. Wilson C. Promising chronic fatigue drug fails trials. 2019. Available from: https://www.newscientist.com/article/2199217-promising-treatment-for-chronic-fatigue-syndrome-fails-large-trial/ Last accessed: 04/02/2020.

106. Young JL. Use of lisdexamfetamine dimesylate in treatment of cognitive impairment (chronic fatigue syndrome): A double blind, placebo controlled study. 2010. Available from: https://clinicaltrials.gov/ct2/show/NCT01071044 Last accessed: 28/10/2019.

107. Young JL. Use of lisdexamfetamine dimesylate in treatment of executive functioning deficits and chronic fatigue syndrome: A double blind, placebo-controlled study. Psychiatry Research. 2013; 207(1-2):127-133

108. Zachrisson O, Colque-Navarro P, Gottfries CG, Regland B, Mollby R. Immune modulation with a staphylococcal preparation in fibromyalgia/ chronic fatigue syndrome: Relation between antibody levels and clinical improvement. European Journal of Clinical Microbiology and Infectious Diseases. 2004; 23(2):98-105

109. Zachrisson O, Regland B, Jahreskog M, Jonsson M, Kron M, Gottfries CG. Treatment with staphylococcus toxoid in fibromyalgia/chronic fatigue syndrome--A randomised controlled trial. European Journal of Pain. 2002; 6(6):455-466