Management of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS): An Updated Systematic Evidence Review

A Systematic Evidence Review prepared for:
Centers for Disease Control and Prevention
April 11, 2022

Contract No. 75D30118P03408

Investigators:
Roger Chou, MD, FACP
Marian McDonagh, PharmD
Jessica C. Griffin, MS
Sara Grusing, MPH

Pacific Northwest Evidence-based Practice Center
Roger Chou, MD, Director
Marian McDonagh, PharmD, Co-Director

Structured Abstract

Objectives. Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) can have profound effects on function and quality of life. This report updates a 2014 Agency for Healthcare Research Quality (AHRQ)-funded review and 2016 addendum in order to synthesize the evidence on evaluation and management of ME/CFS. It also expands upon the prior AHRQ review by including children as well as adults, evaluating harms as well as benefits of diagnosis, and evaluating effects of treatment on depression, anxiety, sleep quality, pain, and other symptoms associated with ME/CFS in addition to fatigue, function, and quality of life.

Data Sources. MEDLINE (1988 to January 2019), PsycINFO (1988 to January 2019), Embase (through January 2019) and the Cochrane Library (through January 2019); supplemented by review of reference lists and the 2014 AHRQ review. Searches were updated through February 16, 2021.

Review Methods. Articles were selected for review if they included: 1) evaluation of patients with fatigue, 2) diagnosis of ME/CFS, or 3) treatments (pharmacological, nonpharmacological, dietary, or complementary and alternative therapies) of ME/CFS. We abstracted data on the frequency of non-ME/CFS conditions in patients presenting with fatigue; benefits and harms of diagnosis of ME/CFS versus non-diagnosis; and benefits and harms of treatments. Two investigators reviewed abstracts and full-text articles for inclusion based on predefined criteria. Risk of bias was assessed using predefined criteria. Discrepancies were resolved through discussion and consensus, with a third investigator if needed. Random effects meta-analyses were conducted on trials of exercise and cognitive therapy; where evidence was unsuitable for combining, it was synthesized qualitatively. The strength of evidence was assessed using methods recommended by the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.

Results. We identified 5,525 potentially relevant articles, selected 687 for full-text review, and included 73 studies in 91 publications (1 systematic review and 6 observational studies on diagnosis and 66 trials of treatments). A systematic review of patients with fatigue or tiredness in primary care settings found that the most common non-ME/CFS conditions were depression (18.5%), serious somatic diseases (4.3%), anemia (2.8%), and malignancy (0.6%). In specialty settings of patients referred for evaluation of possible ME/CFS, the most common non-ME/CFS conditions were psychological (15% to 51%) and sleep disorders (6% to 30%). No study evaluated benefits or harms of ME/CFS diagnosis versus non-diagnosis.

Sixty-six trials evaluated treatments for ME/CFS. Thirty-three trials were included in the prior AHRQ report and 33 trials were new since the prior report. Cognitive Behavioral Therapy (CBT) and exercise therapy were associated with improved fatigue, function, and other outcomes versus inactive control therapies, but the magnitude of effects based on average benefits was small to moderate. These trials demonstrated unexplained statistical heterogeneity in pooled estimates and contained methodological limitations. Additionally, trials varied in the ME/CFS case definitions used, though findings were similar, with no statistically significant differences, when analyses were stratified according to the case definition used. However, stratified analyses based on case definitions were limited by the small number of trials and few trials used more current ME/CFS case definitions (including definitions requiring presence of post-exertional malaise); moreover, trials may not have evaluated patients with more severe ME/CFS. Therefore,

the applicability of findings to patients with severe ME/CFS, ME/CFS diagnosed using more current case definitions, or post-exertional malaise was uncertain. Other pharmacological, nonpharmacological, dietary, and complementary and alternative therapies were ineffective, or evidence of effectiveness was too limited to reliably evaluate benefits and harms. Reporting of harms across trials was suboptimal, with limited evidence that exercise and CBT were not associated with increased risk of serious adverse events or worsening of symptoms. In adolescents with ME/CFS, limited evidence found CBT (family based or involving parents) associated with improved function and school attendance versus inactive therapies, but differences were not statistically significant.

Limitations. Treatment trials had methodological limitations. Most interventions and comparisons were evaluated in few trials, most trials used older ME/CFS case definitions (including case definitions that did not require post-exertional malaise), and there was limited information on how key characteristics and subgroups of patients impacted outcomes. There was unexplained statistical heterogeneity in meta-analyses, study inclusion was restricted to English language publications and formal methods for determining small sample effects were not performed due to small numbers of studies.

Conclusions. This report summarizes the literature through February 2022, documents the gaps and limitations in published clinical trials, and provides evidence that well-designed trials of therapy for ME/CFS utilizing more current ME/CFS case definitions are needed. Evidence on effective treatments for ME/CFS remains limited. The strength of evidence supporting the use of graded exercise and CBT was low and the magnitude of benefits was small to moderate, with inadequate evidence in patients diagnosed with more current case definitions, limited reporting of harms, and inadequate evaluation in severely affected patients. Methodological and other limitations (imprecision, inconsistency, uncertain generalizability) preclude strong conclusions. Other therapies were not shown to be effective or require additional evidence to determine effectiveness.

Contents

List of Tables

List of Figures

List of Appendices

Introduction

This report was commissioned by the Centers for Disease Control and Prevention (CDC) to synthesize the evidence from a systematic review on evaluation and management of myalgic encephalomyelitis (ME)/chronic fatigue syndrome (CFS). It builds upon and updates a 2014 Agency for Healthcare Research and Quality (AHRQ) review and subsequent 2016 addendum that was conducted to support a National Institutes of Health Pathways to Prevention conference,[1]

This report was to inform the development of a guideline on treatment and management of ME/CFS. However, CDC ME/CFS program recognizes that the systematic review did not provide enough evidence to move forward with treatment recommendations at this time. So, despite the fact that improving clinical care remains a critical issue, it is in the best interest to not proceed with developing the guideline. In the interest of transparency, the final report of this systematic review, comments received, and responses to these comments will be posted on CDC ME/CFS website. The posting of the systematic review is to make the information available should others in the ME/CFS field undertake a review of the literature in the future.

Background

ME/CFS is a condition characterized by a constellation of symptoms. Hallmarks of ME/CFS are post-exertional malaise and/or persistent and disabling fatigue, as well as various additional manifestations, including pain, sleep disturbance, orthostatic intolerance, motor impairment, neurological and cognitive manifestations (i.e., impaired concentration, mental processing, and memory), and altered immune and autonomic responses.[2-5] ME/CFS often follows a chronic or relapsing and remitting course and may result in reduced quality of life and loss of independence.[1,6] In 2015, the Institute of Medicine (IOM) recommended renaming the condition to systemic exertion intolerance disease.[5] However, the terms ME and CFS continue to be used, and will be used in this report.

Although similar well-described symptom clusters were reported as early as the 1930s,[7] the term "myalgic encephalomyelitis (ME)" was first used to describe the condition in the 1950s and ME was recognized by the World Health Organization as a disease entity in the 1960s.[8] The term "chronic fatigue syndrome (CFS)" was coined in the 1980s after research failed to identify a clear viral association in what was previously labeled chronic Epstein-Barr virus syndrome.[9-12] Although the terms ME and CFS are often used together or interchangeably, ME may be a subset of CFS or its own distinct disease.

Many case definitions for ME and CFS have been proposed (**Table 1** shows commonly used case definitions and the names we used to refer to them in this report).[13] The first case definitions for ME/CFS were published in the 1980's.[10,13] Other case definitions have been introduced over the years, including the 1991 Oxford criteria,[14] the 1994 Fukuda criteria,[4] the 2003 Canadian criteria,[2] the 2007 National Institute for health and Care Excellence (NICE) criteria,[15] and the 2011 international consensus criteria.[3] An important distinction between case definitions is whether presence of post-exertional malaise is required for diagnosis. Trials of treatments for ME/CFS have primarily utilized older case definitions (Oxford, Fukuda, or NICE) that do not require presence of post-exertional malaise. The 2011 international consensus report advocates for use of the term ME over CFS, to better reflect an underlying pathophysiology involving widespread inflammation and neuropathology, though this position has not been accepted by all.[3]

The proposed IOM definition for systemic exertion intolerance disease requires substantial reduction or impairment in ability to engage in pre-illness levels of activity; post-exertional malaise; unrefreshing sleep; and either cognitive impairment or orthostatic intolerance.[5] The use of multiple case definitions for ME/CFS is an ongoing challenge in the field, as it has resulted in heterogeneous populations in the research literature. For example, a systematic review found that median ME/CFS prevalence varied in studies that used different case definitions.[16] For example, one study found that the prevalence of ME/CFS was 15 times higher in studies that used the earlier Oxford case definition (1.5%), which does not require presence of post-exertional malaise, compared with those that used the Canadian case definition, which does.[16] However, interpretation of these findings is difficult. As noted in the prior AHRQ report, studies have not been able to determine the accuracy of different ME/CFS case definitions, due to the lack of a reliable and universally accepted reference ("gold") standard, methodological limitations in the literature, and limited evidence for some case definitions (including more recently introduced case definitions).[1,16,17]

The IOM report estimates ME/CFS prevalence in the U.S. between 836,000 and 2.5 million.[5] However, the prevalence of ME/CFS is difficult to estimate given the uncertainty and variability in case definitions and differences in study methodology. Even with these challenges in estimating ME/CFS prevalence, it is estimated that as many as 84 to 91 percent of persons with ME/CFS have not been diagnosed.[5] ME/CFS is more common among women than men, with an average age at diagnosis of between 30 and 40 years of age.[18] The prevalence of ME/CFS and pattern of symptoms in children appears similar to adults, though in children an antecedent acute flu- or mononucleosis-like syndrome may be more frequently present and the prognosis may be more favorable.[19-21] Data suggest that about 40 percent (8 to 63%) of adult patients with ME/CFS improve but only 5 percent (0 to 31%) fully recover,[22] compared to recovery in over 50 percent of children within 6 months.[20] However, more research is needed to understand the prognosis of ME/CFS in children.

The goal of treatment for ME/CFS is to reduce symptoms and improve function and quality of life. Although a number of medications have been used to treat ME/CFS, no medication is approved by the U.S. Food and Drug Administration (FDA) for ME/CFS. Treatments for ME/CFS fall into two broad categories: those intended to treat the underlying cause of the disease (pathogenesis-based therapies) and those targeting ME/CFS symptoms (symptom-based therapies).[23] The first category includes immune modulators (e.g., rintatolimod, immunoglobulin, rituximab, and corticosteroids), antiviral and antibiotic medications, and other medications. Symptom-based therapies include medications to treat fatigue, sleep dysfunction, pain, mood disorders, and other symptoms associated with ME/CFS, as well as non-drug therapies such as yoga, stretching and relaxation techniques, mindfulness based training, graded exercise, pacing strategies, cognitive behavioral and other psychological therapies, dietary supplements and interventions, and various complementary and alternative therapies.[23] In practice, there are wide variations in the clinical management of patients with ME/CFS, and many patients receive multiple therapies in various combinations and sequences.

The prior AHRQ report found limited evidence that graded exercise therapy (GET) and counseling therapies (primarily cognitive behavioral therapy [CBT]) were associated with beneficial effects on fatigue and function in some patients, but found that these therapies had not been adequately tested in patients with more severe ME/CFS or in patients identified using more current case definitions.[1] It also found limited evidence that rintatolimod was associated with improved exercise performance in some patients. There was insufficient evidence to determine

the effectiveness of other therapies. The CDC commissioned a review to incorporate new research and address research gaps identified in the 2014 AHRQ report.

Table 1. Case definitions or diagnostic criteria

Case Definition Or Diagnostic Criteria	Reference	Population
Ramsay, 1986[11]	Ramsay M. Postviral fatigue syndrome: The saga of royal free disease. London: Gower Medical; 1986.	Adults
Holmes, 1988[10]	Holmes GP, Kaplan JE, Gantz NM, et al. Chronic fatigue syndrome: a working case definition. Ann Intern Med. 1988;108(3):387-9.	Adults
Oxford Sharpe, 1991[14]	Sharpe MC, Archard LC, Banatvala JE, et al. A report-chronic fatigue syndrome: guidelines for research. J R Soc Med. 1991;84(2):118-21.	Adults
London ME Dowsett, 1994[24]	Dowsett E, Goudsmit E, Macintyre A, et al. Report from the national task force on chronic fatigue syndrome (CFS), post viral fatigue syndrome (PVFS), myalgic encephalomyelitis (ME). Westcare. 1994.	Adults
Fukuda, 1994[4]	Fukuda K, Straus SE, Hickie I, et al. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med. 1994;121(12):953-9.	Adults
Canadian ME/CFS Carruthers, 2003[2]	Carruthers BM, Jain AK, de Meirleir KL, et al. Myalgic encephalomyelitis/chronic fatigue syndrome: clinical working case definition, diagnostic and treatment protocols. J Chronic Fatigue Syndr. 2003;11(1):7-115.	Adults Children
NICE, 2007[15]	National Institute for Clinical Excellence. Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): Diagnosis and management of CFS/ME in adults and children. NHS National Institute for Health and Clinical Excellence, 7. 2007; NICE clinical guideline 53; 2007.	Adults Children
International Pediatric Jason, 2006[25]	Jason LA, Bell DS, Rowe K, et al. A pediatric case definition for myalgic encephalomyelitis and chronic fatigue syndrome. Journal of Chronic Fatigue Syndrome. 2006;13(2-3):1-44.	Children
Revised Canadian ME/CFS Jason, 2010[26]	Jason L, Evans M, Porter N, et al. The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. Am J Biochem Biotechnol. 2010;6(2):120-35.	Adults
International ME Carruthers, 2011[3]	Carruthers BM, van de Sande MI, De Meirleir KL, et al. Myalgic encephalomyelitis: International Consensus Criteria. J Intern Med. 2011;270(4):327-38.	Adults Children
IOM ME/CFS criteria 2015[5]	Institute of Medicine. Beyond myalgic encephalomyelitis/chronic fatigue syndrome: Redefining an illness. Washington, DC: National Academies Press; US; 2015.	Adults Children

Abbreviations: CFS = chronic fatigue syndrome; IOM = Institute of Medicine; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence.

Purpose

The purpose of this systematic review is to synthesize the evidence on benefits and harms of treatment for ME/CFS; benefits and harms of diagnosing ME/CFS; and the prevalence of non-ME/CFS conditions in persons presenting for evaluation of potential ME/CFS.

Methods

This systematic review follows the methods of the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[27]

Topic Development and Refinement

The scope and key questions used to guide the current review were developed by the Pacific Northwest Evidence-based Practice Center (EPC) with input from the CDC and eight Key Informants representing clinical, research, or patient perspectives in ME/CFS. The protocol for this review was registered in the PROSPERO international database of prospectively registered systematic reviews.[28] The following key questions were used to guide this report:

Key Question 1: In patients undergoing evaluation for possible ME/CFS, what is the frequency of non-ME/CFS conditions?
Key Question 2: What are the benefits and harms of diagnosing ME/CFS, versus non-diagnosis?
Key Question 3: What are the benefits and harms of therapeutic interventions for patients with ME/CFS, and how do they vary by patient subgroups?

> *Key Question 3a:* Interventions for treating ME/CFS
> *Key Question 3b:* Interventions for treating symptoms commonly present in persons with ME/CFS (fatigue, poor sleep, orthostatic intolerance, pain, cognitive problems, depression, multiple chemical sensitivity, gastrointestinal symptoms, urinary symptoms, etc.)

The scope of this report differs from the prior report in several ways. Whereas the prior report focused on adults with ME/CFS, this update also includes children. The prior report included questions on the accuracy and concordance of case definitions and diagnostic criteria used to diagnosis ME/CFS. This update does not address diagnostic accuracy of case definitions for ME/CFS, due to the lack of a reliable, universally accepted reference standard, which is necessary to estimate diagnostic accuracy. Instead, this report addresses a new Key Question on the frequency of non-ME/CFS conditions (without a diagnosis of ME/CFS) in persons presenting for evaluation for possible ME/CFS (Key Question 1). The prior report included a question on the harms of diagnosing ME/CFS and included qualitative and noncomparative studies. This update addresses both the benefits and harms of diagnosis (Key Question 2), in order to present a more balanced perspective, and is restricted to comparative studies that assessed outcomes in persons diagnosed and not diagnosed with ME/CFS. Finally, for evaluation of ME/CFS treatments the prior report focused on effects on fatigue, function, and quality of life. This update also evaluates effects of treatments on other outcomes (depression, anxiety, sleep quality, pain, and others [e.g., cognitive functioning, gastrointestinal symptoms, orthostatic intolerance, and symptoms associated with multiple chemical sensitivity]). Because numerous trials evaluated outcomes addressed in Key Questions 3a and 3b, we report results for both sub-questions in the same section. This report also seeks to determine how effects of ME/CFS treatments varied in

subgroups defined by patient characteristics, including the ME/CFS case definition used, severity of symptoms, duration of symptoms, type of onset (e.g., sudden versus gradual), demographic factors, and others.

Data Sources and Searches

A research librarian conducted searches in Ovid MEDLINE (1988 to January 9, 2019), PsycINFO (1988 to January Week 1 2019), the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through January 9, 2019), and Embase (through January 11, 2019) (search strategies shown in **Appendix A**). We supplemented searches of electronic databases with review of reference lists of relevant studies. We also reviewed the excluded studies list of the prior AHRQ report to identify studies potentially relevant to the revised scope of this update. Searches were updated on February 16, 2021.

Process for Study Selection

Criteria for inclusion and exclusion of studies were developed for the Key Questions using the populations, interventions, comparators, outcomes, timing, and setting/study design (PICOTS) framework (**Appendix B**). Articles were selected for review if they were about evaluation of patients with fatigue, diagnosis of ME/CFS, or treatment of ME/CFS in adults or children; were relevant to a Key Question; and met the prespecified inclusion criteria. Studies of nonhuman subjects and studies without original data were excluded. Abstracts were independently reviewed by two investigators and full-text articles were obtained for all studies that either investigator classified as potentially meeting inclusion criteria. Discrepancies were resolved through discussion and consensus, with a third investigator to resolve discrepancies if necessary. Two investigators independently reviewed all full-text articles for final inclusion. Inclusion was restricted to English-language articles. A list of the included studies appears in **Appendix C;** a list of excluded studies and primary reasons for exclusion can be found in **Appendix D.**

Inclusion criteria

For Key Question 1, we included systematic reviews and cohort studies of adults or children presenting with possible ME/CFS due to fatigue or post-exertional malaise that reported the proportion of patients with non-CFS symptoms/conditions. We excluded studies on the prevalence of symptoms in patients diagnosed with ME/CFS, which was not the topic of this Key Question.

For Key Question 2, we included randomized trials and cohort studies of patients presenting with fatigue or post-exertional malaise. The studies compared those diagnosed with ME/CFS versus those not given an ME/CFS diagnosis and reported any potential benefit or harm from diagnosis (including access to treatment, psychological harms, labeling, risk from diagnostic testing, misdiagnosis, or other). We also included studies that evaluated these outcomes before and after diagnosis of ME/CFS.

For Key Question 3, we included randomized trials of patients diagnosed with ME/CFS using published case definitions. Included interventions were various forms of counseling and behavioral therapy (e.g., CBT, cognitive therapy, relaxation, mindfulness-based stress reduction, biofeedback), exercise (e.g., graded exercise or anaerobic exercise), adaptive pacing, orthostatic training, complementary and alternative therapies (e.g., acupuncture, massage, tuina, Qigong,

distant healing, and others), pathogenesis-based medications (e.g., immune modulators, antivirals, or antibiotics), and symptom-based medications (beta blockers, antidepressants, anxiolytics, stimulants, mineralocorticoids, ivabradine, and others). Trials compared an included intervention versus inactive treatment (defined as placebo, no treatment, usual care/usual specialist care, wait list, or an attention control) or versus another included intervention. Wait list refers to trials in which the inactive treatment is delayed initiation of the studied intervention. Attention controls are not intended to have an important therapeutic effect but control for some of the attentional and time aspects of active therapy (e.g., in a trial with CBT as the active intervention, simple education or advice without a cognitive behavioral component). Outcomes were continuous measures of fatigue, function, quality of life, school attendance (children), sleep, depression, anxiety, and other outcomes associated with specific ME/CFS-associated symptoms (gastrointestinal, autonomic dysfunction, orthostatic intolerance, urinary symptoms, symptoms associated with multiple chemical sensitivity). The review focuses on patient-reported outcomes; however, we included the 6-minute walk test, as it is the most commonly reported "objective" measure of function in trials of treatments for ME/CFS. We also included dichotomous measures for improvement in fatigue, improvement in function, overall efficacy, and recovery, as these outcomes were defined in the trials. Harms were serious adverse events, withdrawals due to harms, withdrawal due to symptom worsening, post-exertional malaise, worsening of function, and specific drug-related adverse events. Subgroups of interest were based on age, sex, race/ethnicity, duration of symptoms, severity of symptoms, and ME/CFS case definition used.

For treatment of ME/CFS, inclusion was restricted to studies that utilized a formal, published case definition for diagnosis of ME/CFS. We included only studies of patients that met criteria for ME/CFS, and not those that only had conditions often present in patients with ME/CFS (e.g., fibromyalgia, irritable bowel syndrome, orthostatic intolerance). No duration or timing restriction was applied, other than that treatment trials had to assess outcomes at least 12 weeks after initiation of therapy, because short-term outcomes may not be maintained and may be less meaningful than longer-term outcomes, given the chronic and fluctuating nature of ME/CFS.[29] Studies conducted in inpatient settings or in institutionalized individuals were excluded, to increase applicability to outpatient management, where ME/CFS is typically treated.

Data Extraction and Data Management

The following information was extracted from included studies into evidence tables: study design, setting, ME/CFS case definition, inclusion and exclusion criteria, population characteristics (including sex, age, race, duration of ME/CFS, baseline fatigue, baseline function, presence and severity of depression, and other co-morbidities), sample size, duration of follow-up, attrition, characteristics of treatments and control interventions, funding source, and results, including outcomes at baseline and at follow-up. For studies that reported population characteristics by treatment arm, mean values and standard deviations were calculated for the overall sample from the data provided. For each study, data extraction was performed by two investigators: the first investigator extracted the data, and the second investigator independently reviewed the extracted data for accuracy and completeness.

Risk of Bias of Individual Studies Assessment

The risk of bias of each study was assessed based on predefined criteria adapted from methods proposed by the U.S Preventive Services Task Force. The criteria used are consistent

with the approach recommended by AHRQ in the AHRQ Methods Guide.[27] Two investigators independently assessed the risk of bias of each study. Discrepancies were resolved through discussion and consensus, with a third investigator making the final decision if necessary.

For randomized trials of interventions, risk of bias assessment was based on the methods used for randomization, allocation concealment, and blinding; the similarity of compared groups at baseline; attrition; and use of intent-to-treat analysis.[27,30] For observational studies, risk of bias assessment was based on the methods for selecting patients, ascertaining exposures and outcomes, attrition, and analysis, including control for confounders (when applicable). Based on these factors, each study was assigned an overall "low," "medium," or "high" risk of bias.[27,30]

Low risk of bias studies are considered likely to be valid. Low risk of bias studies clearly describe the population, setting, interventions, and comparison groups; use a valid method for allocating patients to interventions; clearly report dropouts and have low dropout rates; use appropriate methods to control for confounders (observational studies); blind patients and care providers to treatments (randomized trials); assess outcomes blinded to intervention status; and appropriately measure outcomes and fully report results.

Medium risk of bias studies have some methodological deficiencies, but no flaw or combination of flaws judged likely to cause major bias. The study may be missing information, making it difficult to assess its methods or assess limitations and potential problems. The medium risk of bias category is broad, and studies with this rating vary in their strengths and weaknesses: the results of some medium risk of bias studies are likely to be valid, while others are probably invalid.

High risk of bias studies have significant flaws that may invalidate the results. They have a serious or "fatal" flaw in design, analysis, or reporting; large amounts of missing information; or discrepancies in reporting. The results of these studies are at least as likely to reflect flaws in the study design as true effects of the interventions under investigation. High risk of bias studies were not excluded a priori but were considered highly unreliable.

Assessing Research Applicability

Applicability is the extent to which the effects observed in published studies are likely to reflect the expected results when a specific intervention is applied to the population of interest under "real-world" conditions.[27] It indicates the extent to which research included in a review might be useful for informing clinical decisions in specific situations. We recorded factors relevant for understanding applicability, such as the characteristics of the patients (e.g., severity or duration of ME/CFS, ME/CFS case definition used, presence and severity of associated conditions and symptoms, and demographic characteristics), interventions, and settings.[31] To interpret the magnitude of benefits, we defined a minimum clinically important difference for fatigue as 2.3 points on the 11-item 0 to 33 Chalder scale, 0.6 points on the 1 to 7 Fatigue Severity Scale, or 11.5 points on the 1 to 50 Multidimensional Assessment of Fatigue;[32] for function as 10 points on the 0 to 100 Short Form (SF)-36 physical function subscale;[33] and for psychiatric outcomes as 1.7 points on the 0 to 21 Hospital Anxiety and Depression Scale (HADS) depression or anxiety scales.[34] For pooled standardized mean difference (SMD) estimates for outcomes reported using different scales, we defined an SMD of 0.2 to <0.5 as small, 0.5 to <0.8 as moderate, and ≥0.8 as large.[35]

Data Synthesis

Meta-analysis was performed for exercise and CBT, the treatments evaluated in the largest number of trials, using the Dersimonian-Laird random-effects models in RevMan 5.3 (the Nordic Cochrane Centre, Copenhagen).[36] Separate meta-analyses were performed for exercise versus inactive controls, exercise versus active treatments, CBT versus inactive controls, and CBT versus active treatments. We pooled results separately for each active treatment comparator, due to clinical heterogeneity and potential differences in effects; stratified results are presented for each active treatment comparator without pooling results across comparators.

Meta-analyses were also conducted on continuous measures for fatigue, functional impairment, depression, anxiety, sleep quality, pain, and the 6-minute walk test. Separate analyses were performed for outcomes assessed at the end of treatment and for outcomes assessed after the completion of therapy (post-intervention follow-up). Analyses of continuous outcomes were based on the raw (unadjusted) mean difference or SMD (for outcomes assessed using different scales) in follow-up scores.[37] To enable calculation of pooled raw mean differences, continuous pain scales were converted to a common 0 to 10 scale. For function, results using the SF-36 physical function subscale were converted to the standard 0 to 100 scale if necessary. Estimates based on the difference in change from baseline were similar or slightly larger than the difference in follow-up scores and are not discussed further. We utilized the difference in change from baseline when follow-up scores were not reported. Studies that reported adjusted estimates reported results similar to the raw mean differences. When fatigue or function were reported using different scales, we reported the results as the SMD; we also reported stratified results based on each of the original scales. When standard deviations for follow-up scores were not reported, we imputed them based on the average from the other studies in the analysis. Unless indicated otherwise, for all continuous outcomes except for functional impairment, lower scores indicate a better outcome; for functional impairment, lower scores indicate a worse outcome, based on the method for scoring the Short-Form Physical Component Summary score or Physical Function subscale (the most commonly reported measures of function). If necessary, for the purpose of meta-analysis we reversed other scales used to measure function so that the direction of effects (e.g., higher scores indicating worse outcomes) was the same for all studies in an analysis. Although some trials reported results at long-term, post-trial follow-up, we restricted meta-analyses to outcomes assessed during the trial, due to potential crossover and contamination following trial completion.

We also conducted meta-analyses on dichotomous measures for improvement in fatigue, improvement in function, recovery, serious adverse events, withdrawal due to adverse events, withdrawal due to worsening, post-exertional malaise, and school attendance (for studies of children), based on the pooled relative risk (RR). If necessary, RR's were calculated from data reported in the trial publication. For the Pacing, graded Activity, Cognitive behavior therapy; a randomized Evaluation (PACE) trial, the primary analyses of dichotomous outcomes were based on data reported in the main trial publication,[38] which utilized definitions modified from the original protocol.[39] We conducted sensitivity analyses using data based on the original protocol definitions.[40]

For studies with more than two treatment arms relevant for an analysis, we combined the arms for the main analysis, so that each study was represented once, in order to avoid overweighting. However, one study could be represented in multiple subgroups in stratified analyses. Heterogeneity was assessed using the I-squared statistic (the proportion of variation in study estimates due to heterogeneity).[41,42] We conducted subgroup analyses based on the inactive

control type (usual care, specialist care, attention control, wait list, or placebo), ME/CFS case definition, and CBT type (individual/face-to-face, individual/web or telephone, group/face-to-face) and evaluated for the subgroup differences with a statistical test (fixed effect analysis based on the inverse-variance method in RevMan 5.3). We also performed sensitivity analyses in which high risk of bias and outlier trials (trials that qualitatively differed substantially from others in the analysis) were excluded. We did not evaluate for potential publication bias using graphical or statistical methods for small sample effects, because no analysis had at least 10 trials.[43]

Grading the Body of Evidence for Each Key Question

We assessed the strength of evidence for treatment comparisons and outcomes addressed in Key Question 3, in accordance with the AHRQ Methods Guide.[27,44] The strength of evidence was based on risk of bias/study limitations (graded low, moderate, or high); the consistency of results between studies (graded consistent, inconsistent, or consistency unknown when only one study was available); the directness of the evidence linking the intervention and health outcomes (graded direct or indirect); the precision of the estimate of effect, based on the number and size of studies and confidence intervals (CI) for the estimates (graded precise or imprecise); and whether reporting bias was suspected (graded suspected or undetected). We did not evaluate the strength of evidence for Key Question 1 because it provided descriptive information regarding the prevalence of non-CFS conditions and we did not evaluate the strength of evidence for Key Question 2 because no studies met inclusion criteria for this question.

The strength of evidence was rated using the four categories recommended in the AHRQ Methods Guide:[27,44] A "high" grade indicates high confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies and the findings are stable (i.e., another study is unlikely to change the conclusions). A "moderate" grade indicates moderate confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies and findings are likely to be stable, but there is some uncertainty. A "low" grade indicates low confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both) and additional evidence is needed to determine that the findings are stable or that the estimate of effect is close to the true effect. An "insufficient" grade indicates that evidence is too limited to estimate an effect, there is no confidence in the effect estimate, no evidence is available, or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Peer Review and Public Commentary

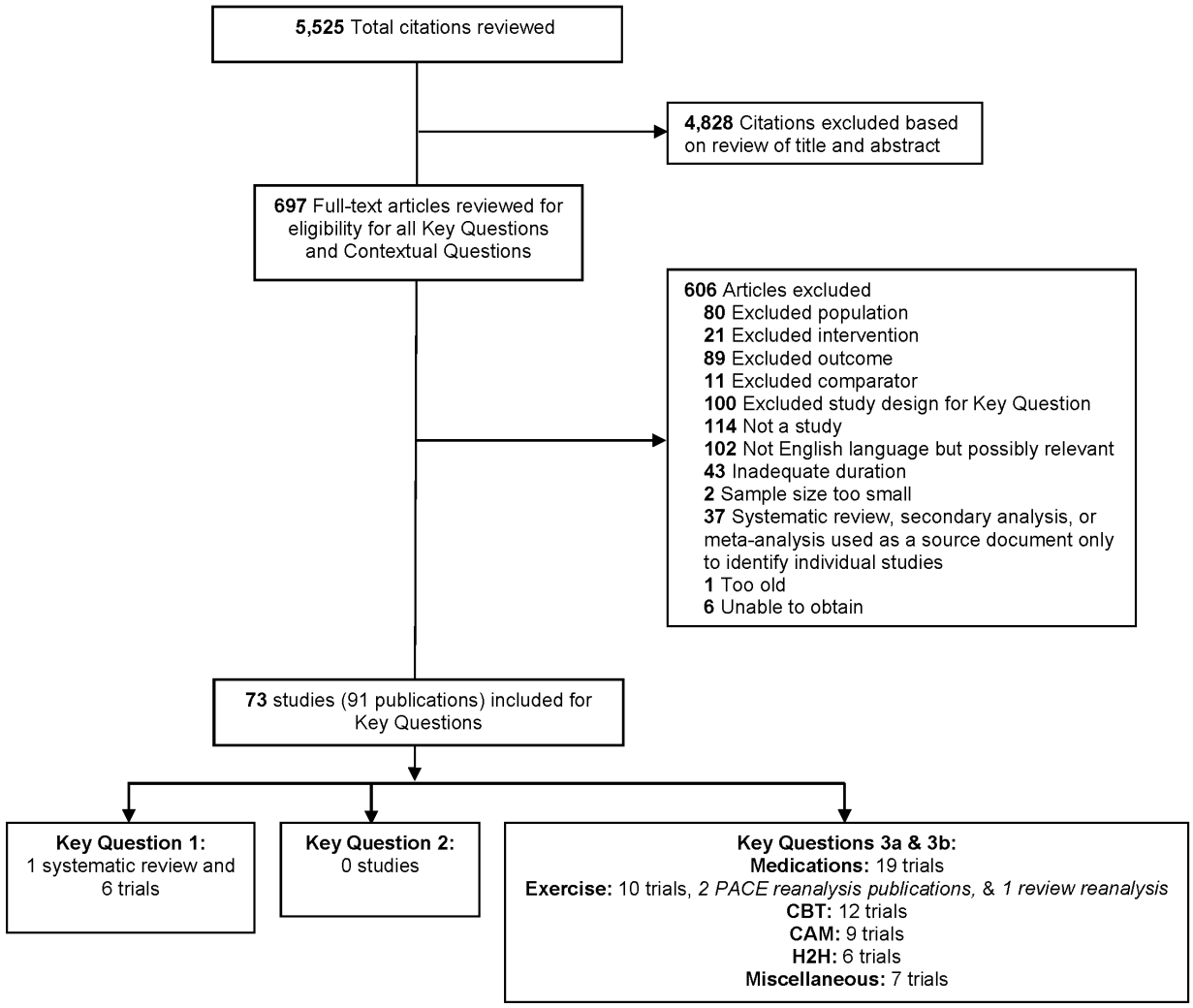
Experts in ME/CFS and Key Informants were invited to provide external peer review of the draft report. The draft report was also posted on the Federal Register Notice and regulations.gov for 90 days (May 17, 2021 to August 16, 2021) to facilitate a public comment period through Federal Register Notice and Regulations.gov. The draft report was edited in response to peer review and public comment, including additional clarification of limitations in the evidence, prior to finalization.

Results

Results of Literature Searches

Results of the literature search and selection process are summarized in the literature flow diagram (**Figure 1**). Database searches and searches of reference lists resulted in 5,525 potentially relevant citations. After dual review of abstracts and titles, 697 articles were selected for full-text review. After dual review of full text articles, 73 studies (in 91 publications) were included. Thirty-three studies were included in the prior AHRQ report and 40 studies were added for this update. The new studies include one systematic review and six observational studies for Key Question 1 and 33 studies for Key Question 3 (treatments for ME/CFS). No study met inclusion criteria for Key Question 2. Detailed evidence tables with data abstraction and risk of bias assessment tables for included studies by Key Question are available in **Appendices E** and **F**.

Figure 1. Literature flow diagram



Note: Some studies included multiple interventions or were reported in multiple publications.
Abbreviations: CAM= complementary and alternative medicine; CBT = cognitive behavioral therapy; H2H= head-to-head comparisons of active interventions; RCT = randomized controlled trial

Key Question 1. In patients undergoing evaluation for possible ME/CFS, what is the frequency of non-ME/CFS conditions?

Key Points

- A systematic review of studies of patients presenting with tiredness/fatigue found a pooled prevalence of anemia of 2.8% (95% CI 1.6% to 4.8%; 3 studies, N=1091), malignancy 0.6% (0.3% to 1.3%; 3 studies, N=1091), depression 18.5% (16.2% to 21.0%; 6 studies, N=1000), and serious somatic diseases 4.3% (2.7% to 6.7%; 3 studies, N=436).
- A study of primary care adult patients with a primary symptom of fatigue (n=571) found that the most frequent diagnostic categories (present in >5% of patients) were musculoskeletal (19.4%), infection (18.0%), psychological or social (16.5%, most commonly depression, strain/burnout, or anxiety), gastrointestinal (8.1%), and neurologic (6.7%, most commonly headache or dizziness).
- Four European studies of adult patients with fatigue (total N = 2293) undergoing evaluation for possible ME/CFS in specialty settings found that when ME/CFS was excluded, the most common non-ME/CFS conditions were psychiatric (15% to 51%) and sleep disorders (6% to 30%), although one study found neurological (21%), neurodegenerative (15%), and immunologic (13%) conditions to be most common. A U.S. study (N=104) found that the most common non-ME/CFS conditions were alcohol abuse (8.2%), anemia (6.1%), diabetes mellitus (16%), high C-reactive protein (20%), hypothyroidism (20%), depression (8.2%), urinary tract infection (8.2%), restless legs syndrome (6.1%), and substance abuse (6.1%).

Detailed Synthesis

We included a systematic review[45] and six additional studies[46-51] on the prevalence of non-ME/CFS conditions in adult patients presenting with fatigue and possible ME/CFS. The systematic review addressed the differential diagnosis of tiredness/fatigue and included 26 studies[45] **(Evidence Table Appendix E1)**. The review restricted inclusion to studies conducted in primary care settings in which patients sought care for tiredness or fatigue (i.e., tiredness symptoms were not elicited from patients through a review of symptoms or other method). Although all patients had fatigue, they did not necessarily present specifically for evaluation of ME/CFS and methods used to exclude or diagnose ME/CFS were not well-described. The review pooled prevalence data for common causes of tiredness based on all studies that met inclusion criteria, as well as a more rigorous subset of studies that used precisely defined diagnostic criteria and described an appropriate diagnostic work-up. Based on this subset of studies in the review, the pooled prevalence of anemia was 2.8% (1.6% to 4.8%; 3 studies, N=1091), malignancy 0.6% (0.3% to 1.3%; 3 studies, N=1091), depression 18.5% (16.2% to 21.0%; 6 studies, N=1000), and serious somatic diseases 4.3% (2.7% to 6.7%; 3 studies, N=436). The serious somatic disease category overlapped with the other categories and included diabetes, anemia, hypothyroidism, and malignancy. The systematic review also included three studies on the prevalence of CFS in those with fatigue. The rates were 1.9% and 0.7% in two studies and 31.2% in the third. The higher prevalence in the latter study could be related to the inclusion criteria: it restricted inclusion to patients with tiredness for at least 6 months without a diagnosis associated with the tiredness (2 of the criteria in the 1994 Fukuda case definition for CFS that

was utilized in the study). The review did not pool estimates of prevalence of CFS, due to the small number of studies, in addition to the inconsistency across studies.

A prospective Dutch study of primary care patients (N=571) evaluated diagnoses following a new episode in which fatigue was the main symptom (duration >6 months in 58%) (**Table 2**).[50] It found that 0.7% of patients were diagnosed with CFS. The most frequent diagnostic categories were musculoskeletal (19.4%), psychological or social (16.5%, most commonly depression, strain/burnout, or anxiety), gastrointestinal (8.1%), neurologic (6.7%, most commonly headache or dizziness), general conditions (4.2%, including anemia, adverse drug effects, pain, and allergies), infection (18%), respiratory (4.9%), endocrine (2.8%, most commonly hypothyroidism), cardiovascular (1.9%), climacteric symptoms (1.1%), malignancy (0.7%), and dermatological (0.5%).

Five studies (N=2293) reported the prevalence of CFS in patients evaluated in specialty settings for possible CFS[46-49] (**Evidence Table Appendix E1**). As in the systematic review, a limitation of these studies is that methods use dot exclude or diagnose ME/CFS were not well-described. One study[46] was conducted in the U.S. and four studies[47-49] in Europe. In these studies, the proportion of patients who met criteria for CFS ranged from 7% to 60% (**Table 2**). In the four European studies, the most common non-ME/CFS conditions among patients who did not meet criteria for CFS were psychiatric (15% to 51%) and sleep disorders (6% to 30%). Other diagnoses included cardiovascular, pain, endocrine, nutritional, musculoskeletal, gastrointestinal, and neurological conditions (<5% for each of these categories). One study[49] found that 47% of patients had chronic diseases but did not specify the conditions further and one study[51] found neurological (21.4%), neurodegenerative (15%), and immunologic (13.5%) disorders to account for most non-ME/CFS conditions. The U.S. study (N=104) enrolled patients who were self-referred or referred by a clinician for evaluation of possible ME/CFS.[46] A diagnosis of ME/CFS was excluded in 47% due to the presence of non-ME/CFS conditions. The most common non-ME/CFS conditions were alcohol abuse (8.2%), anemia (6.1%), diabetes mellitus (16%), high C-reactive protein (20%), hypothyroidism (20%), depression (8.2%), restless legs syndrome (6.1%), substance abuse (6.1%), and urinary tract infection (8.2%). Less common conditions (diagnosed in <5% of the sample) were active inflammation, anorexia, autoimmune disorder, bipolar disorder, spinal disease, hepatitis C virus infection, high blood urea, hypertension, mitochondrial myopathy, obesity, obstructive sleep apnea, osteoarthritis, narcolepsy, rheumatoid arthritis, sleep problems, schizophrenia, sickle cell disease, and uncontrolled high blood pressure.

We identified no studies on the prevalence of non-ME/CFS conditions in children presenting with fatigue symptoms.

Table 2. Studies reporting diagnosis rates for ME/CFS and non-ME/CFS conditions

Author, Year	Presentation	Setting Country	N	ME/CFS (%)	Non-ME/CFS Conditions (%)
Brimmer, 2013[46]	Self-referred from CFS support group or referred by clinician for evaluation of fatigue	Healthcare settings: primary care and multiple specialties United States	104	36%	Non-CFS due to insufficient fatigue or number of symptoms: 17.3% Non-CFS due to presence of other conditions: 47.1%* Active inflammation: 4.1% Alcohol abuse: 8.2% Anemia: 6.1% Anorexia: 2.0% Autoimmune disorder: 2.0% Bipolar: 4.1% Spinal disease: 2.0% Diabetes mellitus: 16.3% Hepatitis C virus: 2.0% High blood urea: 4.1% High C-reactive protein: 20.4% Hypertension: 2.0% Hypothyroidism: 20.4% Depression: 8.2% Mitochondrial myopathy: 2.0% Obesity: 4.1% Obstructive sleep apnea: 4.1% Osteoarthritis: 4.1% Narcolepsy: 2.0% Restless legs syndrome: 6.1% Rheumatoid arthritis: 2.0% Sleep problems: 2.0% Schizophrenia: 2.0% Sickle cell: 2.0% Substance abuse: 6.1% Uncontrolled high blood pressure: 2.0% Urinary tract infection: 8.2%
Devasahayam, 2012[47]	Referral to CFS service	ME/CFS specialty United Kingdom	250 (assessed)	54%	Psychiatric: 22% Sleep disorder: 6% Pain: 2% Endocrine: 3% Nutritional: 3% Musculoskeletal: 1% Gastrointestinal: 2% Neurological: 1% Others: 2% Miscellaneous/other: 2.4%
Mariman, 2013[48]	Presumed CFS	Multidisciplinary setting Belgium	279	Unequivocal CFS: 23% CFS with comorbidity: 21% Psychiatric disorder: 2.5% Sleep disorder: 16% Both: 2.5%	≥4 minor Fukuda criteria, CFS excluded: 35.8% Psychiatric disorder: 12.5% Sleep disorder: 6.5% Both: 14.7% Internal disease: 1.4% Other: 0.7% <4 minor Fukuda criteria: 19.7% Psychiatric: 6.5% Sleep: 3.2% Both: 6.1% Other: 4.0%

Author, Year	Presentation	Setting Country	N	ME/CFS (%)	Non-ME/CFS Conditions (%)
Newton, 2010[49]	Referral to CFS service	ME/CFS Specialty United Kingdom	260	60%	Chronic disease: 47% Sleep disorder: 20% Psychological: 15% Idiopathic: 13% Cardiovascular: 4% Other: 1%
Nijrolder, 2009[50]	Fatigue	Primary care The Netherlands	571	0.7% (4/571)	Musculoskeletal: 19.4% Psychological or social: 16.5% Gastrointestinal: 8.1% Neurologic: 6.7% General (anemia, adverse drug event, pain, allergies): 4.2% Infection: 18% Respiratory: 4.9% Endocrine: 2.8% Cardiovascular: 1.9% Menopause: 1.1% Cancer: 0.7% Skin: 0.5%`
Slomko, 2019[51]	Self-identified as meeting Fukuda criteria	ME/CFS Specialty Poland	1400	7%	Other chronic conditions: 1308/1400 (93%) Neurological: 280/1308 (21.4%) Neurodegenerative: 200/1308 (15%) Psychiatric: 654/1308 (50%) Immunologic: 174/1308 (13.5%)

*Individual could have more than one exclusion condition and the exclusion could be based on one or multiple conditions
Abbreviations: CFS = chronic fatigue syndrome; ME = myalgic encephalomyelitis

Key Question 2. What are the benefits and harms of diagnosing ME/CFS vs. non-diagnosis?

Key Points

No study measured benefits or harms of diagnosing ME/CFS versus non-diagnosis.

Detailed Synthesis

We identified no study that measured benefits or harms of diagnosing ME/CFS versus non-diagnosis. The prior AHRQ report included fourteen studies on the consequences of the diagnostic process or diagnosis of ME/CFS.[1] The studies primarily used descriptive or qualitative methods, and did not meet inclusion criteria for this review because no study measured patient outcomes (e.g., quality of life, function, mood) using validated measures or compared outcomes in persons diagnosed with ME/CFS compared with those not diagnosed with ME/CFS. The AHRQ report included five studies that found that patients with ME/CFS feel stigmatized by their diagnosis in multiple aspects of their life. Two studies in the AHRQ report described prejudices and stereotypes in medical trainees and mental health practitioners related to the diagnosis assigned to an identical case presentation (CFS, ME, or other). The AHRQ report also included six studies that indicated a substantial burden due to failure to diagnosis ME/CFS, due to misdiagnosis or not meeting case definitions for ME/CFS due to presence of an exclusionary condition. Although the prior AHRQ report focused on harms of ME/CFS diagnosis, it included one study in which patients reported that a CFS diagnosis reduced uncertainty and provided social and medical legitimacy by providing a coherent diagnosis for their symptoms.

Key Question 3. What are the benefits and harms of therapeutic interventions for patients with ME/CFS and how do they vary by patient subgroups?

Key Question 3a. Interventions for treating ME/CFS
Key Question 3b. Interventions for treating symptoms commonly present in persons with ME/CFS (poor sleep, orthostatic intolerance, pain, fatigue, cognitive problems, depression, multiple chemical sensitivity, gastrointestinal symptoms, urinary symptoms, etc.)

Key Points

Exercise Therapy

- In adults diagnosed with ME/CFS, graded exercise therapy (GET) was associated with decreased fatigue severity and improved function versus inactive controls at the end of therapy and at post-intervention follow-up, but the magnitude of benefits was small to moderate. The trials had methodological limitations, most trials used the Oxford case definition, and there was unexplained statistical heterogeneity (low strength of evidence).
- Graded exercise was associated with increased likelihood of improvement in fatigue, improvement in function, and recovery versus inactive controls, based on the modified or original PACE trial definitions for these outcomes; however, the definition for recovery did not exclude patients with persistent symptoms (low strength of evidence).
- Graded exercise was associated with decreased depression severity, decreased anxiety severity, and improved sleep quality versus inactive controls, but the magnitude of benefit was small (low strength of evidence).
- Harms of graded exercise versus inactive controls were not well reported, but two trials found no difference in risk of serious adverse events, withdrawal due to worsening of symptoms, or physical function worsening, though estimates were imprecise. One trial found graded exercise associated with decreased risk of post-exertional malaise versus usual specialist care (low strength of evidence).
- There were no differences between GET versus CBT in fatigue, function, depression, anxiety, sleep quality, pain, or likelihood of recovery, but findings were based on 1 or 2 trials and most estimates were imprecise (low strength of evidence).
- Comparisons of exercise therapy versus other (non-CBT) active therapies (relaxation, adaptive pacing, biofeedback, or fluoxetine) were limited to 1 or 2 trials each with no differences for most outcomes; however, estimates were frequently imprecise. One trial found graded exercise associated with improved outcomes versus adaptive pacing (low strength of evidence).
- One small trial found no difference between home orthostatic training versus sham training in fatigue severity. Home orthostatic training was associated with a small improvement in blood pressure changes with standing, but orthostatic symptoms were not reported.

Cognitive Behavioral Therapy (CBT)

- In adults diagnosed with ME/CFS, CBT was associated with decreased fatigue severity and improved function versus inactive controls at the end of therapy and at post-intervention follow-up, but the magnitude of benefits was small to moderate, the trials had methodological limitations, some trials used the Oxford case definition, and there was unexplained statistical heterogeneity (low strength of evidence).
- CBT was associated with greater likelihood of improvement in fatigue and recovery, based on the modified or original PACE trial definitions for these outcomes; however, the definition for recovery did not exclude patients with persistent symptoms. There was no difference in likelihood of improvement in function (low strength of evidence).
- CBT was associated with decreased depression severity, decreased anxiety severity, and improved sleep quality versus inactive controls, but the magnitude of benefit was small (low strength of evidence).
- Harms of CBT versus inactive controls were not well reported, but two trials found no difference in risk of serious adverse events, withdrawal due to worsening, or physical function worsening, though estimates were imprecise. One trial (PACE) found CBT associated with decreased risk of post-exertional malaise versus usual specialist care (low strength of evidence).
- Comparisons of CBT versus other active therapies (relaxation, adaptive pacing, cognitive therapy, or mirtazapine) were limited to 1 or 2 trials each, with no differences for most outcomes but imprecise estimates. One trial (PACE) found CBT associated with improved outcomes versus adaptive pacing (low strength of evidence).
- In adolescents diagnosed with ME/CFS, CBT (family focused or with parental involvement) was associated with decreased fatigue severity at the end of the intervention; effects on severity of functional impairment and school attendance favored CBT but differences were not statistically significant.
- One trial found an intensive but brief osteopathy, life coaching, and neurolinguistics programming intervention ("Lightning Process") in adolescents diagnosed with ME/CFS associated with improved function versus usual specialist care, but there were no statistically significant effects on fatigue, pain, anxiety, depression, or quality of life. School attendance was improved at 12 months but not at 6 months (low strength of evidence).

Other Behavioral Approaches

- There was insufficient evidence to determine effects of other behavioral approaches in adults (illness management and peer counseling, mindfulness-based cognitive therapy, or self-management interventions) due to small numbers of trials, imprecise estimates, methodological limitations, and inconsistency in findings (for self-management interventions).

Medications

- The immune modulating biologic drug rintatolimod was associated with small improvements in exercise ability and overall function, with greater frequency of infusion-related headache, flu-like symptoms, chills, vasodilation and dyspnea versus placebo (low strength of evidence). Fatigue was not measured.

- Immunoglobulin G (IgG) infusions were not associated with improvements in fatigue or function versus placebo in 2 trials of adults (low strength of evidence). A small trial of adolescents found no difference between IgG versus placebo in overall improvement in function, but significantly more patients had greater than 25% improvement 3 months post treatment using an unvalidated method (insufficient strength of evidence). IgG infusions were associated with increased likelihood of withdrawal due to adverse events and headache versus placebo.
- Small placebo-controlled trials of other drugs, including other immune modulators, antidepressants, corticosteroids, and single studies of an antiviral, an acetyl-cholinesterase inhibitor, an alpha-adrenergic agonist, and a stimulant, did not find statistically significant effects on fatigue or function outcomes (low strength of evidence).

Dietary Interventions, Herbal Supplements, or Homeopathy

- There was insufficient evidence to determine the effects of dietary interventions/herbal supplements (insulin-like growth factor, antioxidant, acetyl-carnitine, homeopathy, melatonin, low-sugar/low-yeast diet).

Complementary and Alternative Therapies

- There was insufficient evidence to determine the effects of yoga, abdominal tuina, or distant healing.
- Although single small studies found qigong exercise associated with decreased fatigue severity, the evidence was insufficient due to small sample sizes and methodological limitations of the studies.

Detailed Synthesis

Exercise Therapy

Ten trials evaluated exercise therapy in adult patients with ME/CFS (**Tables 3 and 4, Evidence Table Appendix E2**).[38,52-60] Sample sizes ranged from 24 to 630 (total N=1688). Six trials compared exercise versus inactive controls (usual care, usual specialist care, or an attention control [advice or supportive listening]) and six trials compared exercise versus an active intervention (CBT, adaptive pacing, relaxation, biofeedback, or fluoxetine). Six trials were included in the prior AHRQ report[38,53-55,58,59] and four trials were added for this update.[52,56,57,60] Of the new trials, two trials compared exercise versus inactive controls[52,56] and two trials compared exercise versus active therapies (relaxation[57] and biofeedback[60]).

One trial was conducted in the United States, seven trials in Europe, and 1 trial each in New Zealand and Australia. The mean age of participants ranged from 28 to 51 years and the proportion female ranged from 69% to 100%. The case definition for ME/CFS was the Oxford criteria in five trials, the Fukuda criteria in four trials, and the National Institute for Clinical Excellence (NICE) criteria in one trial. The duration of ME/CFS ranged from 28 to 52 months in four trials that reported this information. Baseline fatigue was measured using a variety of scales **(Table 3)**. One trial[52] required patients to have post-exertional fatigue or malaise (proportion with post-exertional malaise not reported); in the other trials, the proportion of patients with post-exertional fatigue or malaise was not described and none of the trials described the severity of post-exertional fatigue or malaise. Two trials excluded patients with major depression.[53,60] In

the other trials, the proportion of patients with depression or an axis I psychiatric diagnosis ranged from 10% to 39%. Depression severity was most commonly reported using the HADS depression score (0 to 21 scale, higher scores indicate more severe depression). In eight trials, mean HADS depression scores ranged from 6.2 to 9.6. Functional impairment was most commonly reported using the SF-36 physical function subscale (0 to 100 scale, lower scores indicate more functional impairment). In eight trials, baseline mean SF-36 physical function subscale scores ranged from 30.0 to 49.4.

The exercise intervention was graded exercise in all of the trials except for one,[54] which evaluated anaerobic exercise. The duration of the exercise therapy intervention ranged from 8 to 26 weeks. In most trials, the frequency of exercise was weekly or biweekly. The session length and exercise intensity varied, with details not reported in some trials (Table 3). Outcomes were assessed at 12 to 70 weeks; eight trials evaluated patients at the end of the intervention and seven trials evaluated patients 13.5 to 52 weeks following the completion of therapy.

Eight trials were rated medium risk of bias and two trials[57,60] were rated high risk of bias (**Risk of Bias Table Appendix F**). In all trials, blinding of patients and care providers to the exercise intervention was not feasible. Other methodological limitations included high attrition, failure to report attrition, inadequate description of randomization or allocation concealment methods, and failure to blind or unclear blinding status of outcomes assessors and data analysts.

Table 3. Exercise therapy RCTs: study characteristics

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Clark, 2017[52] GETSET United Kingdom Medium	n: 199 Age: 38.4 % Female: 79	Criteria: NICE Duration: Median 46 and 42 months	Chalder Fatigue Scale 11-item (0 to 33): Baseline: 26.2 (SD 4.7) Post-exertional fatigue: Post-exertional fatigue or malaise required (proportion with post-exertion malaise not reported)	Major depression: 10% Baseline depression: HADS depression (0 to 21): 8.9 (SD 4.0) Baseline function: SF-36 physical function (0 to 100): 48.7 (SD 22.4)	A: Graded exercise (guided graded exercise self-help) plus specialist medical care B: Specialist medical care Frequency: Once, then up to 3 more sessions Session length: 30 minutes (initial), 20 minutes (follow-up) Exercise intensity: Not specified Duration of treatment: ~8 weeks Duration of follow-up: 12 weeks

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Fulcher, 1997[53] United Kingdom Medium	n: 59 Age: 37.2 % Female: 74	Criteria: Oxford Duration, median: 2.7 years	Chalder Fatigue Scale 14-item (0 to 42) Baseline, mean: 29.7 (SD 6.4) Post-exertional fatigue or malaise: not reported	Major depression: Excluded Baseline depression: HADS depression (0 to 21), median: 5.0 (range 1.5 to 8.5) Baseline function: SF-36 physical function (0 to 100), mean: 47.8 (SD 20.5)	A: Graded exercise B: Flexibility/relaxation Frequency: Weekly visits, plus home exercise 5 days weekly Session length: Visit length not reported, at home practice increased to maximum of 30 minutes Exercise intensity: maximum 30 minutes daily at 60% of peak oxygen consumption Duration of treatment: 12 weeks Duration of follow-up: 12 weeks for 1 year. Flexibility group was permitted to cross over to exercise treatment after 12-week follow-up
Jason, 2007[54] United States Medium	n: 114 Age: 43.8 % Female: 83	Criteria: Fukuda Duration: not reported	Fatigue Severity Scale 9-item (1 to 7) Baseline: 6.1 (SD 0.71) Post-exertional fatigue or malaise: not reported	Major depression: Current axis I diagnosis: 39% Baseline depression: Beck Depression Inventory (0 to 63): 18.7 (SD 9.9) Baseline function: SF-36 physical function (0 to 100): 46.2 (SD 23.8)	A: Anaerobic Exercise (Anaerobic Activity Therapy [ACT]/progressive relaxation) B: Relaxation (RELAX) C: Cognitive-behavioral therapy D. Cognitive therapy Frequency: Biweekly Session length: 45 minutes Exercise intensity: Not specified Duration of treatment: 6 months Duration of follow-up: 1 year
Moss-Morris, 2005[55] New Zealand Medium	n: 43 Age: 41.0 % Female: 69	Criteria: Fukuda Duration: Median 2.67 and 5.00 (unclear if months or years)	Chalder Fatigue Scale 14-item (0 to 42) Baseline: 24.9 (SD 8.4) Post-exertional fatigue or malaise: not reported	Major depression: not reported Baseline depression: HADS depression (0 to 21): 6.2 (SD 19.8) Baseline function: SF-36 physical function (0 to 100): 49.4 (SD 19.8)	A: Graded exercise B: Usual care Frequency: Weekly Session length: 1-hour initial session, length of follow-up sessions not reported Exercise intensity: Goal 30 minutes 5 days a week at 80% of maximum heart rate Duration of treatment: 12 weeks Duration of follow-up: 12 weeks

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Powell, 2001[56] Powell, 2004[61] United Kingdom Medium	n: 148 Age: 33.2 % Female: 78	Criteria: Oxford Duration: median 51.7 months	Chalder Fatigue Scale 11-item (0 to 11) Baseline: 10.3 (SD 1.4) Post-exertional fatigue or malaise: not reported	On antidepressants: 18% Baseline depression: HADS depression (0 to 21): 9.4 (SD 3.9) Baseline function: SF-36 physical function (10 to 30): 16.0 (SD 3.5)	A: Graded exercise (maximum intervention) B: Graded exercise (minimum intervention) C: Graded exercise (minimum intervention + telephone contacts) D. Standardized medical care (assessment, advice, and booklet) Frequency: 9 sessions over 3 months (maximum intervention), 2 initial sessions only (minimum intervention), 2 initial sessions plus 7 telephone contacts over 3 months (minimum intervention + telephone contacts) Session length: 2 initial sessions totaled 3 hours in all groups; 1-hour follow-up (maximum intervention), or 20-minute follow-up (telephone follow-up) Exercise intensity: not specified Duration of treatment: 3 months Duration of follow-up: 12 months/24 months
Wallman, 2004[57] Australia High	n: 61 Age: not reported (range 16 to 74 years) % Female: not reported	Criteria: Fukuda Duration: not reported	Chalder Fatigue Scale 11-item (0 to 33) Baseline: Mental fatigue: 6.0 (SD 1.8) Physical fatigue: 11.5 (SD 3.5) Post-exertional fatigue or malaise: not reported	Major depression: 12% Baseline depression: HADS depression (0 to 21): 6.8 (SD 3.2) Baseline function: not reported	A: Graded exercise B: Flexibility/relaxation Frequency: Biweekly Session length: not described Exercise intensity: not described Duration of treatment: 12 weeks Duration of follow-up: 12 weeks

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Wearden, 1998[59] United Kingdom Medium	n: 136 Age: 38.7 % Female: 71	Criteria: Oxford Duration of fatigue: median 28.0 months	Chalder Fatigue Scale 14-item (0 to 42) Baseline: 34.5 (SD 5.7) Post-exertional fatigue or malaise: not reported	Major depression: 10% Baseline depression: HADS depression (0 to 21): 8.8 (SD 3.5) Baseline function: not reported	A: Graded exercise B: Attention control (advice) C. Fluoxetine D. Graded exercise + fluoxetine Frequency: at weeks 0, 1, 2, 4, 8, 12, 20, and 26 Session length: 20 minutes Exercise intensity: At least 3 times weekly, at 75% of maximum oxygen uptake. Increased after reduction of 10 bpm in post-exercise heart rate for 1 week and two points on the perceived exertion scale. Duration of treatment: 26 weeks Duration of follow-up: 26 weeks
Wearden, 2010[58] FINE United Kingdom Medium	n: 274 Age: 44.6 % Female: 78	Criteria: Oxford Duration: 7 years	Chalder Fatigue Scale 11-item (0 to 11) Baseline: 10.5 (SD 1.1) Post-exertional fatigue or malaise: not reported	Any depression diagnosis: 18% Baseline depression: HADS depression (0 to 21): 9.6 (SD 4.1) Baseline function: SF-36 physical function (0 to 100): 30.1 (SD 18.9)	A: Graded exercise (pragmatic rehabilitation, including relaxation exercises) B: Usual care C. Supportive listening Frequency: 10 sessions over 18 weeks Session length: 90 minutes (initial session), 1 hour (home visits weeks 2, 4, 10, and 19), 30 minutes (telephone calls weeks 3, 6, 8, 12, and 15) Exercise intensity: not specified Duration of treatment: 18 weeks Duration of follow-up: 20 weeks/70 weeks

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
White, 2011[38] PACE United Kingdom Medium	n: 630 Age: 28 % Female: 77	Criteria: Oxford Duration: median 32 months	Chalder Fatigue Scale 11-item (0 to 33) Baseline: 28.2 (SD 3.8) Post-exertional fatigue or malaise: not reported	Any depressive disorder: 34% Baseline depression: HADS depression (0 to 21) 8.2 (SD 3.8) Baseline function: SF-36 physical function (0 to 100): 38.0 (SD 15.8)	A: Graded exercise + specialist medical care B: Specialist medical care C: CBT + specialist medical care D. Adaptive pacing therapy + specialist medical care Frequency: Weekly for 4 weeks, then biweekly, plus one booster at 36 weeks Session length: not described Exercise intensity: Target 30 minutes 5 times weekly Duration of treatment: 23 weeks (booster at 36 weeks) Duration of follow-up: 52 weeks
Windthorst, 2017[60] Germany High	n: 24 Age: 50.7 % Female: 100	Criteria: Fukuda Duration of symptoms: >2 years in 92%	MFI 20-item (20 to 100) Baseline: 64.8 (SD 9.9) Post-exertional fatigue or malaise: not reported	Major depression: Excluded Baseline depression: PHQ-9 (0 to 27): mean: 8.2 (SD 4.3) Baseline function: SF-36 physical function (0 to 100): 40.4 (SD 8.8)	A: Graded exercise B: Heart rate variability biofeedback therapy Frequency: weekly Session length: 50 minutes Exercise intensity: Target 20 to 30 minutes 2 to 3 times weekly at 70% of maximum heart rate Duration of treatment: 8 weeks Duration of follow-up: 5 months

Abbreviations: ACT = anaerobic activity therapy; CFS = chronic fatigue syndrome; FINE = Fatigue Intervention by Nurses Evaluation; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; HADS-D = Hospital Anxiety and Depression Scale-depression; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; PHQ = patient health questionnaire; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey

Table 4. Exercise therapy RCTs: study results

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes* (fatigue and post-exertional fatigue)	Depression Outcomes*	Function Outcomes*
Clark, 2017[52] GETSET NICE	A: Graded exercise (guided graded exercise self-help) plus specialist medical care (107) B: Specialist medical care (104) Duration of treatment: ~8 weeks Duration of follow-up: 12 weeks	Fatigue at end of intervention, mean (SD): Chalder Fatigue Scale 11-item (0 to 33): 19.1 (7.6) vs. 22.9 (6.9), AMD: -4.2 (95% CI, -6.1 to -2.3) p<0.0001 Meeting Fukuda criteria (n=138), mean difference in Chalder fatigue scale score at end of intervention: -4.1 (95% CI, -6.5 to -1.7) p=0.001 Meeting Oxford criteria (n=141), mean difference in Chalder fatigue scale score: -3.5 (95% CI, -5.7 to -1.3) p=0.002	HADS depression (0 to 21), mean (SD): 7.4 (4.3) vs. 8.6 (4.7), mean difference: -1.1 (-2.0 to -0.3), p=0.006	Physical Function at end of intervention, mean (SD) SF-36 physical function (0 to 100): Overall: 55.7 (23.3) vs. 50.8 (25.3), AMD: 6.3 (95% CI, 1.8 to 10.8) p=0.006 Meeting Fukuda criteria (n=141), mean difference in SF-36: 6.3 (95% CI, 1.1 to 11.6) p=0.019 Meeting Oxford criteria (n=159), mean difference in SF-36: 5.6 (95% CI, 0.8 to 10.4) p=0.024
Fulcher, 1997[53] Oxford	A: Graded exercise (33) B: Flexibility/relaxation (33) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks/1 year	Fatigue at end of intervention, mean (SD): Chalder Fatigue Scale 14-item (0 to 42): 20.5 (8.9) vs. 27.4 (7.4); p=0.004 VAS total fatigue (sum of subscales, (200=normal): 253 (48) vs. 286 (67); p=0.04 VAS physical fatigue subscale (, 100=normal): 130 (28) vs. 154 (34); p=0.006 VAS mental fatigue subscale (100=normal): 124 (31) vs. 132 (39); p=0.38	HADS depression at end of intervention (0 to 21), median (IQR): 5.5 (2.9 to 8.1) vs. 4 (0.6 to 7.4), p=0.92	Physical Function at end of intervention, mean (SD) SF-36 physical function (0 to 100): 69 (18.5) vs 55 (21.8); p=0.01

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes* (fatigue and post-exertional fatigue)	Depression Outcomes*	Function Outcomes*
Jason, 2007[54] Fukuda	A: Anaerobic exercise (Anaerobic Activity Therapy [ACT]/progressive relaxation (29) B: Relaxation (RELAX) (28) C: Cognitive-behavioral therapy (29) D: Cognitive therapy (28) Duration of treatment: 6 months Duration of follow-up: 1 year	Fatigue at follow-up, mean (SD): Fatigue Severity Scale scores 9-item (1 to 7): 5.77 (1.43) vs. 5.62 (1.06) vs. 5.37 (1.19) vs. 5.87 (1.01); p=NR	Beck Depression Inventory (0 to 63) at follow-up: 16.94 (11.82) vs. 13.50 (9.97) vs. 13.95 (13.08) vs. 11.86 (7.36), p<0.001	Physical Function at follow-up, mean (SD) SF-36 physical function (0 to 100): 39.72 (27.63) vs. 61.20 (27.70) vs. 58.64 (30.44) vs. 61.09 (23.74) p<0.01 for cognitive-behavioral therapy and cognitive therapy over time vs. ACT over time % Achieving clinically significant improvement: 11.1 vs. 21.7 vs. 18.2 vs. 30.4; p=0.49
Moss-Morris, 2005[55] Fukuda	A: Graded exercise (25) B: Usual care (24) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	Fatigue at end of intervention, mean (SD): Chalder Fatigue Scale 14-item (0 to 42): 13.91 (10.88) vs. 24.41 (9.69); p=0.02 Chalder Fatigue Scale physical fatigue: 7.91 (7.06) vs. 14.27 (5.75); p=0.02 Chalder Fatigue Scale mental fatigue: 6.00 (4.06) vs. 10.14 (4.27); p=0.03	Not reported	Physical Function at end of intervention, mean (SD) SF-36 physical function (0 to 100): 69.05 (21.94) vs. 55.00 (22.94); p=0.49

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes* (fatigue and post-exertional fatigue)	Depression Outcomes*	Function Outcomes*
Powell, 2001[56] Powell, 2004[61] Oxford	A: Graded exercise (maximum intervention) (38) B: Graded exercise (minimum intervention) (37) C: Graded exercise (minimum intervention + telephone contacts) (39) D: Standardized medical care (assessment, advice, and booklet) (34) Duration of treatment: 3 months Duration of follow-up: 12 months/24 months	Fatigue, mean (95% CI): Chalder Fatigue Scale 11-item (0 to 11): 3 months: 4.3 (2.9 to 5.8) vs. 5.0 (3.4 to 6.6) vs. 3.7 (2.3 to 5.2) vs. 10.4 (10.1 to 10.8) 6 months: 3.4 (2.2 to 4.6) vs. 3.8 (2.5 to 5.2) vs. 4.0 (2.5 to 5.5) vs. 9.9 (9.1 to 10.8) 1 year: 3.1 (1.8 to 4.4) vs. 3.2 (1.8 to 4.7) vs. 3.5 (2.1 to 4.9) vs. 10.1 (9.3 to 10.8), p<0.001 (initial scores and depression scores used as covariates) 2 year, Mean score, (SD): 2.84 (3.67) vs. 4.46 (4.78) vs. 3.59 (4.69) vs. 6.07 (4.60)	Depression, mean (95% CI) HADS depression (0 to 21) 3 months: 5.8 (4.8 to 6.9) vs. 6.1 (4.7 to 7.4) vs. 5.9 (4.5 to 7.3) vs. 11.2 (9.6 to 12.9) 6 months: 5.0 (3.8 to 6.2) vs. 5.4 (3.9 to 6.9) vs. 5.6 (4.3 to 6.9) vs. 11.0 (9.2 to 12.9) 12 months: 4.2 (2.9 to 5.5) vs. 4.2 (3.0 to 5.5) vs. 4.6 (3.2 to 6.0) vs. 10.1 (8.4 to 11.7), p<0.001 (initial scores used as a covariate) 2-year Mean score, (SD): 4.08 (4.33) vs. 5.11 (5.12) vs. 4.77 (4.67) vs. 8.37 (5.75)	Physical Function, mean (95% CI) SF-36 physical function (10 to 30): 3 months: 22.8 (21.2 to 24.3) vs. 22.8 (21.1 to 24.4) vs. 22.3 (20.6 to 24.0) vs. 16.3 (14.9 to 17.7) 6 months: 24.1 (22.6 to 25.6) vs. 24.0 (22.4 to 25.6) vs. 23.0 (21.2 to 24.7) vs. 17.2 (15.6 to 18.7) 1 year: 24.9 (23.4 to 26.4) vs. 25.1 (23.3 to 26.8) vs. 24.3 (22.5 to 26.0) vs. 16.9 (15.4 to 18.4), p<0.001 (initial scores and depression scores used as covariates) 2-year Mean score, (SD): 25.45 (4.72) vs. 24.11 (5.94) vs. 23.64 (6.39) vs. 22.47 (7.02)
Wallman, 2004[57] Fukuda	A: Graded exercise (32) B: Flexibility/relaxation (29) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	Mental fatigue at end of intervention, maximum score 12, average score (range): 4.5 (3.9 to 5.2) vs. 4.8 (4.2 to 5.5) Chalder Fatigue Scale physical fatigue, average score (range): 8.1 (6.9 to 9.4) vs. 9.6 (8.3 to 10.9)	HADS depression (0 to 21) at end of intervention, mean (95% CI): 4.8 (3.6 to 5.9) vs. 6.5 (5.5 to 7.6), p=0.041	Overall Function at end of intervention: Ratings of perceived exertion (estimated from figure): 1.3 vs. 1.8 (p=0.013)

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes* (fatigue and post-exertional fatigue)	Depression Outcomes*	Function Outcomes*
Wearden, 1998[59] Oxford	A: Graded exercise + fluoxetine (33) B: Graded exercise (34) C: Fluoxetine (35) D: Attention control (advice) (34) Duration of treatment: 26 weeks Duration of follow-up: 26 weeks	Fatigue, mean (95% CI): Chalder Fatigue Scale 14-item (0 to 42) 0-12 weeks: -5.7 (-9.2 to -2.2) vs. -2.0 (-4.1 to 0.1) vs. -1.6 (-4.4 to 1.2) vs. -2.1 (-4.9 to 0.6) 26 weeks: -6.0 (-9.7 to -2.3) vs. -2.7 (-5.4 to 0.01) vs. -3 (-5.9 to -0.2) vs. -5.7 (-9.5 to -1.9) % non-cases of fatigue (Chalder fatigue scale score <4 positive items on 14-item scale) 12 weeks: 18 (6/33) vs. 6 (2/34) vs. 3 (1/35) vs. 3 (1/34) 26 weeks: 18 (6/33) vs. 6 (2/34) vs. 6 (2/ 35) vs. 18 (6/34) p=0.025 for exercise interventions combined vs. others Exercise improved Chalder Fatigue Scale scores Mean change 0 to12 weeks: 2.1 (95% CI -0.6 to 4.8), p=0.13 Mean change 26 weeks: 2.9 (95% CI -0.2 to 6.1), p=0.07	HADS-Depression, mean change (95% CI) at 26 weeks: -2.0 (3.3 to -0.7) vs. -1.2 (-2.5 to 0.2) vs. -1.7 (-3.0 to -0.5) vs. -1.3 (-2.3 to -0.3)	Overall Function, mean (95% CI) functional work capacity (amount of O2 consumed in the final minute of exercise per kg of body weight) 0-12 weeks: 2.2 (1.0 to 3.4) vs. 2.6 (1.0 to 43) vs. 0.4 (-1.2 to 2.0) vs. 0.4 (-0.9 to 1.7). 26 weeks: 2.0 (0.4 to 3.5) vs. 2.8 (0.8 to 4.8) vs. 1.0 (-0.9 to 3.0) vs. -0.1 (-1.7 to 1.6) Effect of exercise on functional work capacity Mean change 0-12 weeks: 2.0 (95% CI 0.60 to 3.49), p=0.005 Mean change 0-26 weeks: 1.9 (95% CI 0.15 to 3.69), p=0.03
Wearden, 2010[58] FINE Oxford	A: Graded exercise (pragmatic rehabilitation, including relaxation exercises) (95) B: Usual care (100) C: Supportive listening (101) Duration of treatment: 18 weeks Duration of follow-up: 20 weeks/70 weeks	Fatigue, mean (SD): Chalder Fatigue Scale 11-item (0 to 11) 20 weeks: 8.39 (3.67) vs. 9.32 (3.18) vs. 9.67 (2.76); treatment effect estimate -1.18, 95% CI -2.18 to -0.18; p=0.021 for graded exercise vs. usual care 70 weeks: 8.72 (3.65) vs. 9.48 (2.71) vs. 9.39 (3.21). Graded exercise vs. usual care Chalder Fatigue Scale 11-item (0 to 33) (20 weeks: 22.78 (8.56) vs. 26.27 (7.68) 70 weeks: 23.90 (8.34) vs. 26.02 (7.11)	HADS-Depression, mean (SD): 20 weeks: 7.28 (4.02) vs. 8.48 (4.47) vs. 8.85 (4.01) 70 weeks: 7.88 (4.45) vs. 8.06 (4.75) vs. 8.67 (4.51)	SF-36 physical function (0 to 100), mean (SD) 20 weeks: 39.94 (25.21) vs. 40.27 (26.45) vs. 33.28 (22.94) Treatment effect: -7.54, 95% CI -12.96 to -2.33; p=0.005 for supportive listening vs. usual care 70 weeks: 43.27 (27.38) vs. 39.83 (27.77) vs. 35.72 (25.94); p=NS

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes* (fatigue and post-exertional fatigue)	Depression Outcomes*	Function Outcomes*
White, 2011[38] Dougall, 2014[62] PACE Oxford	A: Adaptive pacing therapy + specialist medical care (160) B: CBT + specialist medical care (161) C: Graded exercise + specialist medical care (160) D: Specialist medical care (160) Duration of treatment: 23 weeks (booster at 36 weeks) Duration of follow-up: 12 months	Mean (SD) Chalder Fatigue Scale 11-item (0 to 33) 12 weeks: 24.2 (6.4) vs. 23.6 (6.5) vs. 22.8 (7.5) vs. 24.3 (6.5) 24 weeks: 23.7 (6.9) vs. 21.5 (7.8) vs. 21.7 (7.1) vs. 24.0 (6.9) 52 weeks: 23.1 (7.3) vs. 20.3 (8.0) vs. 20.6 (7.5) vs. 23.8 (6.6) 52 weeks: CBT vs. control: p=0.0001 APT vs. control: p=NS GET vs. control: p=0.0003 CBT vs. APT: p=0.0027 GET vs. APT: p=0.0059 % Improved from baseline by ≥2 points: 65 (99/153) vs. 76 (113/148) vs. 80 (123/154) vs. 65 (98/152) % Within normal range: 22 (34/153) vs. 41 (60/148) vs. 33 (51/154) vs. 21 (32/152)	HADS-Depression (0 to 21), mean (SD) 52 weeks: 7.2 (4.5) vs. 6.2 (3.7) vs. 6.1 (4.1) vs. 7.2 (4.7); CBT vs. control: p=0.0003; GET vs. control: p=0.0035; CBT vs. APT: p=0.382, GET vs. APT: p=0.23	Mean (SD) SF-36 physical function (0 to 100) 12 weeks: 41.7 (19.9) vs. 51.0 (20.7) vs. 48.1 (21.6) vs. 46.6 (20.4) 24 weeks: 43.2 (21.4) vs. 54.2 (21.6) vs. 55.4 (23.3) vs. 48.4 (23.1) 52 weeks: 45.9 (24.9) vs. 58.2 (24.1) vs. 57.7 (26.5) vs. 50.8 (24.7) 52 weeks: APT vs. control: p=NS CBT vs. control: p=0.0068 GET vs. control: p=0.0005 CBT vs. APT: p=0.0002 GET vs. APT: p<0.0001 % Improved from baseline by ≥8 points: 49 (75/153) vs. 71 (105/148) vs. 70 (108/154) vs. 58 (88/152) % Within normal range: 35 (53/153) vs. 52 (77/148) vs. 53 (81/154) vs. 41 (62/152)
Windthorst, 2017[60] Fukuda	A: Graded exercise (11) B: Heart rate variability biofeedback therapy (13) Duration of treatment: 8 weeks Duration of follow-up: 5 months	MFI 20-item (20 to 100) total baseline vs. end of treatment vs. 5-month follow-up, mean(SD): GET: 68.8 (10.1) vs. 56.6 (18.8) vs. 55.6 (21.3), p=0.319 Biofeedback: 61.5 (9.7) vs. 48.2 (15.9) vs. 43.6 (15.9), p<0.001	PHQ-9 (0 to 27) baseline vs. end of treatment vs. 5-month follow-up, mean (SD): GET: 8.9 (5.4) vs. 8.3 (4.6) vs. 8.8 (6.0), p=0.656 Biofeedback: 7.5 (3.1) vs. 4.3 (3.0) vs. 4.2 (3.1), p=0.006	Overall Function: SF-36 Physical baseline vs. end of treatment vs. 5-month follow-up, mean (SD): GET: 37.7 (7.8) vs. 44.8 (9.7) vs. 46.6 (7.1), p=0.011 Biofeedback: 42.6 (9.2) vs. 45.2 (9.9) vs. 47.1 (12.2), p=0.292

*A vs. B vs. C vs. D, unless otherwise noted

Abbreviations: ACT = anaerobic activity therapy; AMD = adjusted mean difference; APT = adaptive pacing therapy; CBT = cognitive behavioral therapy; CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; GET = graded exercise therapy; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence; NS = not significant; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; RCT = randomized controlled trial; SD = standard deviation; SES = standardized effect sizes; SF-36 = 36-item Short Form Health Survey

Exercise Versus Inactive Controls

Six trials (N=1,430)[38,52,55,56,58,59] compared graded exercise versus usual care (3 trials),[55,56,58] usual specialist care (2 trials),[38,52] or an attention control (advice or supporting listening, 2 trials)[58,59] **(Tables 3 and 4)**. One trial[52] that used the NICE case definition was published subsequent to the AHRQ report and one older trial[56] not included in the prior AHRQ report used the Oxford case definition. The duration of the exercise intervention ranged from 8 to 26 weeks across the trials. All of the trials evaluated patients at the end of the intervention; four trials also evaluated patients 4 weeks to 24 months following the end of the intervention. All of the trials were rated medium risk of bias. Results stratified by the inactive comparator are summarized in **Table 5** and shown in **Figures 2 to 11**.

Fatigue

Graded exercise was associated with decreased fatigue severity versus usual care, specialist care, or an attention control at the end of the exercise intervention, though statistical heterogeneity was high (6 trials, N=1,034, SMD -0.62, 95% CI -0.95 to -0.30, I^2=81%; **Figure 2**).[38,52,55,56,58,59] Fatigue severity was measured using the Chalder fatigue scale, but the trials used different versions and scoring methods. Mean differences were -2.91 (95% CI -4.36 to -1.47, I^2=91%) in two trials (N=501) that used the 11-item 0 to 33 Chalder scale,[38,52] -6.47 (95% CI -13.80 to 0.86, I^2=72%) in two trials (N=111) that used the 14-item 0 to 42 Chalder scale,[55,59] and -3.60 (95% CI -8.49 to 1.29, I^2=98%) in two trials (N=422) that used the 11-item 0 to 11 Chalder scale[56,58] (**Table 5**). Estimates consistently favored exercise when trials were stratified by the control type or ME/CFS case definition and there were no statistically significant subgroup differences; however, stratification on these factors did not reduce heterogeneity **(Table 5)**. The most commonly used case definition was the Oxford criteria (4 trials, N=792, SMD -0.59, 95% CI -1.05 to -0.14, I^2=88%).[38,56,58,59]

An outlier trial by Powell et al. reported substantially greater effects on fatigue (SMD -1.46, 95% CI -1.88 to -1.04) than the other trials (SMD range -0.07 to -1.00).[56] The methods used to select patients (the Oxford case definition), severity of ME/CFS symptoms at baseline (mean 10.3 on the 11-item 0 to 11 Chalder scale and 30.0 on the 0 to 100 SF-36 physical function subscale) and intensity of the exercise intervention (2 to 9 sessions) did not appear to explain the difference in results between this trial and the others. Excluding this trial reduced statistical heterogeneity and attenuated the pooled estimate (5 trials, N=886, SMD -0.32, 95% CI -0.52 to -0.12, I^2=46%).

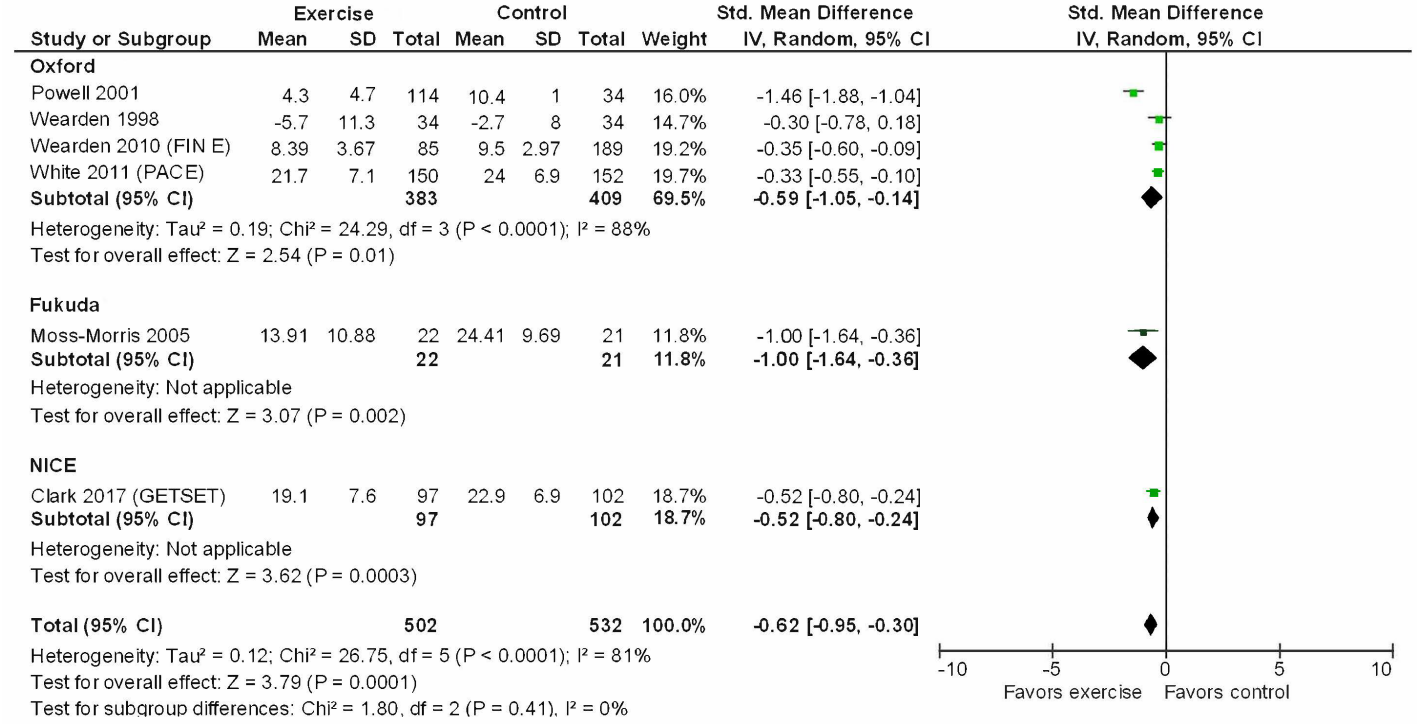
Graded exercise was also associated with decreased fatigue severity versus controls at post-intervention follow-up (29 to 52 weeks following the completion of therapy), though the estimate was based on fewer studies (3 trials, N=625, SMD -0.76, 95% CI -1.48 to -0.05, I^2=94%;[38,56,58] **Figure 3**). Statistical heterogeneity was very high. The outlier trial by Powell et al.[56] reported substantially greater effects on fatigue (SMD -1.69) than the other two trials (SMD -0.21 and -0.45). Excluding this trial reduced statistical heterogeneity and attenuated the pooled estimate (2 trials, N=477, SMD -0.35, 95% CI -0.58 to -0.12, I^2=35%).

The PACE trial evaluated improvement in fatigue as a dichotomous outcome. The main PACE publication found graded exercise associated with an increased likelihood of experiencing a ≥2 point improvement on the 11-item 0 to 33 Chalder scale versus specialist care (80% vs. 65%, RR 1.23, 95% CI 1.07 to 1.42; adjusted risk difference [ARD] 15%, 95% CI 5% to 25%).[38] However, this differed from the original protocol, which defined improvement in fatigue as a score of ≤3 on the 11-item 0 to 11 Chalder fatigue scale or improvement of >50 percent from

baseline.[39] Using the original protocol definition, graded exercise remained associated with increased likelihood of improvement in fatigue (24% vs. 13%, RR 1.81, 95% CI 1.11, to 2.94; ARD 15%, 95% CI 2% to 19%).[40]

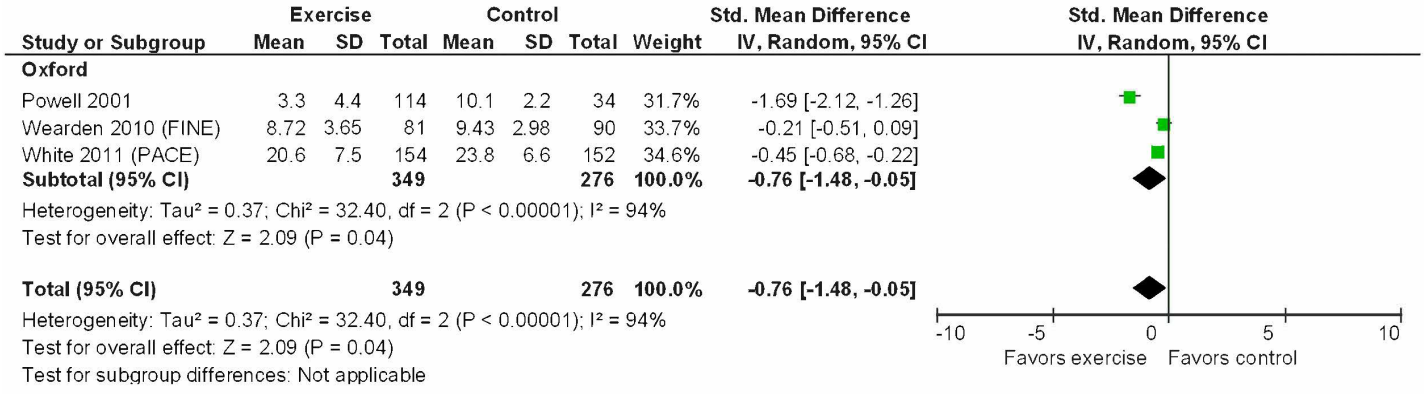
The PACE trial also evaluated longer-term, post-trial outcomes of graded exercise versus specialist care (N=320, median duration from randomization 31 months).[63] 32 percent of patients in the exercise group and 63 percent in the specialist care group received non-randomly allocated therapies between the end of the trial (1 year) and long-term follow-up. Fatigue severity on the 11-item 0 to 33 Chalder fatigue scale was improved at long-term follow-up compared with end-of-trial scores in the exercise (mean change -1.3 points, 95% CI -2.7 to -0.1) and specialist care (-3.9 points, 95% CI -5.3 to -2.6) groups. At long-term post-trial follow-up, there was no difference between graded exercise vs. specialist medical care in fatigue severity (mean difference -0.8, 95% CI -2.8 to 1.2), based on mixed model analyses.

Figure 2. Fatigue severity: graded exercise versus inactive control at end of intervention



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; NICE = National Institute for Health and Care Excellence; PACE = pacing, graded activity, cognitive behavior therapy; ; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 3. Fatigue severity: graded exercise versus inactive control at post-intervention follow-up



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Function

Graded exercise was associated with less severe functional impairment versus usual care, usual specialist care, or an attention control (5 trials, N=965, mean difference 11.73, 95% CI 2.33 to 21.14 on the 0 to 100 SF-36 physical function subscale, I^2=88%; **Figure 4**), but statistical heterogeneity was high.[38,52,55,56,58] Estimates consistently favored exercise when trials were stratified according to the control type or ME/CFS and there were no statistically significant subgroup differences **(Table 5)**. However, stratification on these factors did not reduce heterogeneity. The most commonly used ME/CFS case definition was the Oxford criteria (3 trials, N=723, mean difference 13.60, 95% CI -1.18 to 28.37, I^2=93%).[38,56,58]

An outlier trial by Powell et al.[56] reported substantially greater effects on SF-36 physical function scores (mean difference 31.50, 95% CI 23.03 to 39.97) than the other trials (mean differences ranged from 3.24 to 14.05 points). Excluding this outlier trial eliminated statistical heterogeneity and attenuated the pooled estimate (4 trials, N=817, mean difference 5.89, 95% CI 2.52 to 9.25, I^2=0%).

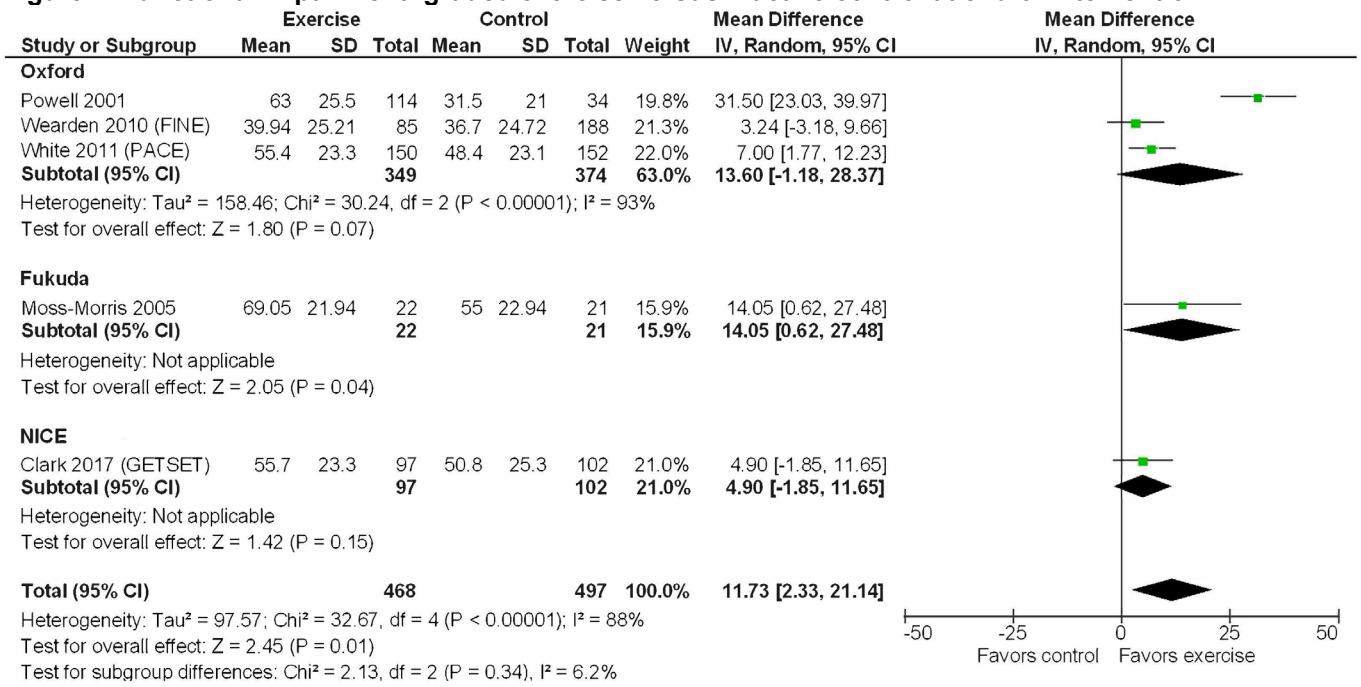
Graded exercise was associated with decreased functional impairment versus controls at post-intervention follow-up, but the difference was not statistically significant (3 trials, N=711, mean difference 17.07, 95% CI -2.02 to 36.16 on the 0 to 100 SF-36 physical function subscale, I^2=95%;[38,56,58] **Figure 5**). Excluding the outlier trial by Powell et al.[56] attenuated the estimate, eliminated statistical heterogeneity, and resulted in a statistically significant effect (2 trials, N=563, mean difference 6.37, 95% CI 1.89 to 10.85, I^2=0%).

Three trials evaluated functional improvement as a dichotomous outcome. Functional improvement was defined in one trial[56] as ≥50 point improvement from baseline on SF-36 physical function (standardized to a 0 to 100 scale) or score ≥75 and in one trial[58] as a ≥50% improvement from baseline or score >70. In PACE, the third trial, the main study publication defined improvement in fatigue as ≥8 point improvement on the SF-36 physical function score from baseline (proportion meeting this definition 70% vs. 58%).[38] However, this differed from the study protocol, which defined functional improvement as an SF-36 physical function score ≥75 or ≥50% improvement from baseline (proportion meeting this definition 61% vs. 44%), similar to the definition used in the other trials.[39] Using the data from the main PACE publication, exercise was associated with increased likelihood of functional improvement versus usual care or specialist care (3 trials, N=618, RR 2.48, 95% CI 0.77 to 7.97, I^2=89%; ARD 28%, 95% CI -7% to 63%) (**Figure 6**).[38,56,58] The pooled estimate was very similar when data based on

the original (protocol) PACE definition for functional improvement were used (3 trials, N=632, RR 2.52, 95% CI 0.90 to 7.02, I^2=85%; ARD 30%, 95% CI -4% to 63%).[40] The trial by Powell et al.[56] reported a much stronger effect on likelihood of functional improvement (RR 11.78, 95% CI 3.05 to 45.45) than the other two trials (RR 1.39 and 1.74). Excluding this outlier trial from the analysis attenuated the pooled estimate and eliminated statistical heterogeneity, but results were based on only two trials (N=484, RR 1.41, 95% CI 1.15 to 1.74, I^2=0%; ARD 13%, 95% CI 5% to 21%).[38,58]

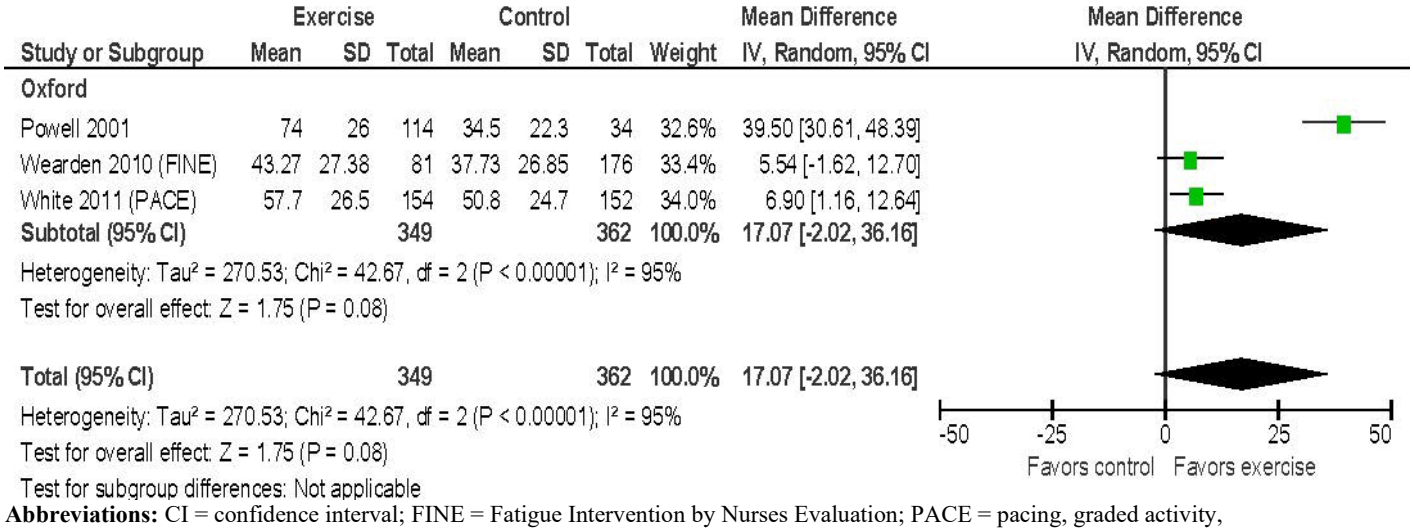
The PACE trial also evaluated longer-term (median duration from randomization 31 months), post-trial outcomes.[63] There was no change in SF-36 physical function at long-term follow-up compared with the end of the trial for exercise (mean change 0.5 point, 95% CI -2.7 to 3.6 on a 0 to 100 scale), but function improved in the specialist care group (mean change 7.1 points, 95% CI 4.0 to 10.3). At long-term post-trial follow-up, there was no difference between graded exercise versus specialist medical care (mean difference 2.0, 95% CI -4.0 to 7.9), based on mixed model analyses.

Figure 4. Functional impairment: graded exercise versus inactive control at end of intervention



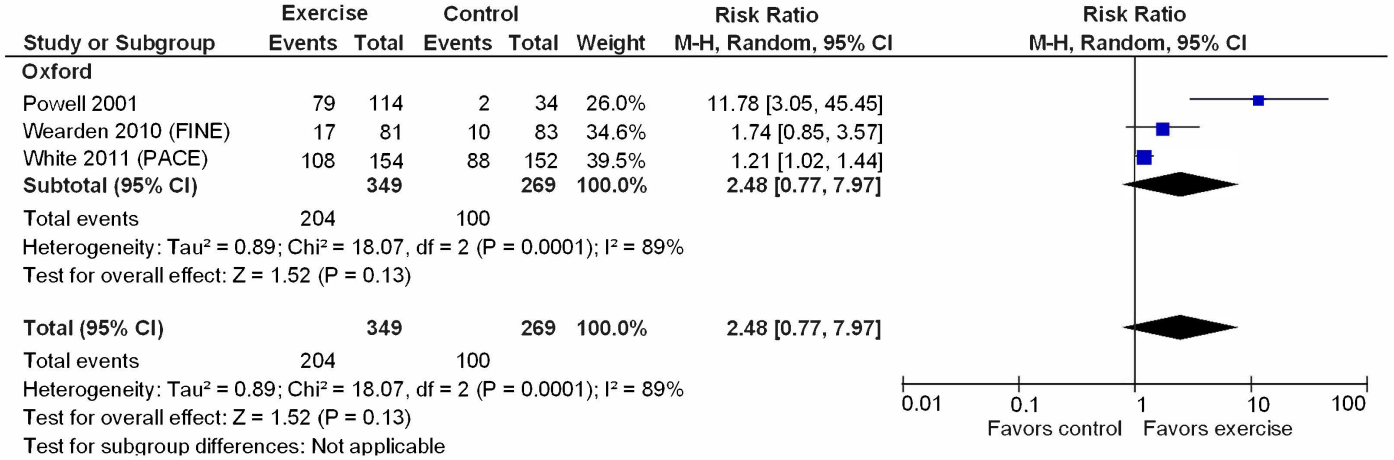
Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; ; NICE = National Institute for Health and Care Excellence; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 5. Functional impairment: graded exercise versus inactive control at post-intervention follow-up



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 6. Likelihood of functional improvement: graded exercise versus inactive control



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Depression and anxiety

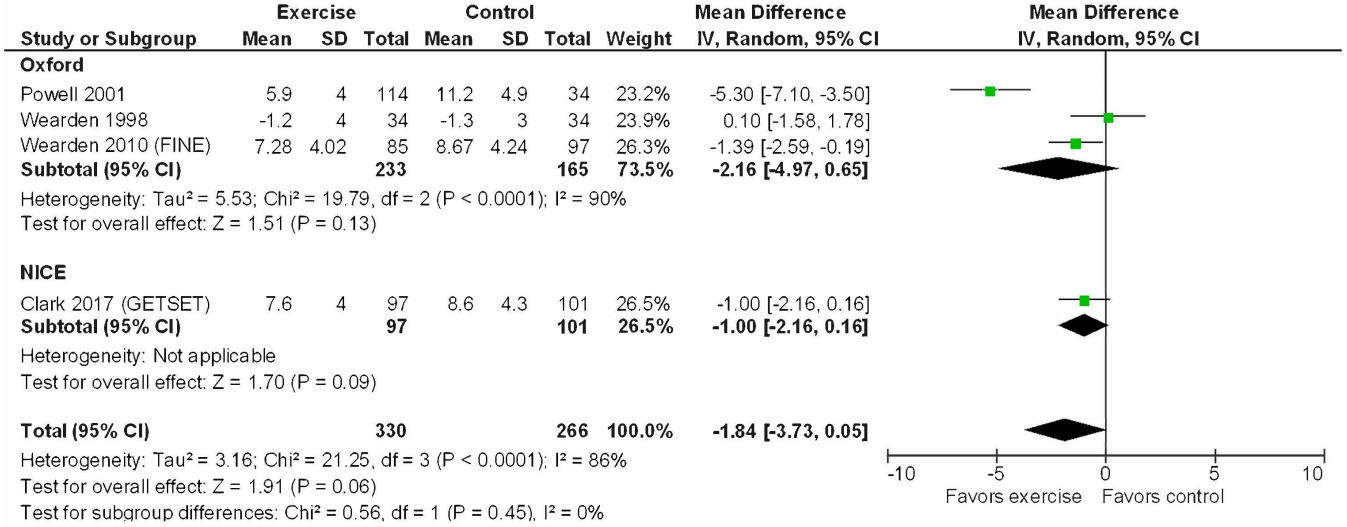
Graded exercise was associated with less severe depression versus usual care, usual specialist care, or an attention control at the end of the intervention (4 trials, N=688, mean difference -1.84, 95% CI -3.73 to 0.05 on the 0 to 21 HADS depression scale, I^2=86%;[52,56,58,59] **Figure 7**). The estimate was similar at post-intervention follow-up, but the difference was not statistically significant (3 trials, N=699, mean difference -2.36, 95% CI -4.98 to 0.27, I^2=92%).[38,56,58] Statistical heterogeneity was high. Excluding the trial by Powell et al. attenuated the estimates and accounted for almost all statistical heterogeneity (3 trials, N=540, mean difference -0.97, 95% CI -1.71 to -0.23, I^2=8%) at end of intervention and (2 trials, N=551, mean difference -0.85, 95% CI -1.61 to -0.08, I^2=0%) at post-intervention follow-up.

Graded exercise was also associated with less severe anxiety versus usual care, usual specialist care, or an attention control at the end of the intervention (3 trials, N=620, mean difference -1.59, 95% CI -2.41 to -0.77 on the 0 to 21 HADS anxiety scale, I^2=16%;[52,56,58]

Figure 8). There was no difference between graded exercise versus controls in anxiety at post-intervention follow-up (3 trials, N=697, mean difference -1.07, 95% CI -2.64 to 0.49, I^2=75%).[38,56,58] Excluding the trial by Powell et al. slightly attenuated the estimate at the end of the intervention (2 trials, N=372, mean difference -1.31, 95% CI -2.12 to -0.51, I^2=0%) and had little effect on the estimate for anxiety at post-intervention follow-up (2 trials, N=566, mean difference -0.38, 95% CI -1.52 to 0.76, I^2=49%).

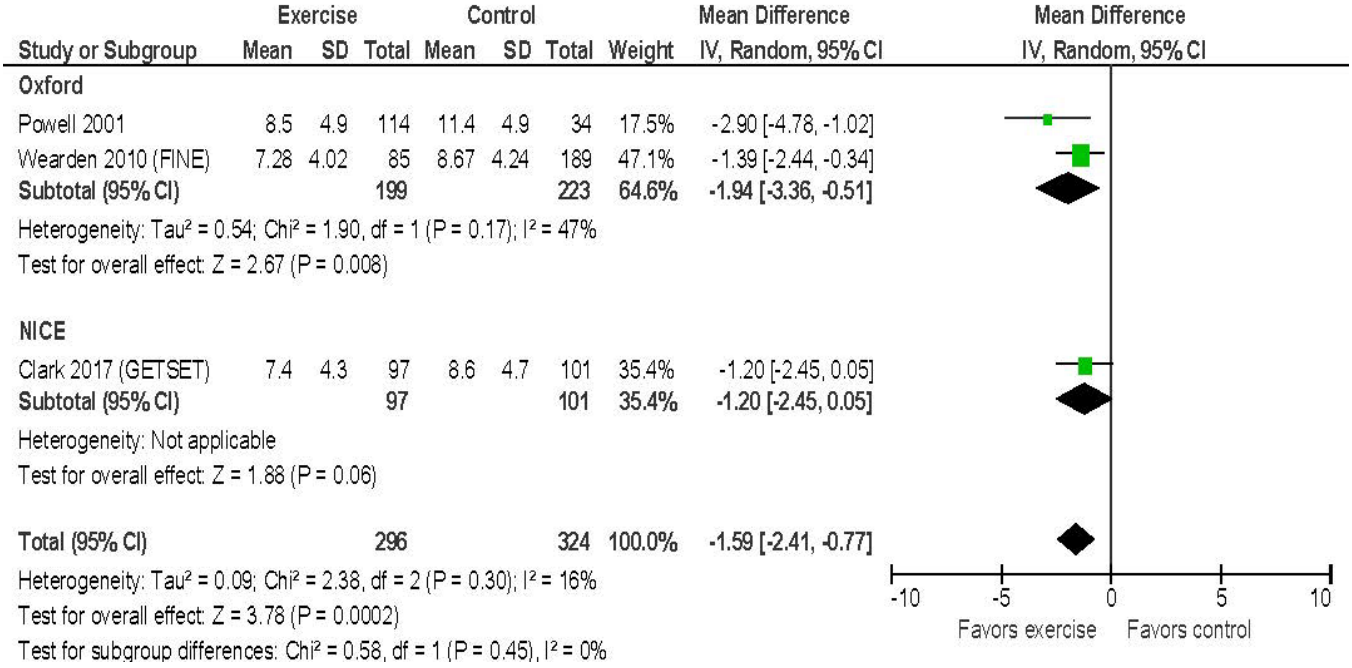
Stratified analyses based on control type and ME/CFS criteria were limited by the small number of trials but indicated no statistically significant subgroup effects **(Table 5)**

Figure 7. Depression severity: graded exercise versus inactive control at end of intervention



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; IV = instrumental variable; ; NICE = National Institute for Health and Care Excellence; SD = standard deviation

Figure 8. Anxiety severity: graded exercise versus inactive control at end of intervention

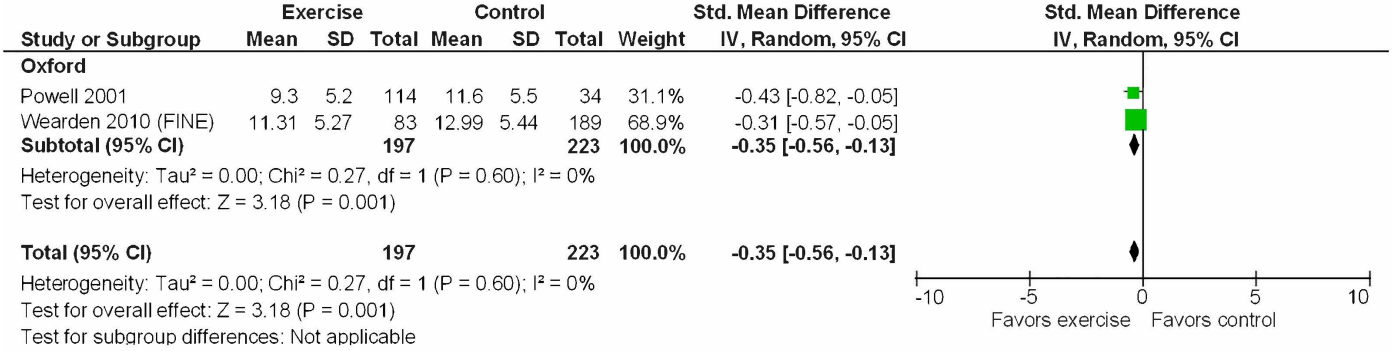


Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; GETSET = guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome; IV = instrumental variable; ; NICE = National Institute for Health and Care Excellence; SD = standard deviation

Sleep

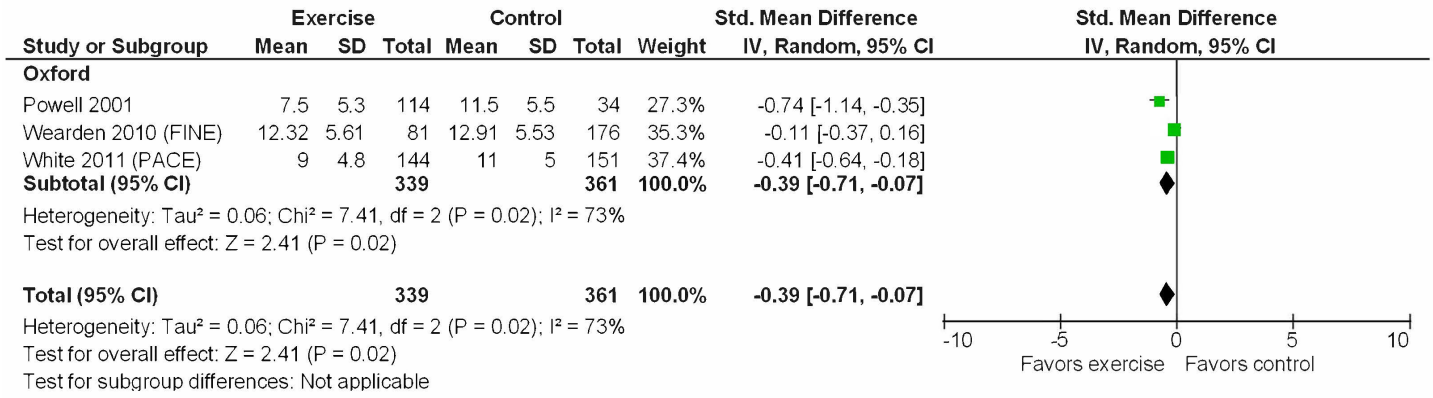
Three trials evaluated effects of graded exercise on sleep quality using the Sleep Problem Questionnaire or the Jenkins Sleep Questionnaire.[38,56,58] Graded exercise was associated with improved sleep quality versus controls at the end of the intervention (2 trials, N=420, SMD -0.35, 95% CI -0.56 to -0.13, I^2=0%;[56,58] **Figure 9**) and at post-intervention follow-up (3 trials, N=700, SMD -0.39, 95% CI -0.71 to -0.07, I^2=73%;[38,56,58] **Figure 10**). On the original 0 to 20 scales, the pooled differences were about 2 points. Subgroup analyses based on the control type or ME/CFS criteria used showed no statistically significant subgroup differences but were limited by the small numbers of trials **(Table 5)**. Excluding the trial by Powell et al. had little effect on the estimates at the end of the intervention (1 trial, N=272, SMD -0.31, 95% CI -0.57 to -0.05)[58] or at post-intervention follow-up (2 trials, N=552, SMD -0.26, 95% CI -0.56 to 0.03, I^2=65%).[38,58]

Figure 9. Sleep quality: graded exercise versus inactive controls at end of intervention



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 10. Sleep quality: graded exercise versus controls at post-intervention follow-up



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; IV = instrumental variable; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; SD = standard deviation; Std = standard

Pain

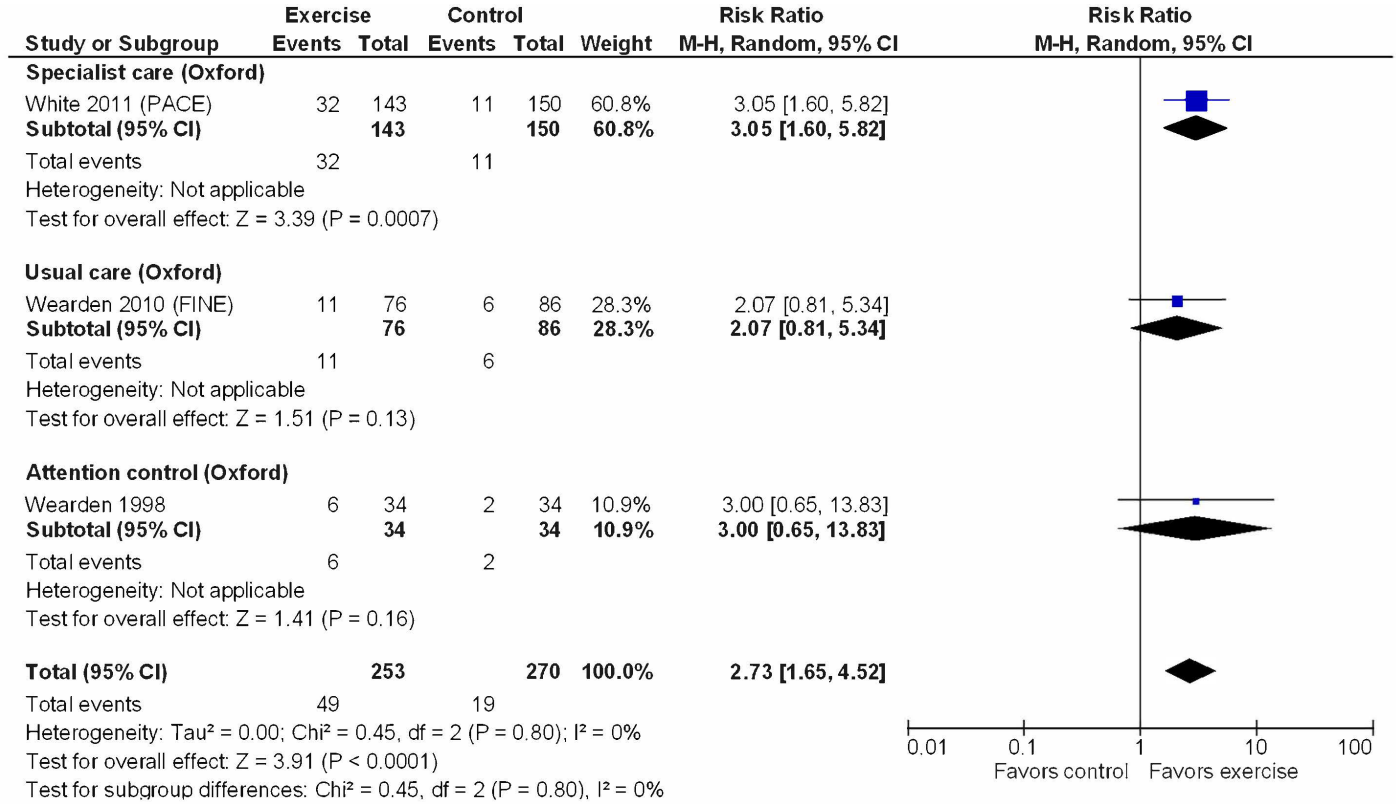
A post-hoc analysis from the PACE trial found exercise associated with decreased severity of muscle pain (mean difference -0.42, 95% CI -0.11 to -0.73) and joint pain (mean difference -0.25, 95% CI -0.70 to -0.57) at post-intervention follow-up (each measured on a 0 to 4 scale).[64]

Recovery

Three trials evaluated effects of exercise on likelihood of recovery.[38,58,59] In two trials, recovery was defined as a score of <4 on the 11-item 0 to 11 Chalder fatigue scale.[58,59] The third trial, PACE, reported results for recovery based on the following definition: SF-36 physical function score ≥60, 11-item 0 to 33 Chalder fatigue score ≤18, Clinical Global Impression rating of better or very much better, and failure to meet one or more case definitions for CFS (the Oxford case definition, SF-36 score ≤65, or positive response on at least 6 of 11 items on the Chalder fatigue scale).[38] This recovery definition has been criticized because the SF-36 physical function threshold includes patients with significant functional impairment; in addition, some patients met the SF-36 physical function and Chalder fatigue scale thresholds for recovery at study entry.[65] Also, the definition used in the main PACE publication differed from the definition for recovery in the original trial protocol: SF-36 physical function score ≥85, 11-item 0 to 11 Chalder score ≤3, Clinical Global Impression rating of very much better, and failure to meet Oxford, Fukuda, and London case definitions for CFS.[39]

Based on the published results from PACE (proportion meeting composite definition for recovery 22% for graded exercise vs. 7% for usual specialist care), graded exercise was associated with increased likelihood of recovery versus usual care, usual specialist care, or attention control (3 trials, N=536, RR 2.73, 95% 1.65 to 4.52, I^2=0%; ARD 12.4%, 95% CI 6.7% to 18.2%;[38,58,59] **Figure 11**). Replacing the data from PACE with results based on the original definition for recovery (proportion meeting definition 4% vs. 3%), resulted in an attenuated, imprecise estimate that was no longer statistically significant (3 trials, N=550, RR 1.86, 95% CI 0.96 to 3.61, I^2=0%).[40,58,59]

Figure 11. Likelihood of recovery: graded exercise versus inactive control



Abbreviations: CI = confidence interval; FINE = Fatigue Intervention by Nurses Evaluation; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Overall improvement

In the PACE trial, a composite outcome for overall improvement (defined as an 11-item 0 to 11 Chalder fatigue scale score ≤3 or >50% improvement from baseline and SF-36 physical function score ≥75 or >50% improvement from baseline at 52 weeks), was described as the primary outcome in the study protocol but not reported in the main publication.[39] In a subsequent publication, the authors reported that graded exercise was associated with greater likelihood of overall improvement than usual specialist care alone, using the protocol definition (N=320, 21% vs. 10%, RR 2.16, 95% CI 1.34 to 3.47).[40]

The Graded Exercise Therapy guided Self-help Treatment (GETSET) trial found guided graded exercise self-help associated with greater likelihood of self-rated Clinical Global Impression of "much better" or "very much better" versus specialist medical care alone, but the difference was not statistically significant (N=198, 14% vs. 5.9%, RR 2.43, 95% CI 0.97 to 6.07).[52]

6-minute walk test

The PACE trial (N=228) found graded exercise associated with longer distance on the 6-minute walk test versus usual specialist care at the end of the intervention, though the difference was small (31.00 meters, 95% CI 4.00 to 58.00).[38] In addition, PACE reported higher loss to follow-up (27.8%) for the 6-minute walk test than for other outcomes (e.g., work and social adjustment scale, 8.1%).

Harms

Data on harms were available from two trials of graded exercise plus usual specialist care versus usual specialist care alone (PACE and GETSET).[38,52] There was no statistically significant difference in the pooled risk of serious adverse events (2 trials, N=518, RR 1.59, 95% CI 0.69 to 3.66, I^2=0%),[38,52] but the estimate was imprecise, with 20 of 23 events reported in one trial (PACE).[38] For withdrawal due to worsening, the PACE trial[38] reported three cases and the GETSET trial[52] reported none, resulting in a very imprecise estimate (1 trial, N=320, RR 2.00, 95% CI 0.18 to 21.84). Exercise was not associated with increased likelihood of physical function worsening, though the pooled estimate was imprecise (2 trials, N=518, RR 0.83, 95% CI 0.52 to 1.34, I^2=0%).[38,52] The PACE trial found graded exercise associated with decreased likelihood of post-exertional malaise versus usual specialist care (N=320, 44% vs. 63%, RR 0.70, 95% CI 0.57 to 0.87).[38]

Table 5. Exercise vs. inactive controls: summary of stratified results

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
Fatigue, end of intervention	6 (1034)	SMD -0.62 (-0.95 to -0.30)	81%	--
By control type: vs. usual care	3 (368)	SMD -0.90 (-1.71 to -0.08)	91%	0.48
vs. usual specialist care	2 (501)	SMD -0.41 (-0.59 to -0.22)	9%	--
vs. attention control	2 (250)	SMD -0.37 (-0.62 to -0.12)	0%	--
On original scale: Chalder (11-item, 0 to 33)	2 (501)	MD -2.91 (-4.36 to -1.47)	24%	--
Chalder (14-item, 0 to 42)	2 (111)	MD -6.47 (-13.80 to 0.86)	72%	--
Chalder (11-item, 0 to 11)	2 (422)	MD -3.60 (-8.49 to 1.29)	98%	--
By ME/CFS criteria: Oxford	4 (792)	SMD -0.59 (-1.05 to -0.14)	88%	0.41
Fukuda	1 (43)	SMD -1.00 (-1.64 to -0.36)	--	--
NICE	1 (199)	SMD -0.52 (-0.80 to -0.24)	--	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	5 (886)	SMD -0.32 (-0.52 to -0.12)	46%	--
Using difference in change from baseline	6 (1034)	SMD -0.68 (-1.01 to -0.35)	82%	--
Fatigue, post-intervention	3 (625)	SMD -0.76 (-1.48 to -0.05)	94%	--
By control type: vs. usual care	2 (315)	SMD -0.95 (-2.38 to 0.47)	97%	0.29
vs. usual specialist care	1 (306)	SMD -0.45 (-0.68 to -0.22)	--	--
vs. attention control	1 (171)	SMD -0.19 (-0.50 to 0.11)	--	--
On original scale: Chalder (11-item, 0 to 33)	1 (306)	MD -3.20 (-4.78 to -1.62)	--	--
Chalder (11-item, 0 to 11)	2 (319)	MD -3.75 (-9.72 to 2.22)	98%	--
By ME/CFS criteria Oxford	3 (625)	SMD -0.76 (-1.48 to -0.05)	94%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (477)	SMD -0.35 (-0.58 to -0.12)	35%	--
Using difference in change from baseline	3 (625)	SMD -0.83 (-1.58 to -0.09)	94%	--
Fatigue improvement (dichotomous)	1 (305)	RR 1.23 (1.07 to 1.42)	--	--
Using original PACE definition	1 (320)	RR 1.81 (1.11 to 2.94)	--	--
SF-36 physical function (0 to 100), end of intervention	5 (965)	MD 11.73 (2.33 to 21.14)	88%	--
By control type: vs. usual care	3 (368)	MD 15.04 (-6.13 to 36.22)	93%	0.72
vs. usual specialist care	2 (501)	MD 6.21 (2.08 to 10.35)	0%	--
vs. attention control	1 (181)	MD 6.66 (-0.40 to 13.72)	--	--
By ME/CFS criteria: Oxford	3 (723)	MD 13.60 (-1.18 to 28.37)	93%	0.34
Fukuda	1 (43)	MD 14.05 (0.62 to 27.48)	--	--

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
NICE	1 (199)	MD 4.90 (-1.85 to 11.65)	--	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	4 (817)	MD 5.89 (2.52 to 9.25)	0%	--
Using difference in change from baseline	5 (969)	MD 12.23 (2.54 to 21.91)	90%	--
SF-36 physical function (0 to 100), post-intervention	3 (711)	MD 17.07 (-2.02 to 36.16)	95%	--
By control type: vs. usual care	2 (315)	MD 21.44 (-13.90 to 56.78)	97%	0.73
vs. usual specialist care	1 (306)	MD 6.90 (1.16 to 12.64)	--	--
vs. attention control	1 (171)	MD 7.55 (-0.47 to 15.57)	--	--
By ME/CFS criteria: Oxford	3 (711)	MD 17.07 (-2.02 to 36.16)	95%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (563)	MD 6.37 (1.89 to 10.85)	0%	--
Using difference in change from baseline	3 (711)	MD 17.91 (-2.00 to 37.81)	97%	--
Functional improvement	3 (618)	RR 2.48 (0.77 to 7.97)	89%	--
Using original PACE definition	3 (632)	RR 2.52 (0.90 to 7.02)	85%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (484)	RR 1.41 (1.15 to 1.74)	0%	--
HADS depression (0 to 21), end of intervention	4 (688)	MD -1.83 (-3.65 to -0.01)	86%	--
By control type: vs. usual care	2 (325)	MD -3.20 (-7.21 to 0.82)	93%	0.56
vs. usual specialist care	1 (198)	MD -1.00 (-2.16 to 0.16)	--	--
vs. attention control	2 (250)	MD -0.85 (-2.47 to 0.77)	61%	--
By ME/CFS criteria: Oxford	3 (398)	MD -2.16 (-4.97 to 0.65)	90%	0.45
NICE	1 (198)	MD -1.00 (-2.16 to 0.16)	--	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	3 (540)	MD -0.97 (-1.71 to -0.23)	8%	--
HADS depression (0 to 21) , post-intervention	3 (699)	MD -2.36 (-4.98 to 0.27)	92%	--
By control type: vs. usual care	2 (314)	MD -2.96 (-8.47 to 2.55)	96%	0.74
vs. usual specialist care	1 (295)	MD -1.10 (-2.11 to -0.09)	--	--
vs. attention control	1 (171)	MD -0.79 (-2.13 to 0.55)	--	--
By ME/CFS criteria: Oxford	3 (699)	MD -2.36 (-4.98 to 0.27)	92%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (551)	MD -0.85 (-1.61 to -0.08)	0%	--
HADS anxiety (0 to 21), end of intervention	3 (620)	MD -1.59 (-2.41 to -0.77)	16%	--
By control type: vs. usual care	2 (325)	MD -1.90 (-3.54 to -0.26)	54%	--
vs. usual specialist care	1 (198)	MD -1.20 (-2.45 to 0.05)	--	--
vs. attention control	1 (182)	MD -1.57 (-2.74 to -0.40)	--	--
By ME/CFS criteria: Oxford	2 (422)	MD -1.94 (-3.36 to -0.51)	47%	0.45
NICE	1 (198)	MD -1.20 (-2.45 to 0.05)	--	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (472)	MD -1.31 (-2.12 to -0.51)	0%	--
HADS anxiety (0 to 21), post-intervention	3 (697)	MD -1.07 (-2.64 to 0.49)	75%	--
By control type: vs. usual care	2 (314)	MD -1.14 (-4.71 to 2.44)	89%	0.64
vs. usual specialist care	1 (293)	MD -0.90 (-1.92 to 0.12)	--	--
vs. attention control	1 (171)	MD -0.08 (-1.52 to 1.36)	--	--
By ME/CFS criteria Oxford	3 (697)	MD -1.07 (-2.64 to 0.49)	75%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (549)	MD -0.38 (-1.52 to 0.76)	49%	--
Sleep, end of intervention	2 (420)	SMD -0.35 (-0.56 to -0.13)	0%	--

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
By control type: vs. usual care	2 (323)	SMD -0.27 (-0.53 to -0.00)	19%	0.33
vs. attention control	1 (180)	SMD -0.46 (-0.76 to -0.17)	--	--
By ME/CFS criteria: Oxford	2 (420)	SMD -0.35 (-0.56 to -0.13)	0%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	1 (272)	SMD -0.31 (-0.57 to -0.05)	--	--
Sleep, post-intervention	3 (700)	SMD -0.39 (-0.71 to -0.07)	73%	--
By control type: vs. usual care	2 (315)	SMD -0.39 (-1.06 to 0.29)	86%	0.41
vs. usual specialist care	1 (295)	SMD -0.41 (-0.64 to -0.18)	--	--
vs. attention control	1 (171)	SMD -0.15 (-0.45 to 0.15)	--	--
By ME/CFS criteria: Oxford	3 (700)	SMD -0.39 (-0.71 to -0.07)	73%	--
Sensitivity analysis: Outlier trial (Powell 2001) excluded	2 (552)	SMD -0.26 (-0.56 to 0.03)	65%	--
Recovery	3 (536)	RR 2.73 (1.65 to 4.52)	0%	--
Original PACE definition	3 (550)	RR 1.86 (0.96 to 3.61)	0%	--
6-minute walk test (meters)	1 (228)	MD 31.00 (4.00 to 58.00)	--	--
Serious adverse events	2 (518)	RR 1.59 (0.69 to 3.66)	0%	--
Withdrawal due to worsening	1 (320)	RR 2.00 (0.18 to 21.84)	--	--
Physical function worsening	2 (518)	RR 0.83 (0.52 to 1.34)	0%	--
Post-exertional malaise	1 (320)	RR 0.70 (0.57 to 0.87)	--	--

Abbreviations: CFS = chronic fatigue syndrome; CI = confidence interval; HADS = Hospital Anxiety and Depression Scale; MD = mean difference; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference

Exercise Versus Active Interventions

Six trials (N=1,024) compared exercise versus active interventions (**Tables 3 and 4, Evidence Table Appendix E2**).[38,53,54,57,59,60] The active interventions were CBT (2 trials),[38,54] relaxation (3 trials),[53,54,57] adaptive pacing (1 trial),[38] heart rate variability biofeedback (1 trial),[60] and fluoxetine (1 trial).[59] Of these, one trial of relaxation[57] and one trial of biofeedback[60] were added for this update. Five trials evaluated graded exercise and one trial[54] evaluated anaerobic exercise. The duration of the exercise intervention ranged from 6 to 26 weeks. Four trials evaluated patients at the end of the intervention and five trials evaluated patients 26 to 70 weeks following the end of the intervention. Four trials were rated medium risk of bias and two trials were rated high risk of bias; in addition to open-label design, methodological limitation for the high-risk of bias trials include unclear randomization methods and failure to conduct intention to treat analysis. One high risk of bias trial[57] evaluated exercise versus relaxation and the other[60] evaluated exercise versus biofeedback. Results stratified by the active comparator are summarized in **Table 6** and shown in **Figures 12 to 29**.

Exercise versus CBT

Two trials compared graded[38] or anaerobic[54] exercise versus CBT in patients who met the Oxford case definition. The duration of the interventions was 5 to 6 months in both trials. One trial[38] evaluated outcomes at the end of the intervention and both trials evaluated outcomes approximately 6 months following the completion of therapy (**Table 6**).

There were no differences between exercise versus CBT in fatigue severity at the end of the intervention (1 trial, N=298, mean difference 0.20, 95% CI -1.49 to 1.89, **Figure 12**)[38] or

between exercise versus CBT in fatigue at post-intervention follow-up (2 trials, N=360, mean difference 0.39, 95% CI -0.24 to 1.02, I^2=0%, **Figure 13**).[38,54] There were also no differences between exercise versus CBT in severity of functional impairment at the end of the intervention (1 trial, N=298, mean difference 1.20, 95% CI -3.90 to 6.30 on the 0 to 100 SF-36 physical component scale, **Figure 14**)[38] or severity of functional impairment (2 trials, N=360, mean difference -8.36, 95% CI -26.21 to 9.50, **Figure 15**),[38,54] depression (2 trials, N=345, SMD 0.02, 95% CI -0.19 to 0.23, I^2=0%, **Figure 16**),[38,54] anxiety (2 trials, N=345, SMD 0.07, 95% CI -0.14 to 0.28, I^2=0%, **Figure 17**),[38,54] sleep quality (2 trials, N=345, SMD -0.17, 95% CI -0.39 to 0.04, I^2=0%, **Figure 18**),[38,54] pain interference (1 trial, N=58, mean difference -0.35, 95% CI -2.02 to 1.32 on the 0 to 10 Brief Pain Inventory [BPI], **Figure 19**),[54] or 6-minute walk test distance (2 trials, N=291, mean difference -4.23 meters, 95% CI -75.99 to 67.52, I^2=71%, **Figure 20**)[38,54] at post-intervention follow-up. One trial found no differences between exercise versus CBT in severity of sore throat, tender lymph nodes, impaired memory, or headaches.[54] The PACE trial found no difference between exercise versus CBT in likelihood of having poor concentration or memory (N=321, 48% vs. 45%, RR 1.05, 95% CI 0.83 to 1.33)[38] and a post-hoc analysis from PACE found no difference between exercise versus CBT in severity of muscle or joint pain.[64]

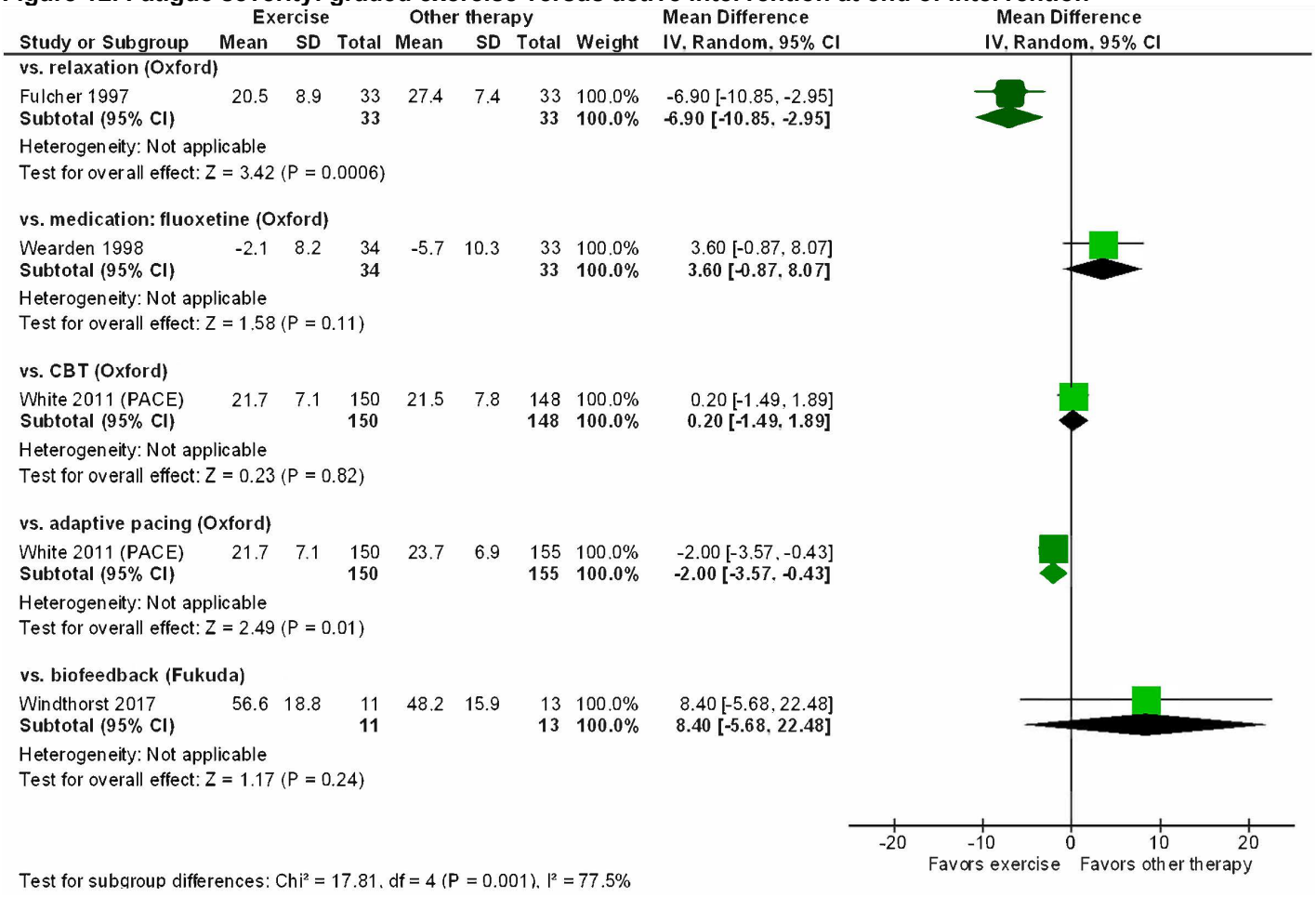
There were no differences between exercise versus CBT in the likelihood of improvement in fatigue using the definition reported in the main PACE publication (1 trial, N=303, 80% vs. 76%, RR 1.05, 95% CI 0.93 to 1.19, **Figure 21**)[38] or the PACE protocol definition (1 trial, N=321, 24% vs. 26% RR 0.91, 95% CI 0.62 to 1.33)[40] (see earlier Results for details regarding PACE outcome definitions). There were also no differences in the likelihood of improvement in function using the definition reported in the main PACE definition (2 trials, N=360, RR 0.98, 95% CI 0.85 to 1.14, I^2=0%, **Figure 22**)[38,54] or the PACE protocol definition (2 trials, N=379, RR 1.17, 95% CI 0.81 to 1.70)[40,54] or in the likelihood of overall improvement (the primary outcome in the PACE protocol), a composite of improvement in fatigue and function (1 trial, N=321, 21% vs. 20%, RR 1.04, 95% CI 0.67 to 1.60).[40] There was no difference between exercise versus CBT in the likelihood of recovery (2 trials, N=360, RR 0.91, 95% CI 0.65 to 1.29, I^2=0%, **Figure 23**).[38,54] As previously noted, the definition for recovery in the PACE trial was modified from the original protocol; there was also no difference in the likelihood of recovery using the original PACE definition for this outcome (2 trials, N=379, RR 1.13, 95% CI 0.33 to 3.84, I^2=45%).[54,66]

Data on harms of exercise versus CBT were largely limited to the PACE trial.[38,62] It found exercise associated with increased likelihood of serious adverse events, but the difference was not statistically significant (N=321, RR 1.87, 95% CI 0.77 to 4.56, **Figure 24**). The estimate for withdrawal due to adverse events was based on only 2 cases and very imprecise (RR 5.03, 95% CI 0.24 to 104.0, **Figure 25**). There was no difference in the likelihood of worsening of function (RR 1.21, 95% CI 0.63 to 2.31, **Figure 26**) or post-exertional malaise (RR 0.90, 95% CI 0.72 to 1.14). However, one other trial found exercise associated with increased severity of post-exertional malaise versus CBT (N=58, mean difference 18.6, 95% CI -31.6 to 7.1 on the 0 to 100 CFS Questionnaire).[54]

The PACE trial also evaluated longer-term, post-trial outcomes (median duration from randomization 31 months).[63] About 30% of patients in both the exercise and CBT groups received non-randomly allocated therapies between the end of the trial at 1 year and long-term follow-up. Fatigue on the 11-item 0 to 33 Chalder scale was slightly improved at long-term follow-up compared with end-of-trial scores in both the exercise (mean change -1.3 points, 95% CI -2.7 to -0.1) and CBT arms (-2.2 points, 95% CI -3.7 to -0.6). For SF-36 physical function (0

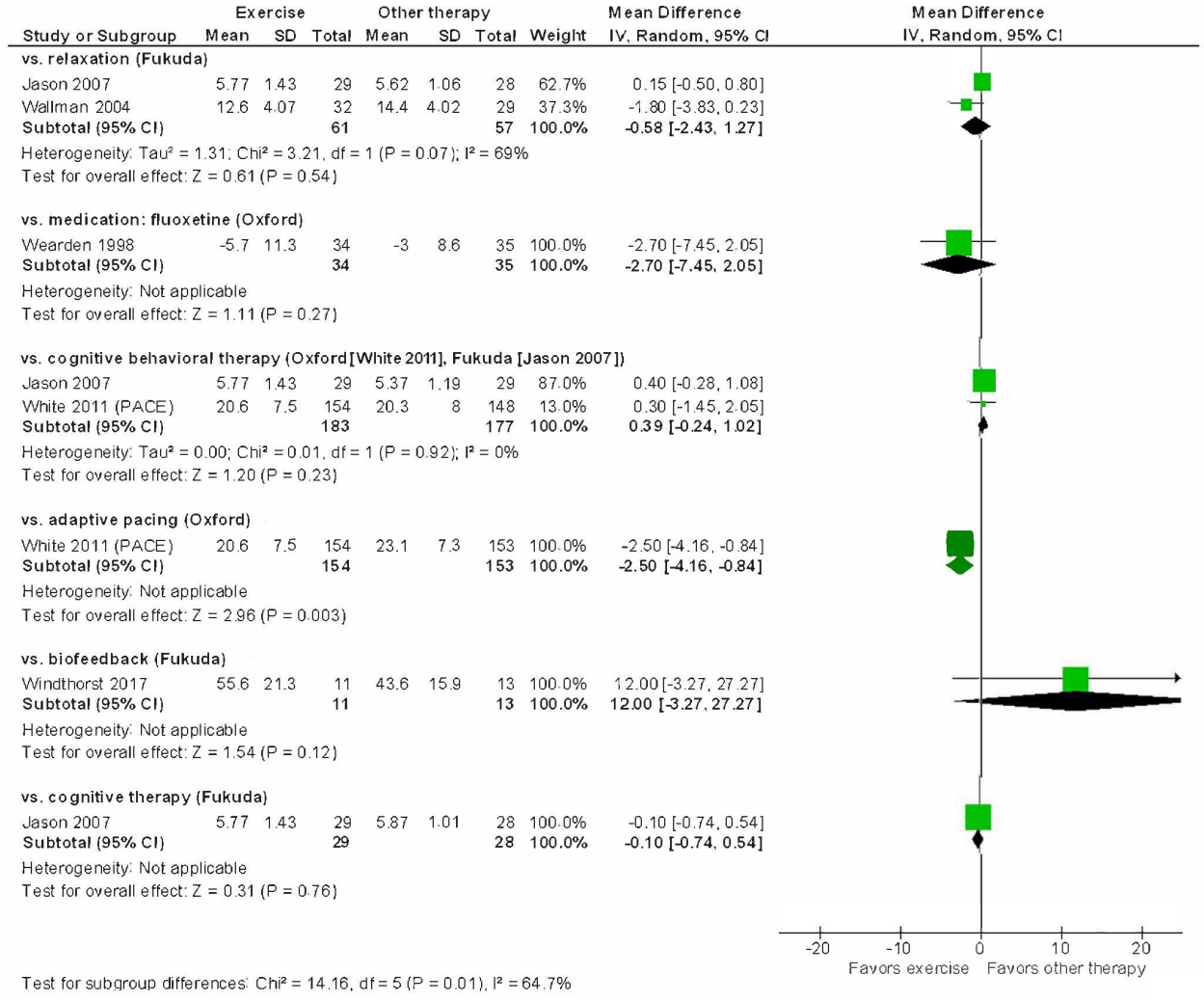
to 100 scale), there was no change at long-term follow-up compared with the end of the trial in the exercise group (mean change 0.5 point, 95% CI -2.7 to 3.6 on a 0 to 100 scale), but function slightly improved in the CBT group (mean change 3.3 points, 95% CI 0.02 to 6.7). At long-term post-trial follow-up, there were no difference between graded exercise or CBT versus specialist medical care in fatigue (N=246, mean difference 0.7, 95% CI -1.4 to 2.8) or function (N=246, mean difference -2.4, 95% CI -9.3 to 4.5), based on mixed model analyses.

Figure 12. Fatigue severity: graded exercise versus active intervention at end of intervention



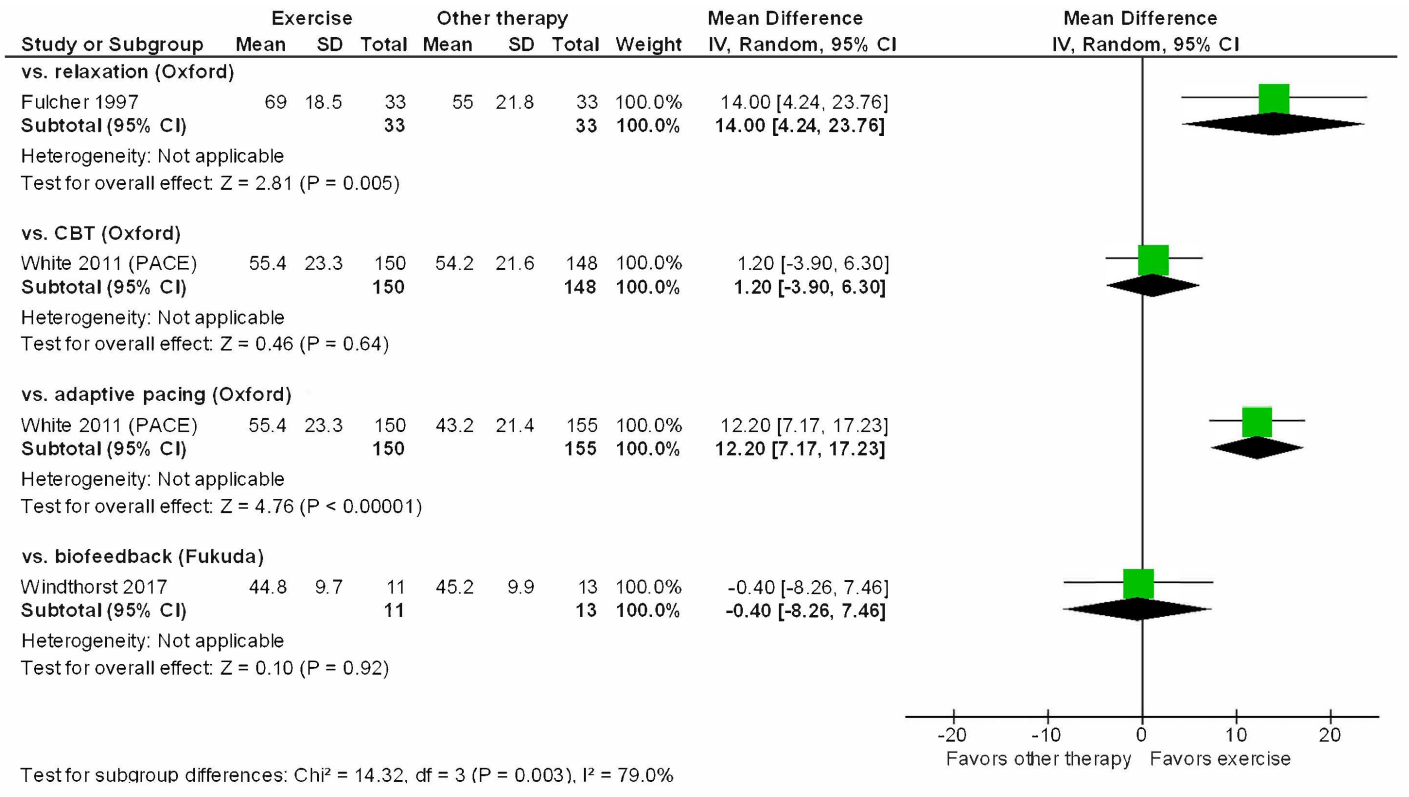
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 13. Fatigue severity: graded or anaerobic exercise versus active intervention at post-intervention follow-up



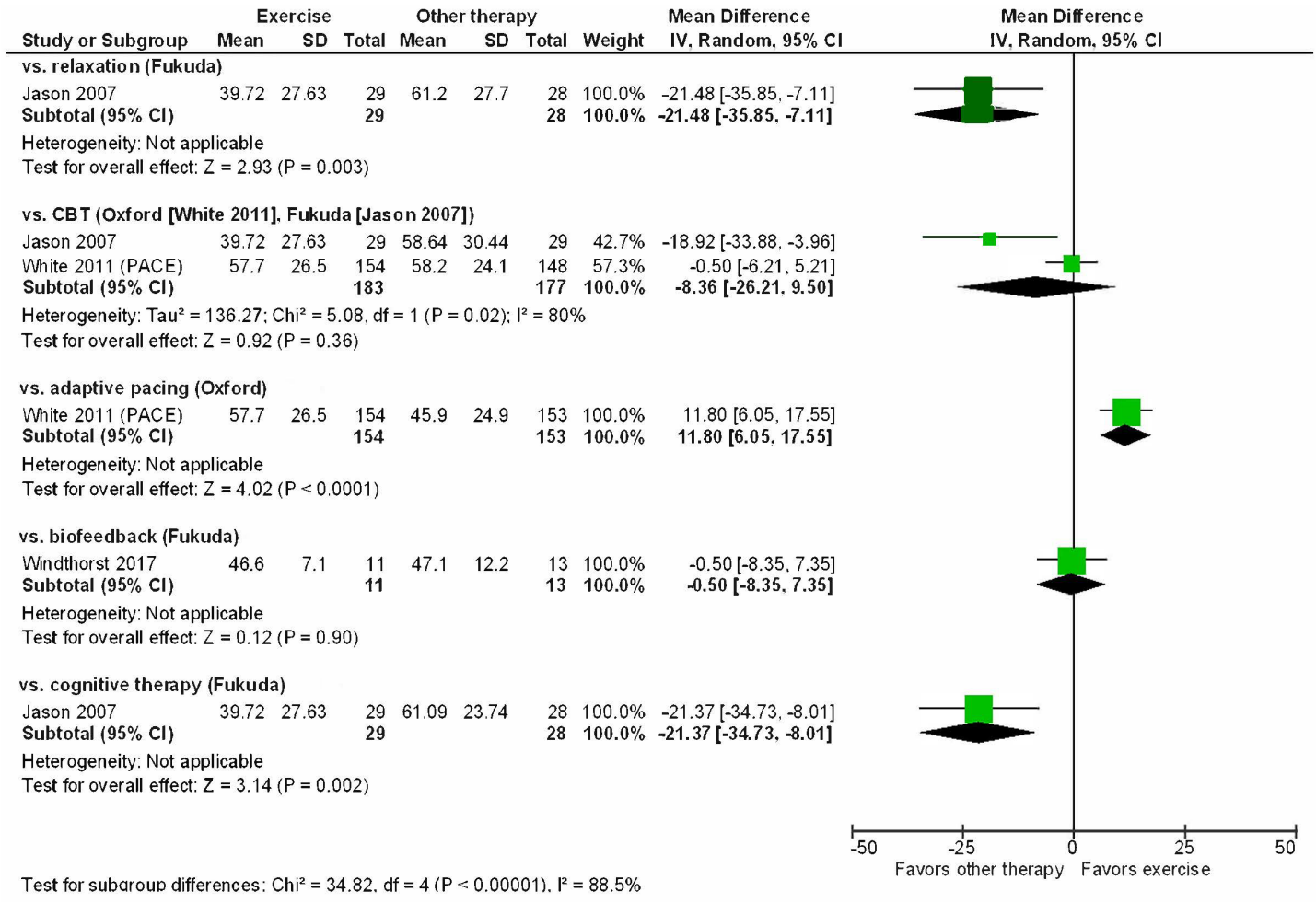
Abbreviations: CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 14. Functional impairment: graded or anaerobic exercise versus active intervention at end of intervention



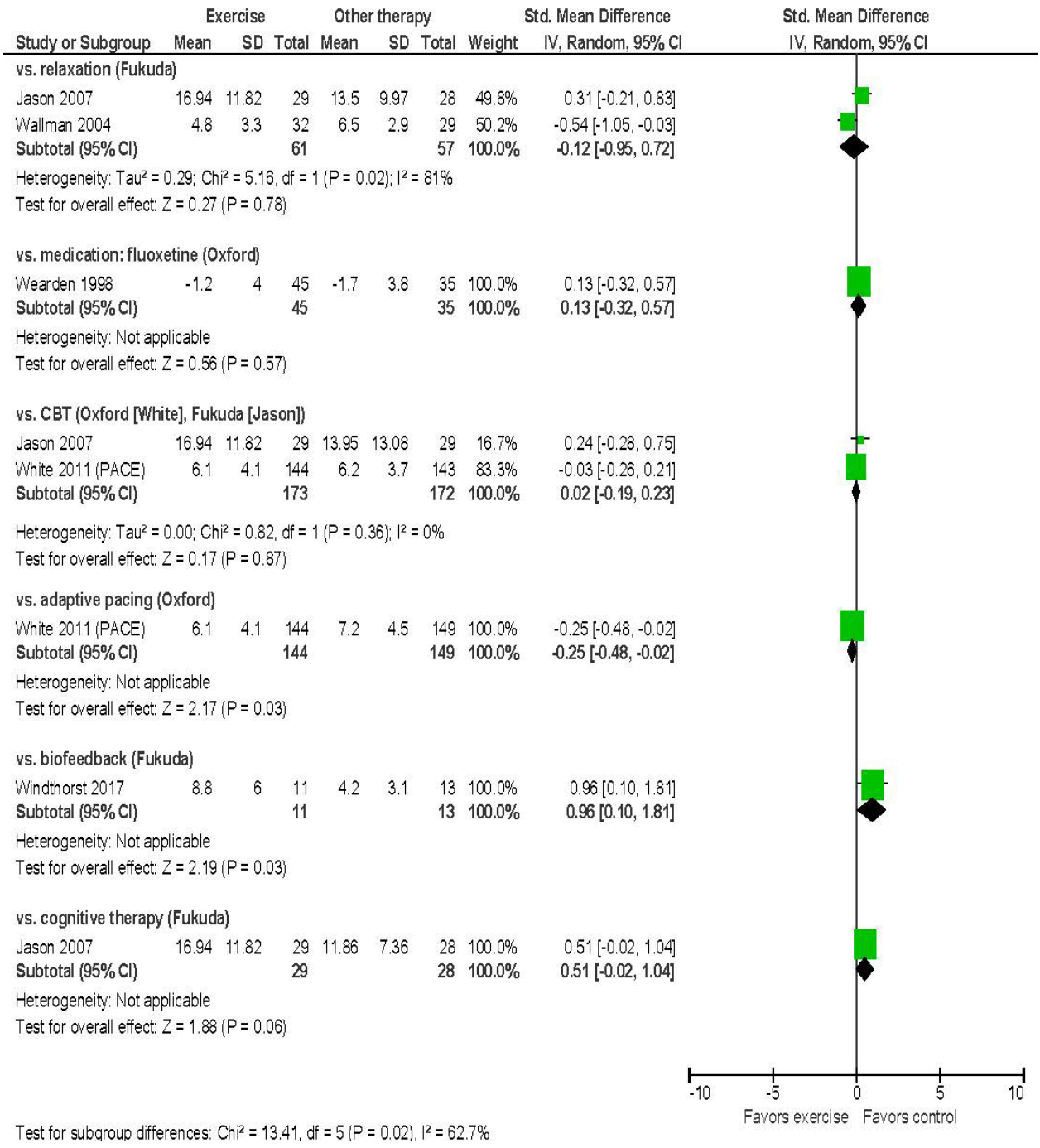
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 15. Functional impairment: graded or anaerobic exercise versus active intervention at post-intervention follow-up



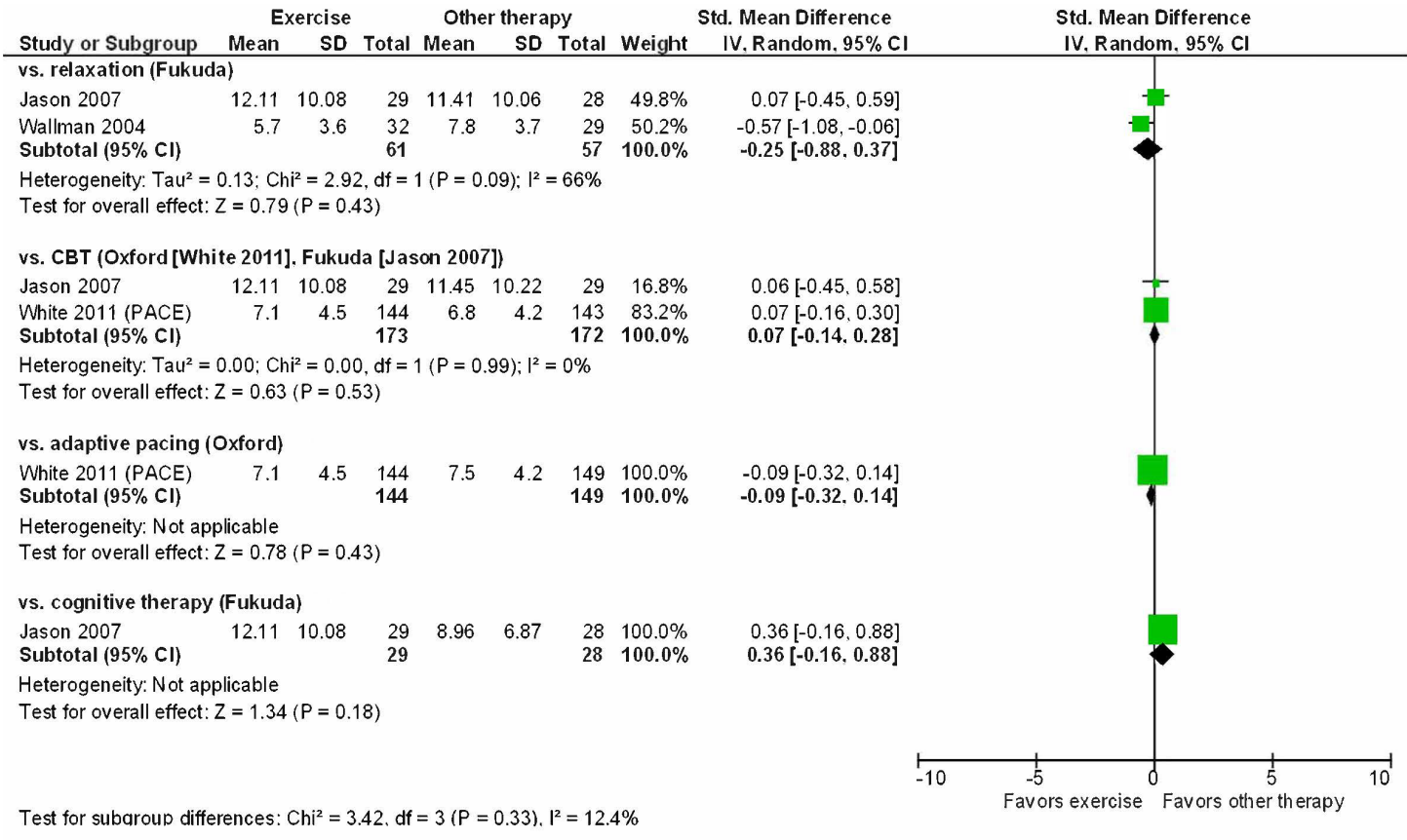
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 16. Depression severity: graded or anaerobic exercise versus active intervention at post-intervention follow-up



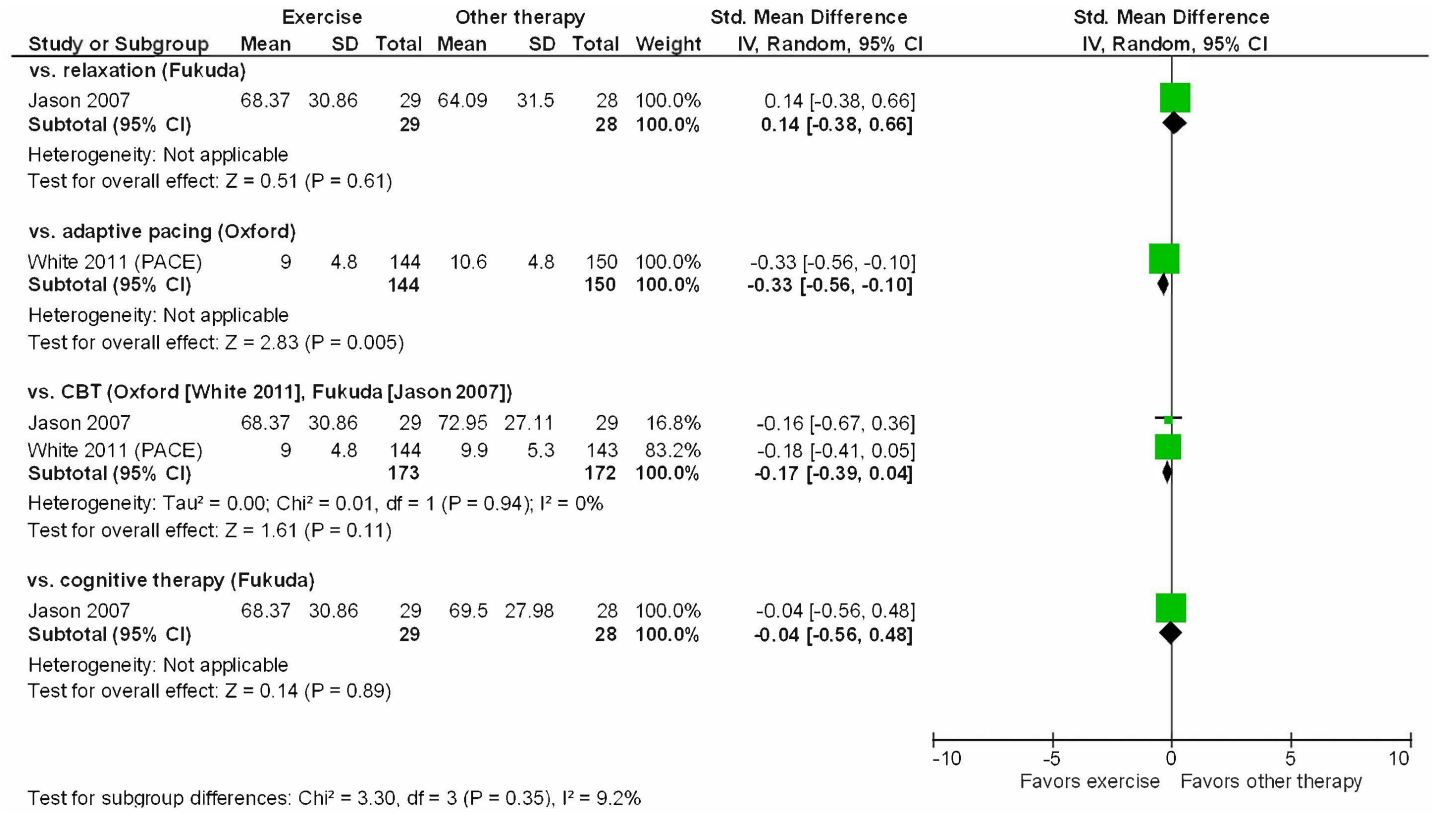
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 17. Anxiety severity: graded or anaerobic exercise versus active interventions at post-intervention follow-up



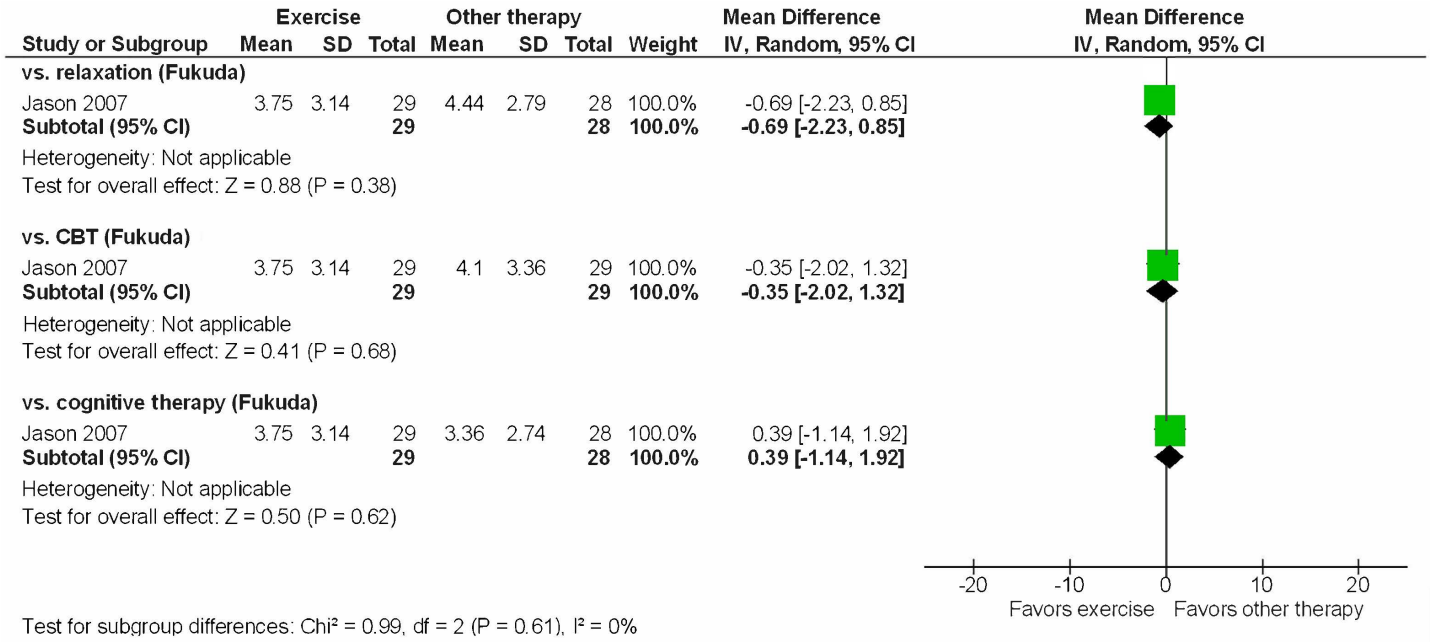
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 18. Sleep: graded or anaerobic exercise versus active intervention at post-intervention follow-up



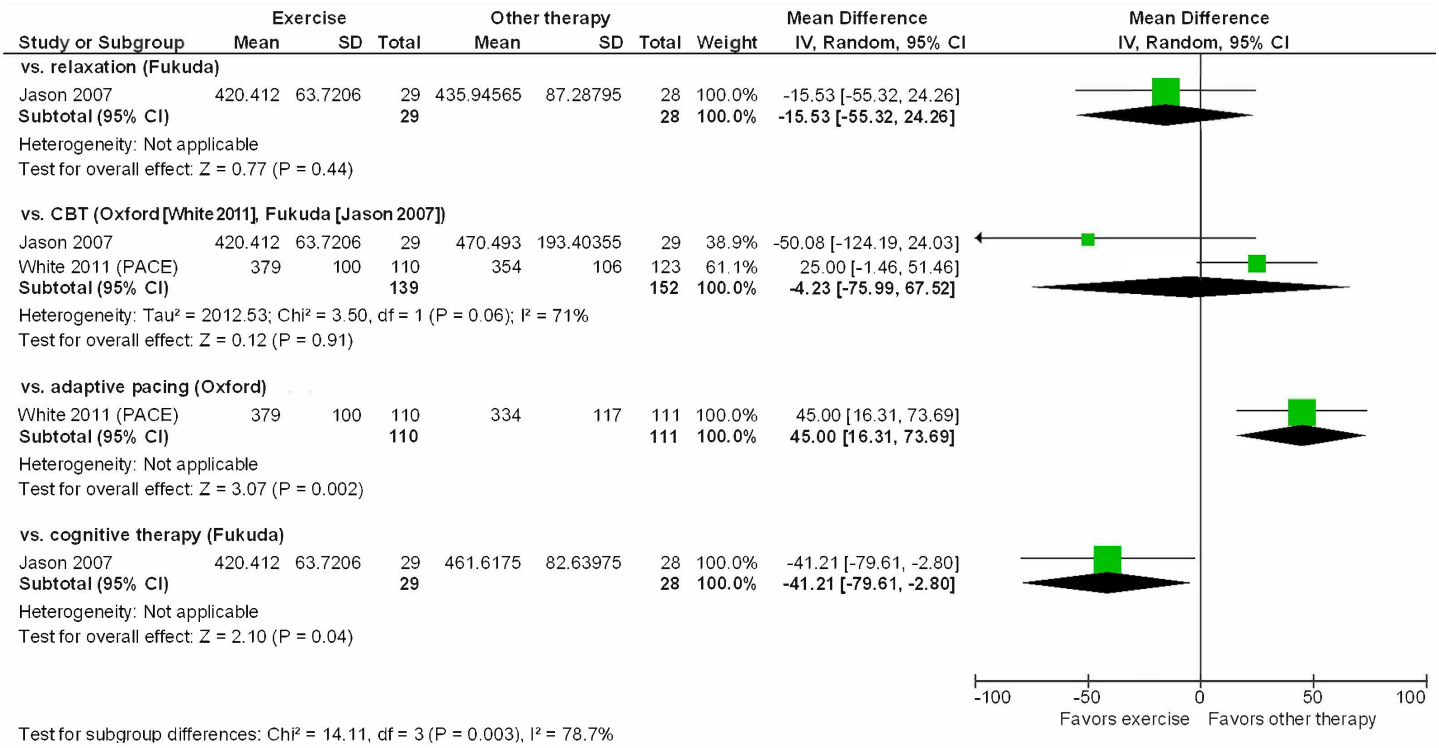
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 19. Pain: anaerobic exercise versus active intervention at post-intervention follow-up



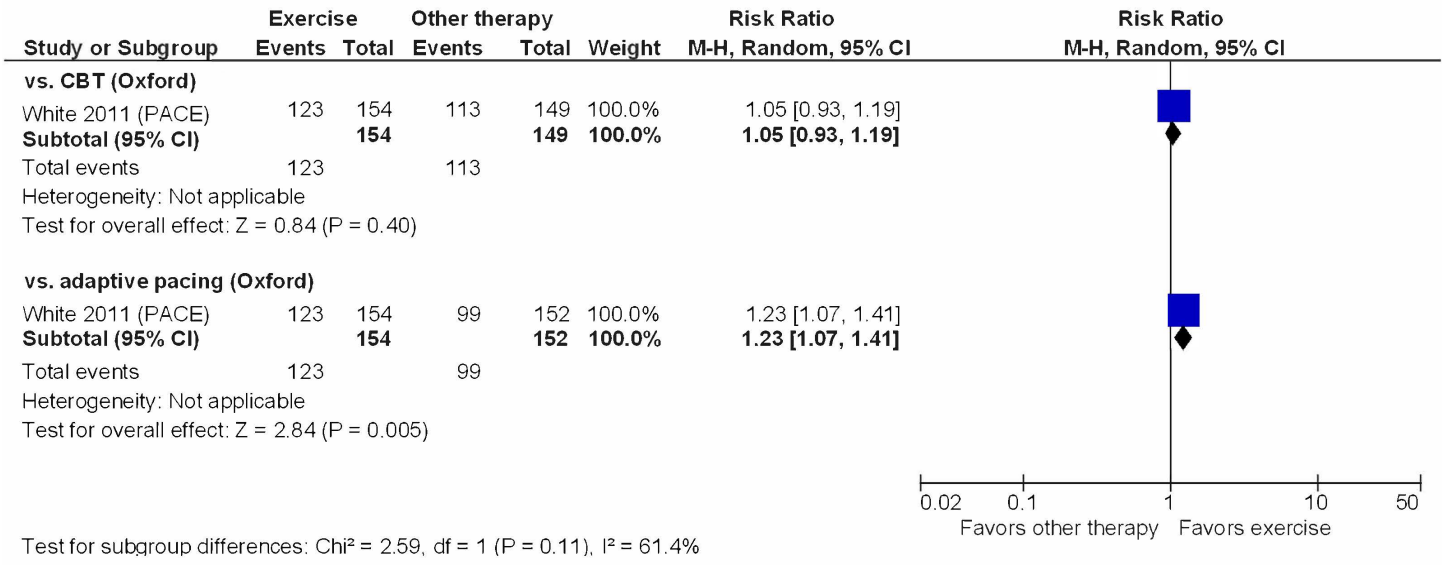
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; IV = instrumental variable; SD = standard deviation

Figure 20. 6-minute walk test: graded or anaerobic exercise versus active intervention at end of intervention



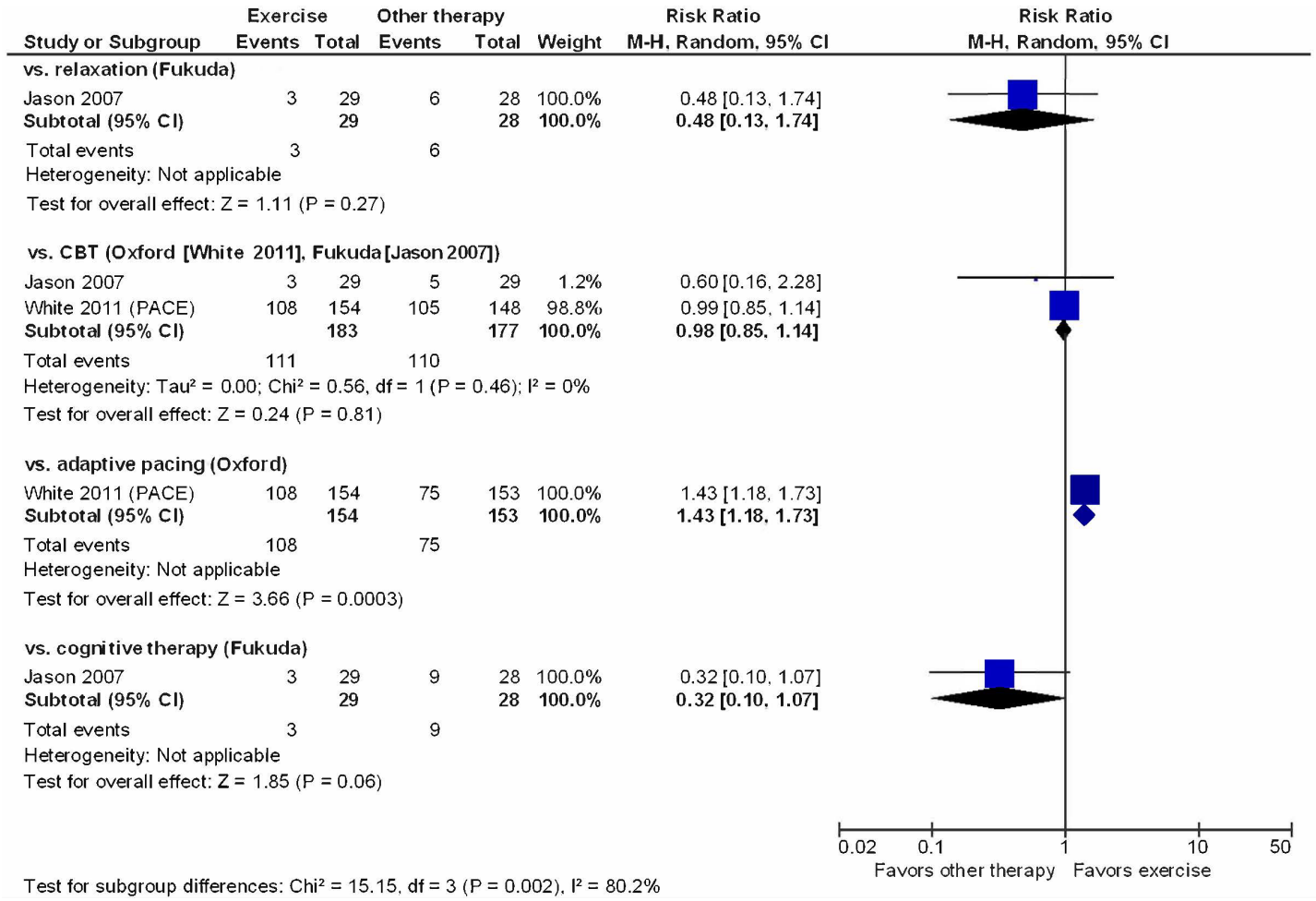
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 21. Likelihood of fatigue improvement: graded exercise versus active interventions



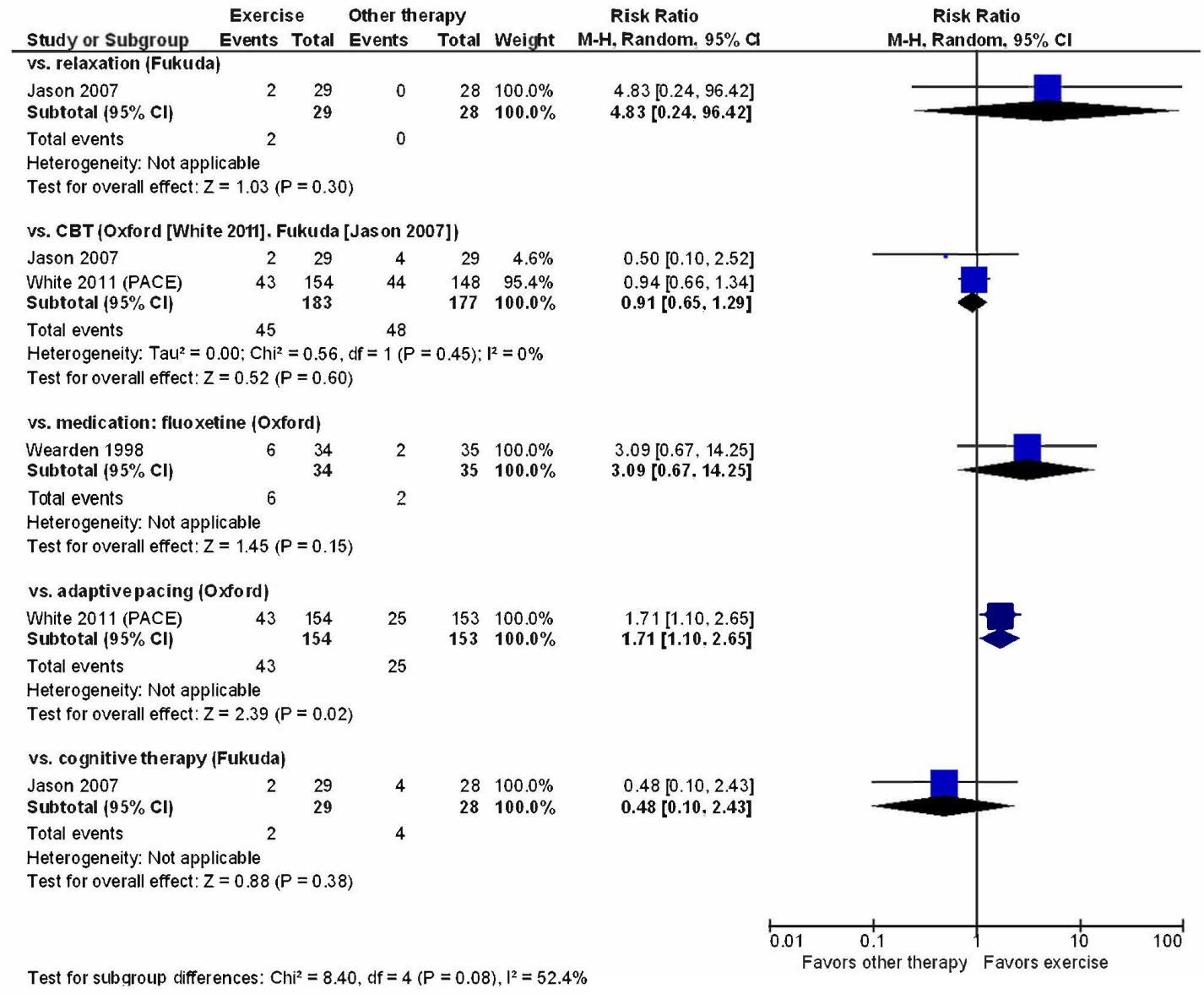
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 22. Likelihood of functional improvement: graded or anaerobic exercise versus active interventions



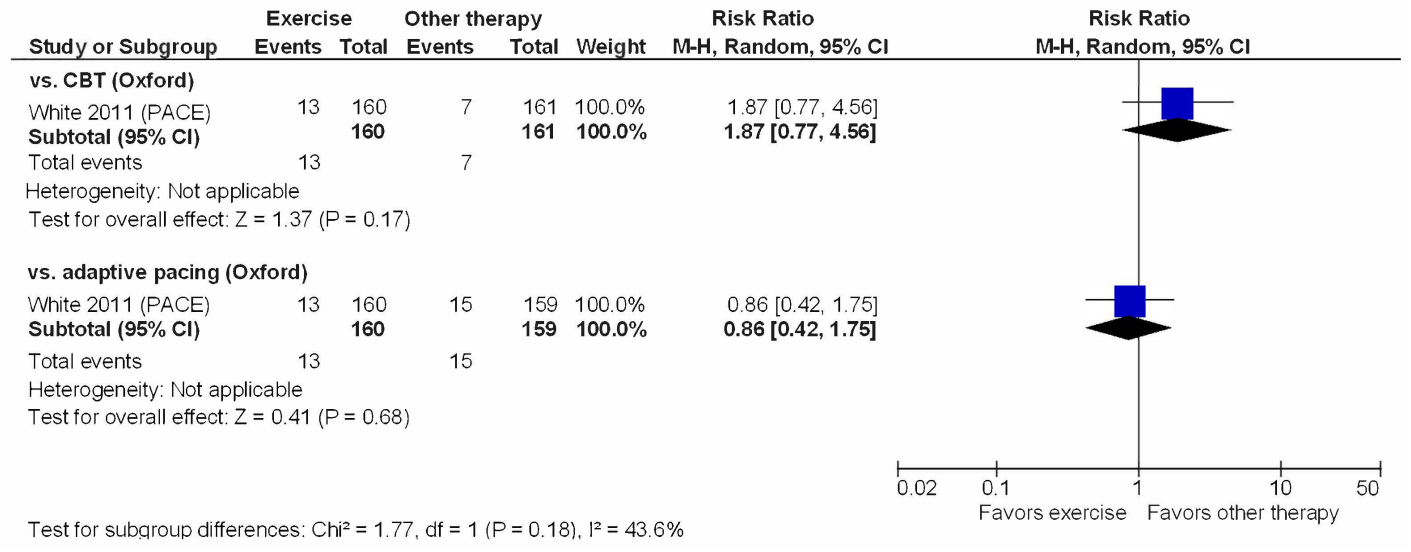
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 23. Likelihood of recovery: graded or anaerobic exercise versus active interventions



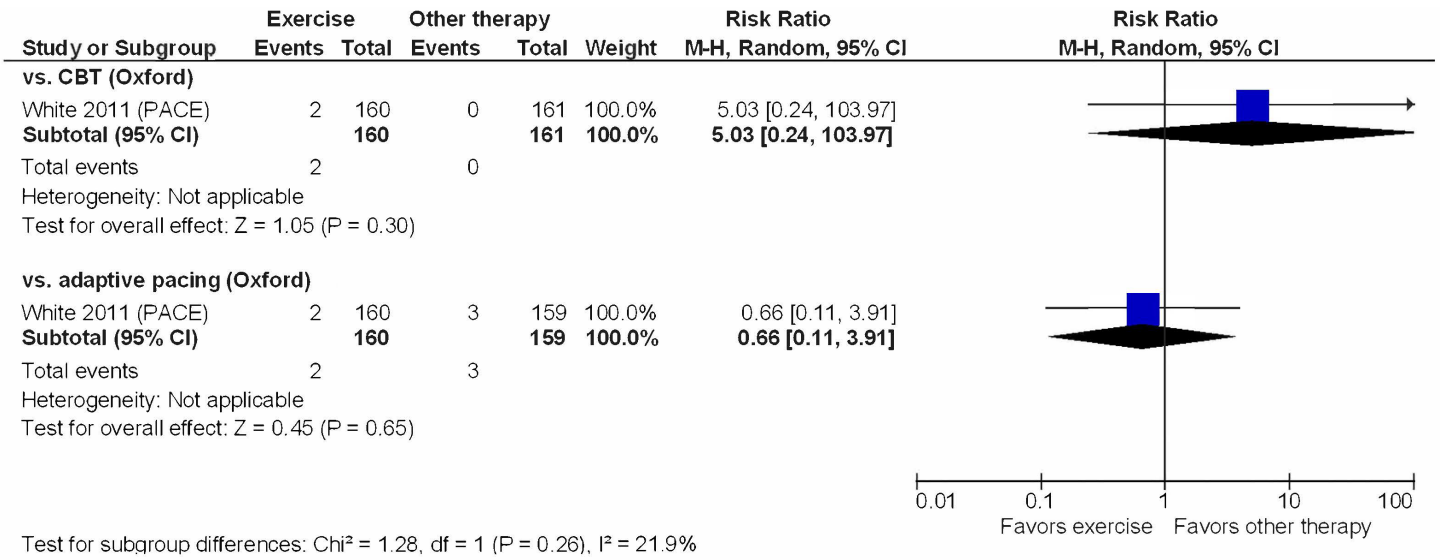
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 24. Likelihood of serious adverse event: graded exercise versus active intervention



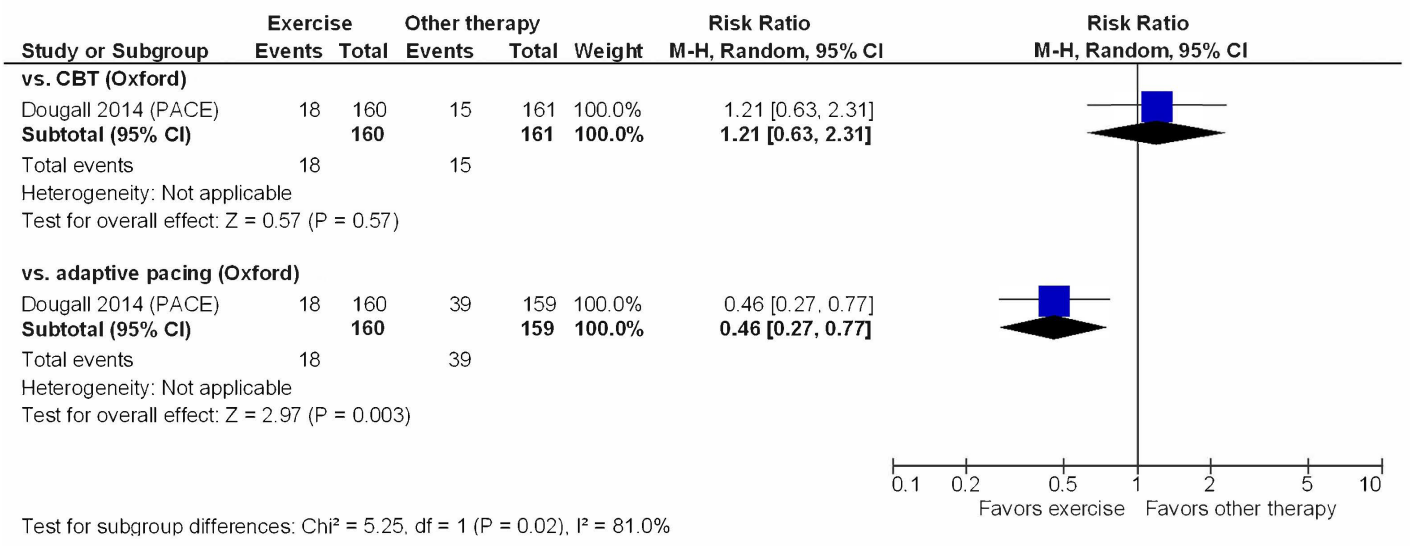
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 25. Likelihood of withdrawal due to adverse event: graded exercise versus active intervention



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 26. Likelihood of function worsening: graded exercise versus active intervention



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Exercise versus cognitive therapy

One trial included a comparison of anaerobic exercise versus cognitive therapy (N=57) in patients who met the Fukuda case definition.[54] Outcomes were evaluated 6 months following completion of 6 months of treatment. There were no differences between exercise versus cognitive therapy in severity of fatigue (**Figure 13**), depression **(Figure 16)**, anxiety **(Figure 17)** or pain **(Figure 19)**; sleep quality **(Figure 18)**; or distance on the 6-minute walk test (**Figure 20, Table 6**). CBT was associated with greater severity of functional impairment (mean difference -21.37, 95% CI -34.73 to -9.01 on the 0 to 100 SF-36 physical function scale, **Figure 15**) and greater severity of post-exertional malaise (mean difference 20.1, 95% CI 3.0 to 37.3 on the 0 to 100 CFS Questionnaire). There were no differences in severity of sore throat (mean difference 3.6, 95% CI -7.2 to 14.4), tender lymph nodes (mean difference 4.7, 95% CI -8.5 to 17.9), impaired memory (mean difference 1.5, 95% CI -13.4 to 16.5), or headaches (mean difference -0.64, 95% CI -20.7 to 19.4). There was also no difference in likelihood of functional improvement **(Figure 22)** or recovery **(Figure 23)**, but estimates were imprecise (**Table 6**).

Exercise versus relaxation

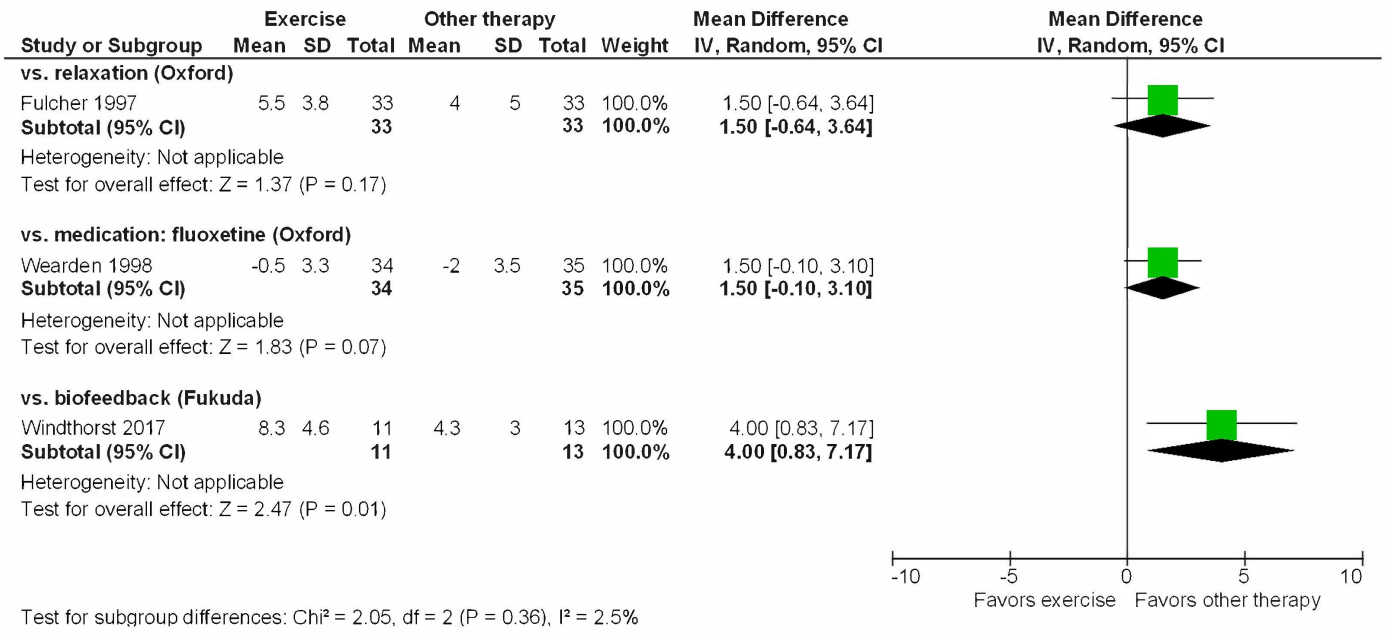
Three trials compared exercise versus relaxation in patients who met the Oxford[53] or Fukuda[54,57] case definition. Two of the trials included a flexibility intervention along with relaxation.[53,57] Two trials[53,57] evaluated graded exercise (treatment duration 12 weeks) and one trial[54] evaluated anaerobic exercise (treatment duration 6 months). One trial[53] evaluated outcomes at the end of the intervention and two trials[54,57] evaluated outcomes 4 weeks and 6 months following the end of the intervention. One trial was rated high risk of bias[57] and the other two were medium risk of bias; conclusions did not change when the high risk of bias trial was excluded.

Exercise was associated with decreased fatigue severity versus relaxation at the end of the intervention (1 trial, N=66, mean difference -6.9, 95% CI -10.85 to 2.95 on the 14-item 0 to 42

Chalder fatigue scale, **Figure 12**)[53] but there was no difference at post-intervention follow-up (2 trials, N=118, SMD -0.16, 95% CI -0.71 to 0.38, I^2=56%, **Figure 13**).[54,57] Exercise was associated with decreased functional impairment at the end of the intervention (1 trial, N=66, mean difference 14.00, 95% CI 4.24 to 23.76 on the 0 to 100 SF-36 physical function subscale, **Figure 14**)[53] but greater functional impairment at post-intervention follow-up (1 trial, N=57, mean difference -21.48, 95% CI -35.85 to -7.11, **Figure 15**);[54] only 1 trial evaluated this outcome at each of these timepoints. There were no differences between exercise versus relaxation in depression (**Figure 27**), anxiety (**Figure 28**), sleep (**Figure 29**), pain **(Figure 19)**, or the 6-minute walk test **(Figure 20)**, at the end of the intervention or at post-intervention follow-up (**Table 6**). One trial found no difference between exercise versus relaxation in severity of sore throat, tender lymph nodes, impaired memory, or headaches.[54] Exercise was associated with increased likelihood of a self-rated Clinical Global Impression rating of much better or very much better (2 trials, N=120, RR 1.64, 95% CI 1.09 to 2.48, I^2=0%).[53,57] There was no difference in the likelihood of functional improvement, defined as improvement in the SF-36 physical function subscale greater than the age adjusted reliable change index and within 1 standard deviation of the normative mean (1 trial, N=57, RR 0.48, 95% CI 0.13 to 1.74, **Figure 22**).[54] The estimate for recovery was very imprecise (1 trial, N=57, RR 4.83, 95% CI 0.24 to 96.42, **Figure 23**).[54]

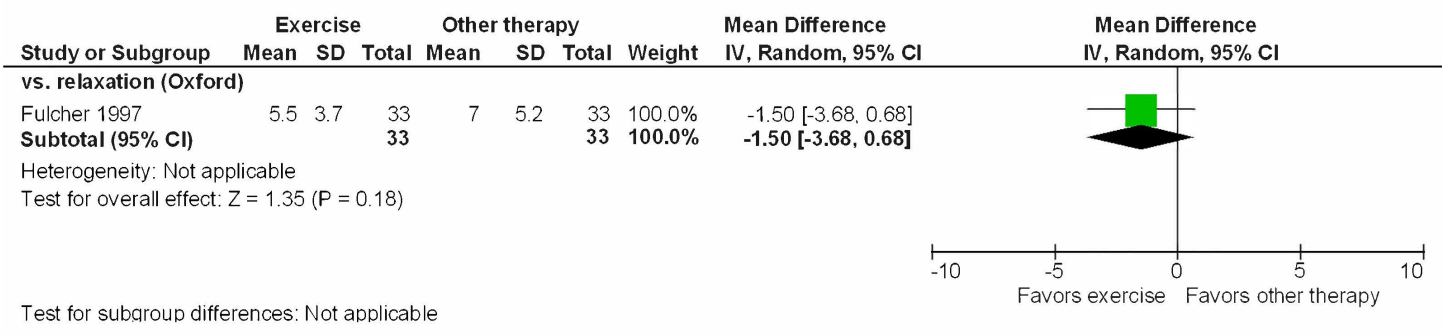
One trial found exercise associated with increased severity of post-exertional malaise versus relaxation (1 trial, N=57, mean difference 22.0, 95% CI 5.7 to 38.4).[54] The trials did not report serious adverse events, withdrawal due to adverse events, or other harms.

Figure 27. Depression severity: graded exercise versus active intervention at end of intervention



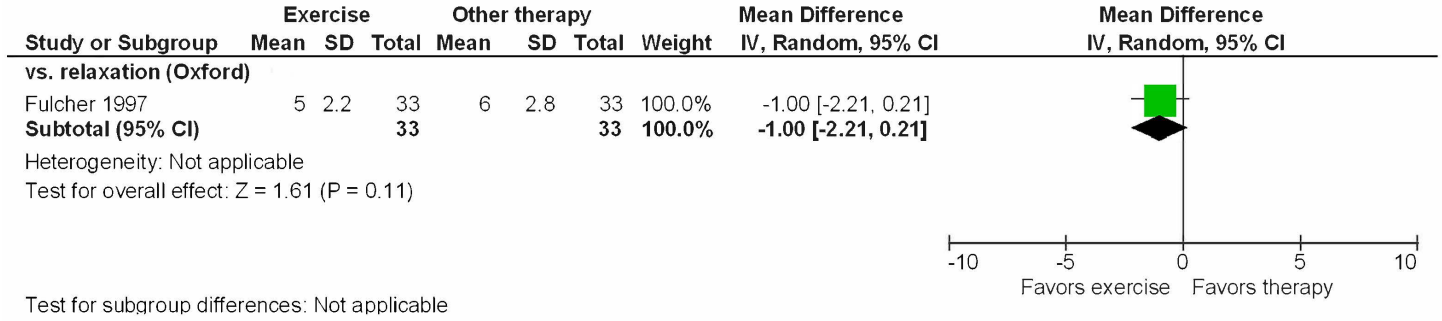
Abbreviations: CI = confidence interval; IV = instrumental variable; SD = standard deviation

Figure 28. Anxiety severity: graded exercise versus active intervention at end of intervention



Abbreviations: CI = confidence interval; IV = instrumental variable; SD = standard deviation

Figure 29. Sleep: graded exercise versus active intervention at end of intervention



Abbreviations: CI = confidence interval; IV = instrumental variable; SD = standard deviation

Exercise versus adaptive pacing

The PACE trial included a comparison of graded exercise versus adaptive pacing (N=319) in patients who met the Oxford case definition. Outcomes were reported at the end of therapy at 24 weeks and 28 weeks following the completion of therapy.[38] Exercise was associated with decreased fatigue severity versus adaptive pacing at the end of the intervention (mean difference -2.0, 95% CI -3.6 to -0.4 on the 11-item 0 to 33 Chalder scale, **Figure 12**) and at post-intervention follow-up (mean difference -2.5, 95% CI -4.2 to -0.84**, Figure 13**). Exercise was also associated with decreased depression severity (mean difference -1.1, 95% CI -2.0 to -0.15 on the 0 to 21 HADS depression scale), better sleep quality (mean difference -1.6, 95% CI -2.7 to -0.54 on the 0 to 20 Jenkins sleep scale), and longer 6-minute walk test (mean difference 45 meters, 95% CI 21 to 69) at post-intervention follow-up, with no difference in anxiety (mean difference -0.4, 95% CI -1.4 to 0.6 on the 0 to 21 HADS anxiety scale).

Exercise was associated with increased likelihood of improvement in fatigue (80% vs. 65%, RR 1.23, 95% CI 1.07 to 1.41, **Figure 21**), improvement in function (70% vs. 49%, RR 1.43, 95% CI 1.18 to 1.73, **Figure 22**), and recovery (28% vs. 16%, RR 1.71, 95% CI 1.10 to 2.65, **Figure 23**) versus adaptive pacing, based on the definitions used in the main PACE publication.[38] Results were similar using the original PACE protocol definitions for improvement in fatigue (24% vs. 14%, RR 1.64, 95% CI 1.03 to 2.62), improvement in function (61% vs. 40%, RR 1.51, 95% CI 1.20 to 1.89), and overall improvement (composite of improvement in fatigue and function, 21% vs. 9%, RR 2.19, 95% CI 1.24 to 3.86).[40] Exercise was also associated with lower likelihood of post-exertional malaise (44% vs. 63%, RR 0.71, 95% CI 0.57 to 0.87) and lower likelihood of worsening of function (RR 0.46, 95% CI 0.27 to 0.77, **Figure 26**).[62] There were no differences between exercise therapy versus adaptive pacing in risk of serious adverse events (**Figure 24)**, withdrawal due to adverse events (**Figure 25)**, but estimates were

imprecise (**Table 6**). In a post-hoc analysis, exercise was associated with decreased severity of muscle pain (mean difference -0.37, 95% CI -0.69 to -0.05) and joint pain (mean difference -0.36, 95% CI -0.69 to -0.05) at post-intervention follow-up (each assessed on a 0 to 4 scale).[64]

The PACE trial also evaluated longer-term, post-trial outcomes (median duration from randomization 31 months).[63] About 30% of patients in the exercise group and 50% in the adaptive pacing group received non-randomly allocated therapies between the end of the trial at 1 year and long-term follow-up. Fatigue severity based on the 11-item 0 to 33 Chalder scale was improved at long-term follow-up compared with end-of-trial scores in both the exercise (mean change -1.3, 95% CI -2.7 to -0.1) and adaptive pacing groups (mean change -3.0, 95% CI -4.4 to -1.6). For SF-36 physical function, there was no change at long-term follow-up compared with the end of the trial for exercise (mean change 0.5 point, 95% CI -2.7 to 3.6 on a 0 to 100 scale), but the severity of functional impairment improved in the adaptive pacing group (mean change 8.5 points, 95% CI 4.5 to 12.5). At long-term follow-up, mixed model analysis showed no differences between graded exercise versus adaptive pacing in fatigue (mean difference -1.1, 95% CI -3.0 to 0.9) or function (mean difference 5.6, 95% CI -0.3 to 11.5), though estimates favored graded exercise.

Exercise versus biofeedback

One small (n=24), high risk of bias trial compared graded exercise versus heart rate variability biofeedback therapy in patients who met the Fukuda case definition.[60] The duration of the intervention was 8 weeks and outcomes were assessed at the end of treatment and at 5 months. Results favored exercise therapy over biofeedback for fatigue at the end of the intervention (**Figure 12**) and at post-intervention follow-up (**Figure 13**), but differences were small and not statistically significant, and the estimates were imprecise (**Table 6**). Exercise was associated with greater depression severity at the end of treatment (mean difference 4.0, 95% CI 0.72 to 1.14 on the 0 to 27 Patient Health Questionnaire, 95% CI 0.72, **Figure 27**) and at post-intervention follow-up (mean difference 4.6, 95% CI 0.65 to 8.55, **Figure 16**) compared with biofeedback. The trial did not report harms.

Exercise versus fluoxetine

One trial (n=69) compared graded exercise versus fluoxetine (a selective serotonin reuptake inhibitor [SSRI]) in patients who met the Oxford case definition.[59] Outcomes were evaluated at the completion of 12 weeks of treatment and 14 weeks following the end of treatment. There were no differences between exercise versus fluoxetine in fatigue or depression at the end of the intervention (**Figures 12, 27**) or at post-intervention follow-up (**Figures 13, 16**). There was also no difference in the likelihood of recovery (defined as a score <4 on the 11-item 0 to 11 Chalder fatigue scale), but the estimate was imprecise (RR 3.09, 95% CI 0.67 to 14.25, **Figure 23**). The trial did not evaluate harms.

Exercise plus medication (fluoxetine) versus exercise or fluoxetine alone

One trial (N=102) compared graded exercise plus fluoxetine versus exercise or fluoxetine alone in patients who met the Oxford case definition.[59] Outcomes were evaluated at the completion of 12 weeks of treatment and 14 weeks following the end of treatment. The combination of exercise and fluoxetine was associated with less fatigue severity at the end of the intervention versus either exercise (mean difference -3.6, 95% CI -8.1 to 0.9 on the 14-item 0 to 42 Chalder fatigue scale) or fluoxetine (mean difference -4.1, 95% CI -8.6 to 0.4) alone, but the

differences were not statistically significant. At post-intervention follow-up, differences in fatigue severity between the combination versus either therapy alone were smaller and remained non-statistically significant. Differences in depression scores were small and not statistically significant at the end of the intervention and at post-intervention follow-up (mean differences on the 0 to 21 HADS depression scale ranged from -1.0 to 0.5 points). The trial did not evaluate harms.

Table 6. Exercise versus active interventions: summary of stratified results

Outcome	Number of studies (N)	Estimate (95% CI)	I^2
Fatigue, end of intervention			
Exercise vs. CBT	1 (298)	SMD 0.03 (-0.20 to 0.25)	--
Exercise vs. relaxation	1 (66)	SMD -0.83 (-1.34 to -0.33)	--
Exercise vs. adaptive pacing	1 (305)	SMD -0.29 (-0.51 to -0.06)	--
Exercise vs. biofeedback	1 (24)	SMD 0.47 (-0.35 to 1.29)	--
Exercise vs. fluoxetine	1 (67)	SMD 0.38 (-0.10 to 0.87)	--
Fatigue, post-intervention			
Exercise vs. CBT	2 (360)	SMD 0.08 (-0.13 to 0.29)	0%
Exercise vs. relaxation	2 (118)	SMD -0.16 (-0.71 to 0.38)	56%
Exercise vs. cognitive therapy	1 (57)	SMD -0.08 (-0.60 to 0.44)	--
Exercise vs. adaptive pacing	1 (307)	SMD -0.34 (-0.56 to -0.11)	--
Exercise vs. biofeedback	1 (24)	SMD 0.62 (-0.20 to 1.45)	--
Exercise vs. fluoxetine	1 (69)	SMD -0.27 (-0.74 to 0.21)	--
SF-36 physical function subscale or physical component score (0 to 100), end of intervention			
Exercise vs. CBT	1 (298)	MD 1.20 (-3.90 to 6.30)	--
Exercise vs. relaxation	1 (66)	MD 14.00 (4.24 to 23.76)	--
Exercise vs. adaptive pacing	1 (305)	MD 12.20 (7.17 to 17.23)	--
Exercise vs. biofeedback	1 (24)	MD -0.40 (-8.26 to 7.46)	--
SF-36 physical function subscale or physical component score (0 to 100), post-intervention			
Exercise vs. CBT	2 (360)	MD -8.36 (-26.21t o 9.50)	80%
Exercise vs. relaxation	1 (57)	MD -21.48 (-35.85 to -7.11)	--
Exercise vs. cognitive therapy	1 (57)	MD -21.37 (-34.73 to -8.01)	--
Exercise vs. adaptive pacing	1 (207)	MD 11.80 (6.05 to 17.55)	--
Exercise vs. biofeedback	1 (24)	MD -0.50 (-8.35 to 7.35)	--
Depression, end of intervention			
Exercise vs. relaxation	1 (66)	SMD 0.33 (-0.15 to 0.82)	--
Exercise vs. biofeedback	1 (24)	SMD 1.01 (0.15 to 1.88)	--
Exercise vs. fluoxetine	1 (69)	SMD 0.44 (-0.04 to 0.91)	--
Depression, post-intervention			
Exercise vs. CBT	2 (345)	SMD 0.02 (-0.19 to 0.23)	0%
Exercise vs. relaxation	2 (118)	SMD -0.12 (-0.95 to 0.72)	81%
Exercise vs. cognitive therapy	1 (57)	SMD 0.51 (-0.02 to 1.04)	--
Exercise vs. adaptive pacing	1 (293)	SMD -0.25 (-0.48 to -0.02)	--
Exercise vs. biofeedback	1 (24)	SMD 0.96 (0.10 to 1.81)	--
Exercise vs. fluoxetine	1 (80)	SMD 0.13 (-0.32 to 0.57)	--
HADS anxiety (0 to 21), end of intervention			
Exercise vs. relaxation	1 (66)	MD -1.50 (-3.68 to 0.68)	--
Anxiety, post intervention			
Exercise vs. CBT	2 (345)	SMD 0.07 (-0.14 to 0.28)	0%
Exercise vs. relaxation	2 (118)	SMD -0.25 (-0.88 to 0.37)	66%
Exercise vs. cognitive therapy	1 (57)	SMD 0.36 (-0.16 to 0.88)	--
Exercise vs. adaptive pacing	1 (293)	SMD -0.09 (-0.32 to 0.14)	--
Pittsburgh Sleep Quality Index (0 to 21), end of intervention			
Exercise vs. relaxation	1 (66)	MD -1.00 (-2.21 to 0.212)	--
Sleep, post intervention			
Exercise vs. CBT	2 (345)	SMD -0.17 (-0.39 to 0.04)	0%

Outcome	Number of studies (N)	Estimate (95% CI)	I^2
Exercise vs. relaxation	1 (57)	SMD 0.14 (-0.38 to 0.66)	--
Exercise vs. cognitive therapy	1 (57)	SMD -0.04 (-0.56 to 0.48)	--
Exercise vs. adaptive pacing	1 (294)	SMD -0.33 (-0.56 to -0.10)	--
Brief Pain Inventory (0 to 10), post intervention			
Exercise vs. CBT	1 (58)	MD -0.35 (-2.02 to 1.32)	--
Exercise vs. relaxation	1 (57)	MD -0.69 (-2.23 to 0.85)	--
Exercise vs. cognitive therapy	1 (57)	MD 0.39 (-1.14 to 1.92)	--
6-minute walk test (meters), end of intervention			
Exercise vs. CBT	2 (291)	MD -4.23 (-75.99 to 67.52)	71%
Exercise vs. relaxation	1 (57)	MD -15.53 (-55.32 to 24.26)	--
Exercise vs. cognitive therapy	1 (57)	MD -41.21 (-79.61 to -2.80)	--
Exercise vs. adaptive pacing	1 (221)	MD 45.00 (16.31 to 73.69)	--
Recovery			
Exercise vs. CBT	2 (360)	RR 0.91 (0.65 to 1.29)	0%
Exercise vs. relaxation	1 (57)	RR 4.83 (0.24 to 96.42)	--
Exercise vs. cognitive therapy	1 (57)	RR 0.48 (0.10 to 2.43)	--
Exercise vs. adaptive pacing	1 (307)	RR 1.71 (1.10 to 2.65)	--
Exercise vs. fluoxetine	1 (69)	RR 3.09 (0.67 to 14.25)	--
Fatigue improvement			
Exercise vs. CBT	1 (303)	RR 1.05 (0.93 to 1.19)	--
Exercise vs. adaptive pacing	1 (306)	RR 1.23 (1.07 to 1.41)	--
Functional improvement			
Exercise vs. CBT	2 (360)	RR 0.98 (0.85 to 1.14)	0%
Exercise vs. relaxation	1 (57)	RR 0.48 (0.13 to 1.74)	--
Exercise vs. cognitive therapy	1 (57)	RR 0.32 (0.10 to 1.07)	--
Exercise vs. adaptive pacing	1 (307)	RR 1.43 (1.18 to 1.73)	--
Serious adverse events			
Exercise vs. CBT	1 (321)	RR 1.87 (0.77 to 4.56)	--
Exercise vs. adaptive pacing	1 (319)	RR 0.86 (0.42 to 1.75)	--
Withdrawal due to worsening			
Exercise vs. CBT	1 (321)	RR 5.03 (0.24 to 103.97)	--
Exercise vs. adaptive pacing	1 (319)	RR 0.66 (0.11 to 3.91)	--
Physical function worsening			
Exercise vs. CBT	1 (321)	RR 1.21 (0.0.63 to 2.31)	--
Exercise vs. adaptive pacing	1 (319)	RR 0.46 (0.27 to 0.77)	--
Post-exertional malaise			
Exercise vs. CBT	1 (321)	RR 0.90 (0.72 to 1.14)	--
Exercise vs. adaptive pacing	1 (319)	RR 0.71 (0.57 to 0.87)	--

Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; HADS = Hospital Anxiety and Depression Scale; MD = mean difference; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference

Other Exercise Therapies

Orthostatic training versus sham training

One medium risk of bias trial (N=25, **Evidence Table Appendix E2**) included in the prior AHRQ report compared home orthostatic training versus sham treatment in patients who met the Fukuda case definition.[67] Home orthostatic training consisted of standing with the upper back against the wall and heels approximately 15 cm from the base of the wall once daily for 30 minutes. Sham training consisted of the same position for 10 minutes and gentle calf flexion and extension exercises. The mean Fatigue Impact Severity score at baseline was 95.3 (0 to 160 scale). At the end of 6 months of therapy, there was no difference between orthostatic training versus no orthostatic training in fatigue severity (mean difference -8.00, 95% CI -33.5 to 1.75 on the Fatigue Impact scale; estimate includes data from two non-adherent patients). Orthostatic

training was associated with reduced drop in blood pressure with standing (mean difference 6 mm Hg, 95% CI 00 to 12.6). The trial did not report orthostatic symptoms or harms.

Cognitive behavioral therapy

Twelve trials evaluated CBT in adult patients with ME/CFS (**Tables 7 and 8**).[38,54,68-77] Sample sizes ranged from 58 to 630 (total N=1888). Nine trials compared CBT versus usual care, usual specialist care, an attention control (education and support), or wait list (advice or supportive listening);[38,70-77] four trials compared CBT versus an active intervention (exercise, adaptive pacing, relaxation, cognitive therapy, or fluoxetine);[38,54,69,75] and one trial compared different CBT modes of delivery (face-to-face or telephone).[68] Nine trials were included in the prior AHRQ report[38,54,68,69,71-74,76] and three trials were added for this update.[70,75,77] All three new trials compared CBT versus inactive controls; one[75] also compared CBT versus active therapy (mirtazapine).

Two trials were conducted in the United States and 10 trials in Europe. The mean age of participants ranged from 35 to 46 years and the proportion female ranged from 60% to 88%. The case definition for ME/CFS was the Oxford criteria in two trials, the Fukuda criteria in seven trials (including one trial[70] that used Dutch criteria in accordance with Fukuda), or both in three trials. The duration of ME/CFS ranged from 37 to 104 months in eight trials that reported this information. Baseline fatigue was measured using a variety of scales (**Table 7**). One trial[70] reported that 90% of patients had post-exertional fatigue at baseline and one trial[76] reported that 38% of patients were "low active" and 62% "relative active" at baseline. Otherwise, details regarding the presence of post-exertional fatigue and activity patterns were lacking. One trial[68] excluded patients with melancholic depression, the proportion of patients with depression or treated for depression ranged from 10% to 39% in six trials,[38,54,69,70,73,74] and five trials[71,72,75-77] did not report the proportion of patients with depression. Two trials excluded patients with major depression. In the other trials, the proportion of patients with depression or an axis I psychiatric diagnosis ranged from 10% to 39%. Depression severity was most commonly assessed with the HADS depression score (0 to 21 scale, higher scores indicate more severe depression). In three trials, mean HADS depression scores at baseline ranged from 8.2 to 9.1. Functional impairment was most commonly reported using the SF-36 physical function subscale (0 to 100 scale, lower scores indicate more functional impairment). In 10 trials, mean SF-36 physical function subscale scores ranged from 26.6 to 62.5.

CBT was administered individually in 10 trials; of these, the mode of delivery was face-to-face in seven trials and another mode of delivery (web-based, telephone, or self-guided) was used in three trials. CBT was administered in group, face-to-face sessions in two trials. The duration of the CBT intervention ranged from 12 weeks to 6 months. In most trials, the frequency of CBT was weekly or biweekly. The session length varied, with details not reported in some trials (**Table 7**). Outcomes were assessed at 16 weeks to 18 months; nine trials[38,68-72,75-77] evaluated patients at the end of the intervention and five trials[38,54,68,73,74] evaluated patients 29 weeks to 12 months following the completion of therapy.

Eleven trials were rated medium risk of bias and one trial[72] was rated high risk of bias (**Risk of Bias Table Appendix F**). In all trials, blinding of patients and care providers to CBT was not feasible. Other methodological limitations included high attrition, failure to report attrition, inadequate description of randomization or allocation concealment methods, and failure to blind or unclear blinding status of outcomes assessors and data analysts.

Table 7. Cognitive behavioral therapy RCTs: study characteristics

Author, year Country Risk of Bias	Study N (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
Burgess, 2012[68] United Kingdom Medium	n: 58 Age: 37.4 % Female: 79	Criteria: Fukuda and Oxford Duration: 3.97 years	Fatigue Scale: Chalder (11-item, 0 to 11) Baseline: 10.2 (SD 1.9) Post-exertional fatigue or malaise: not reported	Major depression: Excluded for melancholic depression Baseline depression: HADS depression (0 to 21): 9.1 (SD 1.7) Baseline function: SF-36 physical function (0 to 100): 51.4 (SD 1.7)	A: Face-to-face CBT B: Telephone CBT Frequency: Face-to-face: 2 initial sessions and 15 follow-up sessions over 6 months Telephone: 1 initial session and 13 follow-up sessions over 6 months Session length: Face-to-face: 1.5 hours for initial sessions and 50 minutes to 1-hour for follow-up sessions Telephone: 3 hours for initial session and 30-minute follow-up sessions Duration of treatment: 6 months Duration of follow-up: 18 months
Deale, 1997[69] United Kingdom Medium	n: 60 Age: 34.5 % Female: 68	Criteria: Oxford and Fukuda Duration: 4 years	Fatigue Scale: Chalder (11-item, 0 to 11) Baseline: 9.8 (SD 2.1) Post-exertional fatigue or malaise: Not reported	Major depression: 15% Baseline depression: Beck Depression Inventory (0 to 63): 14.1 (SD 6.7) Baseline function: SF-36 physical function (0 to 100): 26.6 (SD 23.4)	A: CBT B: Relaxation (10 sessions with twice daily practice) Frequency: 13 sessions weekly or biweekly over 4 to 6 months Session length: Not described (mean total time 15 hours) Duration of treatment: 4 to 6 months Duration of follow-up: 10 to 12 months
Janse, 2018[70] The Netherlands Medium	n: 240 Age: 37.6 % Female: 60	Criteria: National Dutch guidelines (in accordance with Fukuda) Duration: 4 to 6.5 years	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 50.0 (SD 5.2) Post-exertional fatigue or malaise: 90%	Major depression: Any depressive disorder: 10% Baseline depression: SCL-90 (90 to 450): 156.5 (SD 35.3) Baseline function: SF-36 physical function (0 to 100): 62.5 (SD 19.4)	A: Web-based CBT, protocol driven feedback B: Web-based CBT, feedback on demand C: Wait list Frequency: At least biweekly (protocol driven) or individualized (feedback on demand) Session length: Diagnostic sessions 2 hours, otherwise not specified Duration of treatment: 6 months Duration of follow-up: 6 months

Author, year Country Risk of Bias	Study N (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
Jason, 2007[54] United States Medium	n: 114 Age: 43.8 % Female: 83	Criteria: Fukuda Duration: Not reported	Fatigue Scale: Fatigue Severity Scale 9-item (1 to 7) Baseline: 6.1 (SD 0.71) Post-exertional fatigue or malaise: not reported	Major depression: Current axis I diagnosis: 39% Baseline depression: Beck Depression Inventory (0 to 63), mean: 18.7 (SD 9.9) Baseline function: SF-36 physical function (0 to 100): 46.2 (SD 23.8)	A: CBT B: Cognitive therapy C: Relaxation (RELAX) D: Anaerobic exercise (Anaerobic Activity Therapy [ACT]/progressive rehabilitation) Frequency: Biweekly Session length: 45 minutes Duration of treatment: 6 months Duration of follow-up: 1 year
Knoop, 2008[71] The Netherlands Medium	n: 159 Age: 38.0 % Female: 79	Criteria: Fukuda Duration: Median 72 vs. 96 months	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 49.5 (SD 5.4) Post-exertional fatigue or malaise: Not reported	Major depression: Not reported Baseline depression: Not reported Baseline function: SF-36 physical function (0 to 100): 53.2 (SD 20.7)	A: Self-guided CBT B: Wait list Frequency: Not specified for self-guided sessions. Email contact with therapist at least every 2 weeks for at least 16 weeks Session length: Not described Duration of treatment: At least 16 weeks Duration of follow-up: At least 16 weeks
Lopez, 2011[72] United States High	n: 58 Age: 45.9 % Female: 88	Criteria: Fukuda Duration: Not reported	Fatigue Scale: Profile of Mood States fatigue/inertia subscale (0 to 28) Baseline: 17.8 (SD 6.3) Post-exertional fatigue or malaise: Not reported	Major depression: Not reported Baseline depression: POMS total mood disturbance (0 to 200): 36.3 (SD 30.8) Baseline function: Not reported	A: Group cognitive behavioral stress management B: Psycho-educational seminar (half day) Frequency: Intervention: Weekly for 12 weeks Control: single session Session length: Intervention: 2 hours (20 to 30 minutes relaxation, 90 minutes didactic and discussion) Control: half-day Duration of treatment: 12 weeks Duration of follow-up: 12 weeks

Author, year Country Risk of Bias	Study N (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
O'Dowd, 2006[73] United Kingdom Medium	n: 122 Age: 41.1 % Female: 67	Criteria: Fukuda Duration: Mean/median not reported (48% >60 months)	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 24.6 (SD 6.4) Post-exertional fatigue or malaise: Not reported	Major depression: Treated for depression: 16% Baseline depression: HADS depression (0 to 21): 8.7 (SD 3.5) Baseline function: SF-36 physical function (0 to 100): 32.2 (SD 7.8)	A: CBT B: Education and support (8 sessions biweekly, 2 hours each) C: Usual care Frequency: 8 sessions biweekly Session length: 2 hours Duration of treatment: 16 weeks Duration of follow-up: 12 months
Sharpe, 1996[74] United Kingdom Medium	n: 60 Age: 36.0 % Female: 68	Criteria: Oxford Duration: Mean 31.6 months	Fatigue Scale: 0 to 10 Likert scale Baseline: 7.8 (SD 1.7) Post-exertional fatigue or malaise: Not reported	Major depression: Excluded for severe depression or bipolar disorder Major depression: 20% Baseline depression: HADS depression (0 to 21): 6.8 (SD 3.6) Baseline function: KPS (0 to 100): 71.5 (SD 3.4)	A: CBT B: Usual medical care Frequency: 16 sessions over 4 months Session length: 1 hour Duration of treatment: 4 months Duration of follow-up: 12 months
Stubhaug, 2008[75] Norway Medium	n: 72 Age: 46.3 % Female: 82	Criteria: Oxford (90%) or Fukuda (40%) Duration: not reported	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 25.0 (SD 4.5) Post-exertional fatigue or malaise: not reported	Major depression: Not reported Baseline depression: Hamilton Rating Scale for Depression (0 to 52): 14.5 (SD 3.9) Baseline function: SF-36 physical function (0 to 100): 28.9 (SD 11.3)	A: CBT B: Mirtazapine 15 to 45 mg daily C: Placebo Frequency: 2 sessions per week for 12 weeks Session length: 1.5 hours Duration of treatment: 12 weeks (initial therapy) Duration of follow-up: 24 weeks (after crossover)
Tummers, 2012[76] The Netherlands Medium	n: 111 Age: 36.4 % Female: 78	Criteria: Fukuda Duration: 48 vs. 60 months	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 51.3 (SD 5.4) Post-exertional fatigue or malaise: Not reported Low active: 38% Relative active: 62%	Major depression: Proportion with depression not reported Baseline depression: Brief Symptom Inventory psychological distress (0 to 4): 1.02 (SD 0.63) Baseline function: SF-36 physical function (0 to 100): 50.8 (SD 22.3)	A: Self-guided CBT B: Wait list Frequency: Not specified for self-guided sessions. Email contact with therapist at least every 2 weeks for at least 20 weeks Session length: Not described Duration of treatment: At least 20 weeks Duration of follow-up: At least 20 weeks

Author, year Country Risk of Bias	Study N (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
White, 2011[38] United Kingdom Medium	n: 630 Age: 38.0 % Female: 77	Criteria: Oxford Duration: Median 32 months	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 28.2 (SD 3.8) Post-exertional fatigue or malaise: Not reported	Major depression: Any depressive disorder: 34% Baseline depression: HADS depression (0 to 21), mean: 8.2 (SD 3.8) Baseline function: SF-36 physical function (0 to 100): 38.0 (SD 15.8)	A: Adaptive pacing therapy + specialist medical care B: CBT + specialist medical care C: Graded exercise + specialist medical care D: Specialist medical care Frequency: Weekly for 4 weeks, then biweekly, plus one booster at 36 weeks Session length: Not described Duration of treatment: 23 weeks (booster at 36 weeks) Duration of follow-up: 12 months
Wiborg, 2015[77] The Netherlands Medium	n: 204 Age: 37.7 % Female: 77	Criteria: Fukuda Duration: Mean 8.7 years	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 50.6 (SD 4.7) Post-exertional fatigue or malaise: Not reported	Major depression: Proportion with depression not reported Baseline depression: SCL-90 (90 to 450): 163.7 (SD 37.6) Baseline function: SF-36 physical function (0 to 100): 56.9 (SD 19.2)	A: Group CBT, 8 patients per group B: Group CBT 4 patients per group C: Wait list Frequency: 14 session over 6 months Session length: 2 hours Duration of treatment: 6 months Duration of follow-up: 6 months

Abbreviations: ACT = anaerobic activity therapy; CBT = cognitive behavioral therapy; CFS = chronic fatigue syndrome; HADS = Hospital Anxiety and Depression Scale; ME = myalgic encephalomyelitis; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey

Table 8. Cognitive behavioral therapy RCTs: study results

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)*	Depression Outcomes*	Function Outcomes*
Burgess, 2012[68] Fukuda and Oxford	A: Face-to-face CBT (35) B: Telephone CBT (45) Duration of treatment: 6 months Duration of follow-up: 18 months	Fatigue: Chalder Fatigue Scale 11-item (0-11, score of ≥4 is cutoff for caseness), mean (SD); all p values are NS 3 months: 7.08 (3.97) vs. 7.08 (3.56) 6 months: 5.75 (4.49) vs. 7.75 (3.77) 12 months: 6.83 (4.57) vs. 7.89 (3.75)	Not reported	Overall Function: MOS-SF physical function (0 to 100), mean (SD) 3 months: 58.97 (19.38) vs. 62.89 (20.33) 6 months: 65.78 (23.61) vs. 62.96 (20.36) 12 months: 62.32 (24.96) vs. 65.83 (21.73); p=0.043 for change from baseline for both groups, all other p-values NS
Deale, 1997[69] Oxford and Fukuda	A: CBT (30) B: Relaxation (10 sessions with twice daily practice) (30) Duration of treatment: 4 to 6 months Duration of follow-up: 10 to 12 months	Fatigue: Fatigue Problem Rating (0-8 scale), mean (SD): Posttreatment: 4.1 (1.9) vs. 5.5 (1.4) 6-month follow-up: 3.4 (2.2) vs. 5.5 (1.9), p<0.001 for between group differences over time Chalder Fatigue Scale 11-item (0 to 11), score of ≥4 is cutoff for caseness, mean (SD) Posttreatment: 7.2 (4.0) vs. 7.5 (4.1) 6-month follow-up: 4.1 (4.0) vs. 7.2 (4.0) p<0.001 for between group differences over time % With fatigue rating by assessor at 3-month follow-up Better or much better: 72 (18/25) vs. 17 (4/23); p<0.001 Unchanged or worse: 28 (7/25) vs. 83 (19/23) % With score <4 on Chalder Fatigue Scale 6-month follow-up: 63 (17/27) vs. 15 (4/26); p=0.001 5-year follow-up: 28 (7/25) vs. 25 (7/28); p=1.00	Beck Depression Inventory, mean (SD) Posttreatment: 8.9 (5.6) vs. 11.9 (7.4) 6-month follow-up: 10.1 (6.9) vs. 12.3 (8.5), p>0.30	Overall Function: SF-36 physical function (0 to100), mean (SD) Posttreatment: 56.2 (26.2) vs. 34.6 (28.3) 6-month follow-up: 71.6 (28.0) vs. 38.4 (26.9); p<0.03 % With good outcome on SF-36 physical function(increase of ≥50 from baseline to 6 months, or end score of ≥83): 6 months follow-up: 63 (19/30) vs. 17 (5/30); difference of 46 (95% CI 24 to 68) p<0.001 5-year follow-up: 48 (12/25) vs. 32 (9/28); p=0.27 % With rating by assessor at 3-month follow-up Better or much better: 80 (20/25) vs. 26 (6/23); p<0.001 Unchanged or worse: 20 (5/25) vs. 74 (17/23)

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)*	Depression Outcomes*	Function Outcomes*
Janse, 2018[70] National Dutch guidelines (in accordance with Fukuda)	A: 1: Web-based CBT, protocol driven feedback (80) B: Web-based CBT, feedback on demand (80) C: Wait list (80) Duration of treatment: 6 months Duration of follow-up: 6 months	Fatigue: Checklist Individual Strength, fatigue severity subscale (8 to 56), mean (SD): 36.3 (14.6) vs. 37.0 (13.1) vs. 43.9 (10.5) Mean difference compared with control (97.5% CI): iCBT with protocol feedback: -8.3 (-12.7 to -3.9), p<0.0001; iCBT with feedback on demand: -7.2 (-11.3 to -3.1), p<0.0001	Not reported	Overall Function: SF-36 physical function (0 to 100), mean (SD): 73.3 (25.9) vs. 77.0 (21.3) vs. 70.8 (21.0) Difference compared with control: iCBT with protocol feedback: 2.4 (-3.6 to 8.4), p=0.44; iCBT with feedback on demand: 5.8 (0.6 to 11.0), p=0.030
Jason, 2007[54] Fukuda	A: CBT (29) B: Cognitive therapy (28) C: 1: Relaxation (RELAX) (28) D: Anaerobic exercise (Anaerobic Activity Therapy [ACT]/progressive rehabilitation) (29) Duration of treatment: 6 months Duration of follow-up: 1 year	Fatigue Severity Scale 9-item (1 to 7), mean (SD) 12 months: 5.77 (1.43) vs. 5.62 (1.06) vs. 5.37 (1.19) vs. 5.87 (1.01); p=NR	Beck Depression Inventory (0 to 21), mean (SD) 12 months: 13.95 (13.08) vs. 11.86 (7.36) vs. 16.94 (11.82) vs. 13.50 (9.97), p<0.001	Overall Function: SF-36 physical function0 to100), mean (SD): 12 months: 58.64 (30.44) vs. 61.09 (23.74) vs. 61.20 (27.70) vs. 39.72 (27.63) p<0.01 for CBT and COG over time vs. ACT over time % Achieving clinically significant improvement: 18.2 vs. 30.4 vs. 21.7 vs. 11.1; p=0.49
Knoop, 2008[71] Fukuda	A: Self-guided CBT (85) B: Wait list (86) Duration of treatment: At least 16 weeks Duration of follow-up: At least 16 weeks	Fatigue: Checklist Individual Strength, fatigue severity subscale (8 to 56) mean (SD): Second assessment: 38.9 (12.1) vs. 46.4 (8.7); p<0.001 % With reduction in Checklist Individual Strength, fatigue severity subscale (<35 score and reliable change index of >1.96) 27 (23/84; 95% CI, 18 to 37) vs. 7 (6/85; 95% CI, 2 to 13); OR 4.9 (95% CI 1.9 to 12.9); p=0.001	Not reported	Overall Function: SF-36 physical function (0 to100), mean (SD) Second assessment: 65.9 (23.2) vs. 60.2 (23.7); p=0.011 Mean (SD) functional impairment Sickness Impact Profile (SIP-8) (0 to 5,799) Second assessment: 1,079 (690) vs. 1,319 (619); p<0.001
Lopez, 2011[72] Fukuda	A: Group cognitive behavioral stress management (44) B: Psycho-educational (half day) (25) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	Fatigue: POMS fatigue subscale (0 to 28), mean (SD) After treatment: 17.85 (7.34) vs. 20.09 (6.99); p=0.06	Not reported	Overall Function: Not reported

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)*	Depression Outcomes*	Function Outcomes*
O'Dowd, 2006[73] Fukuda	A: CBT (52) B: Education and support (8 sessions every other week, 2 hours each) (50) C: Usual care (51) Duration of treatment: 16 weeks Duration of follow-up: 12 months	Chalder Fatigue Scale 11-item (0 to 33), mean (SD) 6 months: 17.9 (8.41) vs. 21.4 (7.55) vs. 21.8 (6.90); p=0.19 12 months: 17.4 (7.32) vs. 21.4 (7.79) vs. 18.8 (7.19); p=0.19 Difference between groups from baseline at 6 and 12 months pooled CBT vs. group support: -3.16 (95% CI -5.59 to -0.74); p=0.011 CBT vs. usual care: -2.61 (95% CI -4.92 to -0.30); p=0.027 Support vs. usual care: 0.55 (95% CI -1.56 to 2.66); p=NR	HADS-Depression 6 months: 6.84 (3.46) vs. 8.20 (3.81) vs. 7.78 (3.76) 12 months: 6.82 (3.80) vs. 7.74 (4.02) vs. 7.44 (4.42) Mean difference, adjusted for baseline: -0.13 (-1.13 to 0.87) vs. -0.56 (-1.69 to 0.58) vs. -0.43 (-1.56 to 0.70), p=0.52	SF-36 physical function (0 to 100), mean (SD); all p values are NS 6 months: 33.4 (9.04) vs. 32.3 (9.30) vs. 34.5 (9.95) 12 months: 35.2 (8.15) vs. 32.5 (7.91) vs. 35.0 (9.93)
Sharpe, 1996[74] Oxford	A: CBT (30) B: Usual medical care (30) Duration of treatment: 4 months Duration of follow-up: 12 months	Fatigue severity (0 to 10), mean: 12 months: 4.3 vs. 6.3 Change from baseline, -3.5 vs. -1.6; difference 1.9, 95% CI 0.5 to 3.3	HADS-Depression, mean 12 months: 3.6 vs. 5.8 Change from baseline: -3.1 vs. -1.0; difference 2.0, 95% CI 0.0 to 4.1	Achieved KPS score of ≥80 5 months: 27% (8/30) vs. 20% (6/30); difference of 7 (95% CI, -15 to 28) 8 months: 53% (16/30) vs. 30% (9/30); difference of 23 (95% CI, 0 to 48) 12 months: 73% (22/30) vs. 27% (8/30); difference of 47 (95% CI, 24 to 69); p<0.001 Improvement of ≥10 points on KPS 5 months: 23% (7/30) vs. 7% (2/30); difference of 17 (95% CI, 0 to 34) 8 months: 60% (18/30) vs. 20% (6/30); difference of 40 (95% CI, 17 to 63) 12 months: 73% (22/30) vs. 23% (7/30); difference of 50 (95% CI, 28 to 72); p<0.001
Stubhaug, 2008[75] Oxford (90%) or Fukuda (40%)	A: CBT (23) B: Mirtazapine 15 to 45 mg daily (28) C: Placebo (24) Duration of treatment: 12 weeks (initial therapy) Duration of follow-up: 24 weeks (after crossover)	Chalder Fatigue Scale 11-item (0 to 33) score at 12 weeks: 23.7 (21.0 to 26.5) vs. 22.7 (21.4 to 24.1) vs. 23.7 (21.0 to 26.5), p=0.014 Chalder Fatigue Scale 11-item (0 to 33) score at 24 weeks: 23.3 (20.1 to 26.5) vs. 23.7 (22.4 to 25.0) vs. 24.2 (21.4 to 27.1) vs. 18.7 (15.4 to 22.0); p<0.001	Hamilton Rating Scale for Depression, mean (95% CI) 12 weeks: 12.9 (10.1 to 15.7) vs. 12.6 (11.4 to 13.8) vs. 13.5 (10.9 to 16.1)	Not reported

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)*	Depression Outcomes*	Function Outcomes*
Tummers, 2012[76] Fukuda	A: Self-guided CBT (62) B: Wait list (61) Duration of treatment: At least 20 weeks Duration of follow-up: At least 20 weeks	Checklist Individual Strength, fatigue severity subscale (8 to56 scale), mean (SD) Second assessment: 39.6 (14.1) vs. 48.3 (8.1); p<0.01 % With reduction in Checklist Individual Strength, fatigue severity subscale (score <35 and reliable change index of >1.96) 33 (18/55) vs. 9 (5/56); OR 5.0 (95% CI 1.69 to 14.57)	Not reported	SF-36 physical function (0 to 100), mean (SD) Second assessment: 65.4 (24.9) vs. 59.3 (22.9); p=0.08 Subanalysis of baseline group with SF-36 physical function score ≤70 Self-instruction (n=53) vs. wait list (n=50) Mean (SD) SF-36 physical function (0 to 100) Second assessment: 63.0 (25.9) vs. 53.4 (18.7) Change from baseline: 18.5 vs. 9.6, difference: 9.05 (95% CI, 0.2 to 17.9); p<0.05
White, 2011[38] Oxford	A: Adaptive pacing therapy + specialist medical care (160) B: CBT + specialist medical care (161) C: Graded exercise + specialist medical care (160) D: Specialist medical care (160) Duration of treatment: 23 weeks (booster at 36 weeks) Duration of follow-up: 12 months	Fatigue: Chalder Fatigue Scale 11-item (0 to 33), mean (SD) 12 weeks: 24.2 (6.4) vs. 23.6 (6.5) vs. 22.8 (7.5) vs. 24.3 (6.5) 24 weeks: 23.7 (6.9) vs. 21.5 (7.8) vs. 21.7 (7.1) vs. 24.0 (6.9) 52 weeks: 23.1 (7.3) vs. 20.3 (8.0) vs. 20.6 (7.5) vs. 23.8 (6.6) Mean difference (95% CI) at 52 weeks: CBT vs. control: -3.4 (-5.0 to -1.8) p=0.0001 CBT vs. APT: -2.7 (-4.4 to -1.1) p=0.0027 % Improved from baseline (by ≥2 points): 65 (99/153) vs. 76 (113/148) vs. 80 (123/154) vs. 65 (98/152) % Within normal range (score ≤18): 22 (34/153) vs. 41 (60/148) vs. 33 (51/154) vs. 21 (32/152)	HADS-Depression, mean (SD) 52 weeks: 7.2 (4.5) vs. 6.2 (3.7) vs. 6.1 (4.1) vs. 7.2 (4.7); CBT vs. control: p=0.0003 CBT vs. APT: p=0.382	SF-36 physical function (0 to 100), mean (SD): 12 weeks: 41.7 (19.9) vs. 51.0 (20.7) vs. 48.1 (21.6) vs. 46.6 (20.4) 24 weeks: 43.2 (21.4) vs. 54.2 (21.6) vs. 55.4 (23.3) vs. 48.4 (23.1) 52 weeks: 45.9 (24.9) vs. 58.2 (24.1) vs. 57.7 (26.5) vs. 50.8 (24.7) Mean difference at 52 weeks: CBT vs. control: 7.1 (2.0 to 12.1) p=0.0068 CBT vs. APT: 10.5 (5.4 to 15.6) p=0.0002 % Improved from baseline (by ≥8 points): 49 (75/153) vs. 71 (105/148) vs. 70 (108/154) vs. 58 (88/152) % Within normal range (score ≥60): 35 (53/153) vs. 52 (77/148) vs. 53 (81/154) vs. 41 (62/152)

Author, year ME/CFS criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)*	Depression Outcomes*	Function Outcomes*
Wiborg, 2015[77] Fukuda	A: 1. Group CBT, 8 patients per group B. Group CBT, 4 patients per group C: Wait list Duration of treatment: 6 months Duration of follow-up: 6 months	A+B vs. C Fatigue severity, mean (SD): 33.5 (13.6) vs. 46.6 (8.5), treatment effect -13.8 (95% CI, -17.2 to -10.3), p<0.001 Improvement in fatigue severity: 49.3% (67/139) vs. 8.8% (6/68), OR 10.0 (95 CI, 4.1 to 24.8), p<0.001 Normal functioning in fatigue severity: 32.4% (44/136) vs. 2.9% (2/68), OR 15.8 (95% CI, 3.7 to 67.4), p<0.001	Not reported	A+B vs. C SF-36 physical function (0 to 100), mean (SD): 747.7 (22.0) vs. 63.3 (21.1), treatment effect 14.1 (95% CI, 9.0 to 19.3), p<0.001

*A vs. B vs. C vs. D, unless otherwise noted

Abbreviations: ACT = anaerobic activity therapy; APT = adaptive pacing therapy; CBT = cognitive behavioral therapy; CI = confidence interval; CFS = chronic fatigue syndrome; COG = cognitive therapy; GET = graded exercise therapy; iCBT = internet-based cognitive-behavioral therapy; KPS = Karnofsky Performance Scale; ME = myalgic encephalomyelitis; NR = not reported; NS = not significant; OR = odds ratio; POMS = profile of mood states; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey; SIP-8 = Sickness Impact Profile 8-item

Cognitive Behavioral Therapy Versus Inactive Controls

Nine trials (N=1,656) compared CBT versus usual care (2 trials),[73,74] usual specialist care (1 trial),[38] wait list (4 trials),[70,71,76,77] an attention control (education, 2 trials),[72,73] or placebo medication (1 trial)[75] (**Tables 7 and 8, Evidence Table Appendix E2**). Three of these trials were added for this review; two of the trials[70,77] used the Fukuda case definition and one[75] used Fukuda or Oxford. The duration of the CBT intervention ranged from 12 weeks to 6 months. Seven trials evaluated patients at the end of the intervention and three trials[38,73,74] evaluated patients 6.7 to 8.4 weeks following the end of the intervention. One trial[72] was rated high risk of bias and the others were rated medium risk of bias. Results stratified by the inactive comparator are summarized in **Table 9** and shown in **Figures 30 to 40**.

Fatigue

CBT was associated with decreased fatigue severity versus wait list, usual specialist care, or an attention control at the end of treatment (7 trials, N=1,129, SMD -0.61, 95% CI -0.83 to -0.40, I^2=64%;[38,70-72,75-77] **Figure 30**). Mean differences were -2.53 (95% CI -4.06 to -1.00, I^2=0%) in two trials (N=347) that used the 11-item 0 to 33 Chalder scale,[38,75] -9.20 (95% CI -12.09 to -6.31, I^2=66%) in four trials (N=724) that used the 8 to 56 Checklist Individual Strength fatigue severity scale,[70,71,76,77] and -2.24 (95% CI -6.09 to 1.61, I^2=98%) in one trial (N=422) that used the 0 to 28 Profile of Mood States (POMS) fatigue/inertia scale. Estimates consistently favored CBT when trials were stratified by the control type, ME/CFS case definition, and CBT type (**Table 9**). A statistically significant subgroup effect was present for control type (p for subgroup difference=0.05), with the strongest estimate for trials of CBT versus wait list (4 trials, N=724, SMD -0.77, 95% CI -0.99 to -0.55, I^2=50%) compared with trials of CBT versus placebo medication, usual specialist care, or an attention control (SMDs ranged from -0.31 to -0.40). For ME/CFS case definition, there was a subgroup effect of borderline statistical significance

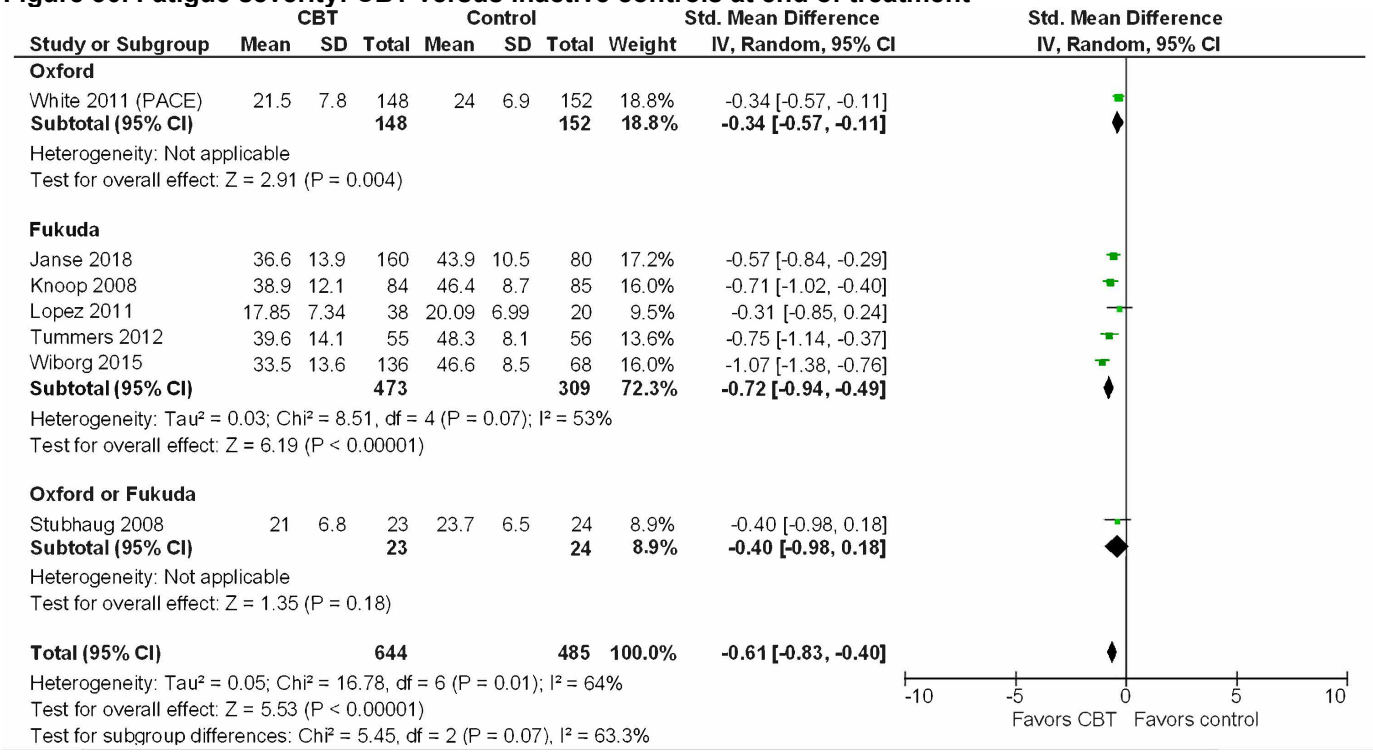
(Fukuda: 5 trials, N=782, SMD -0.72, 95% CI -0.94 to -0.49, I^2=53%; Oxford: 1 trial, N=300, SMD -0.34, 95% CI -0.57 to -0.11; and Oxford or Fukuda: 1 trial: N=47, SMD -0.40, 95% CI -0.98 to 0.18; p for subgroup difference=0.07). There was no subgroup effect for CBT type; the most common CBT type was group/face-to-face (3 trials, N=309, SMD -0.63, 95% CI -1.18 to -0.09, I^2=75%).[72,75,77] Excluding the high risk of bias trial[72] had little effect on the pooled estimate (6 trials, N=1,071, SMD -0.65, 95% CI -0.88 to -0.41, I^2=68%).

CBT was also associated with decreased fatigue severity versus controls at post-intervention follow-up, though the estimate was based on fewer studies (3 trials, N=489, SMD -0.57, 95% CI -0.89 to -0.25, I^2=57%; **Figure 31**).[38,73,74] Statistical heterogeneity was moderate; however, estimates favored CBT in all trials (SMD ranged from -0.36 to -1.06).

CBT was associated with increased likelihood of improvement in fatigue versus wait list or usual specialist care (4 trials, N=784, RR 3.00, 95% CI 0.95 to 9.49, I^2=93%; **Figure 32**).[38,71,76,77] In three trials, improvement in fatigue was defined as a Checklist Individual Strength severity subscale score <35 and reliable change index >1.96.[71,76,77] The fourth trial (PACE) found CBT associated with increased likelihood of improvement in fatigue versus usual specialist care, based on ≥2 point improvement on the 11-item 0 to 33 Chalder fatigue scale (76% vs. 65%, RR 1.17, 95% CI 1.01 to 1.36; ARD 11%, 95% CI 1% to 21%).[38] This differed from the original PACE protocol, which defined improvement in fatigue as a score ≤3 on the 11-item 0 to 11 Chalder fatigue scale or >50% improvement from baseline (26% vs. 13%, RR 1.99, 95% CI 1.23 to 3.20).[40] Using the PACE protocol definition for fatigue improvement, the pooled estimate was very similar (4 trials, N=805, RR 3.30, 95% CI 1.96 to 5.55, I^2=51%).

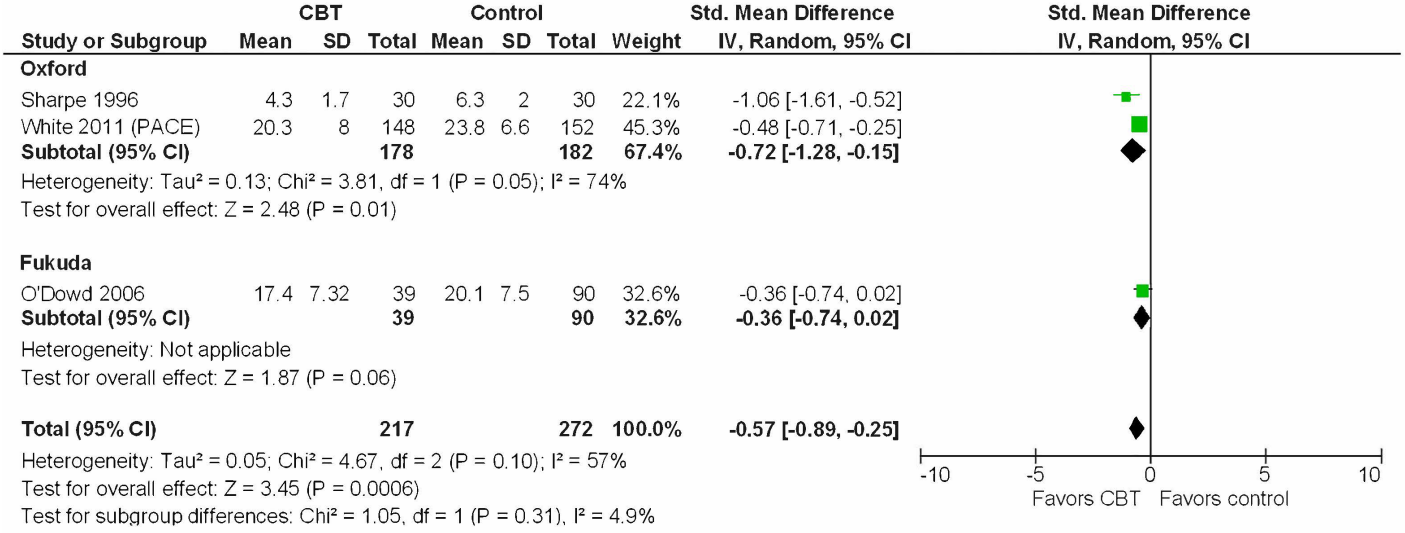
The PACE trial also evaluated longer-term, post-trial outcomes (median duration from randomization 31 months).[63] Thirty-one percent of patients in the CBT group and 63 percent in the usual specialist care group received non-randomly allocated therapies between the end of the trial at 1 year and long-term follow-up. Fatigue severity on the 11-item 0 to 33 Chalder scale was slightly improved at long-term follow-up compared with end-of-trial scores in both the exercise (mean change -2.2 points, 95% CI -3.7 to -0.6) and usual specialist care groups (-3.9 points, 95% CI -5.3 to -2.6). At long-term post-trial follow-up, there were no differences between CBT vs. specialist medical care in fatigue (mean difference -1.4, 95% CI -3.4 to 0.7), based on mixed model analyses.

Figure 30. Fatigue severity: CBT versus inactive controls at end of treatment



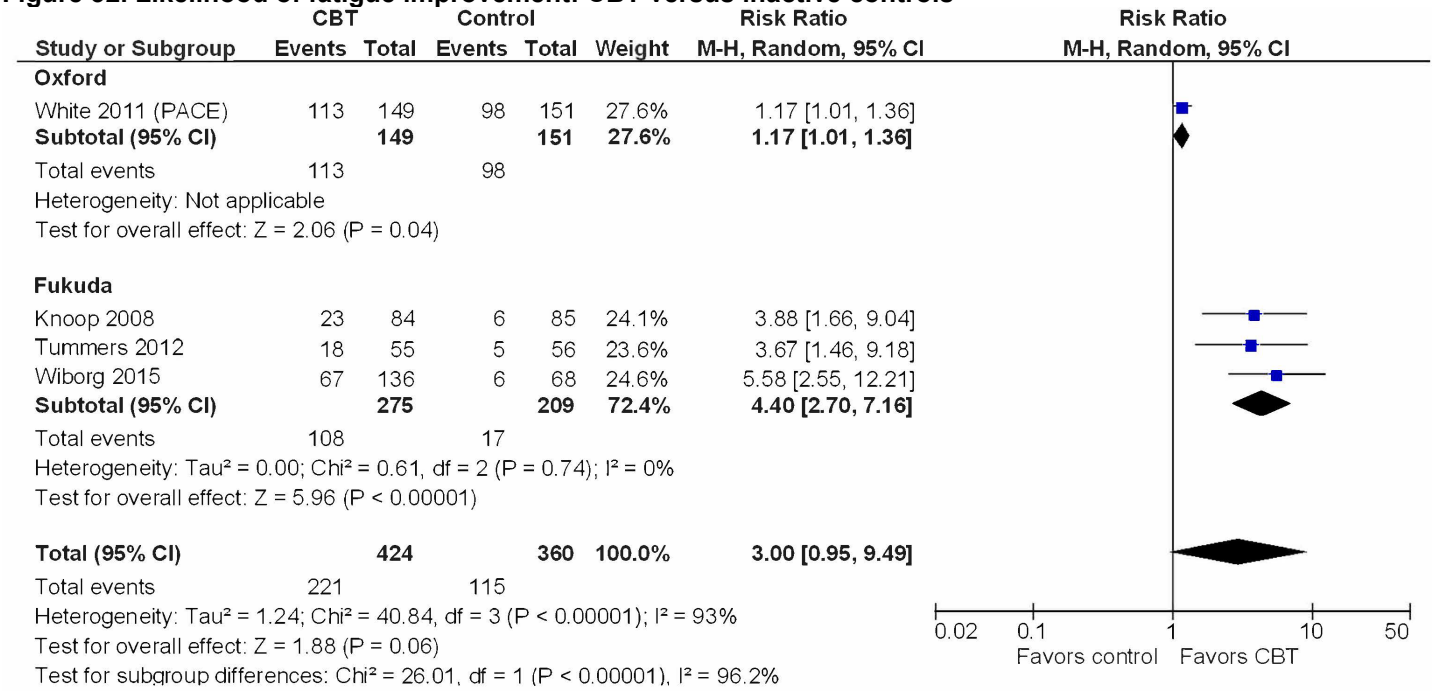
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 31. Fatigue severity: CBT versus inactive controls at post-intervention follow-up



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 32. Likelihood of fatigue improvement: CBT versus inactive controls



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Function

CBT was associated with decreased severity of functional impairment versus wait list or usual specialist care (5 trials, N=1024, mean difference 6.58, 95% CI 3.76 to 9.39, I^2=0%;[38,70,71,76,77] **Figure 33**). Estimates consistently favored CBT when trials were stratified according to the control type, ME/CFS case definition, or type of CBT, and there were no statistically significant subgroup differences (**Table 9**). The most commonly ME/CFS case definition was the Fukuda criteria (4 trials, N=724, mean difference 6.92, 95% CI 3.54 to 10.31, I^2=0%).[70,71,76,77]

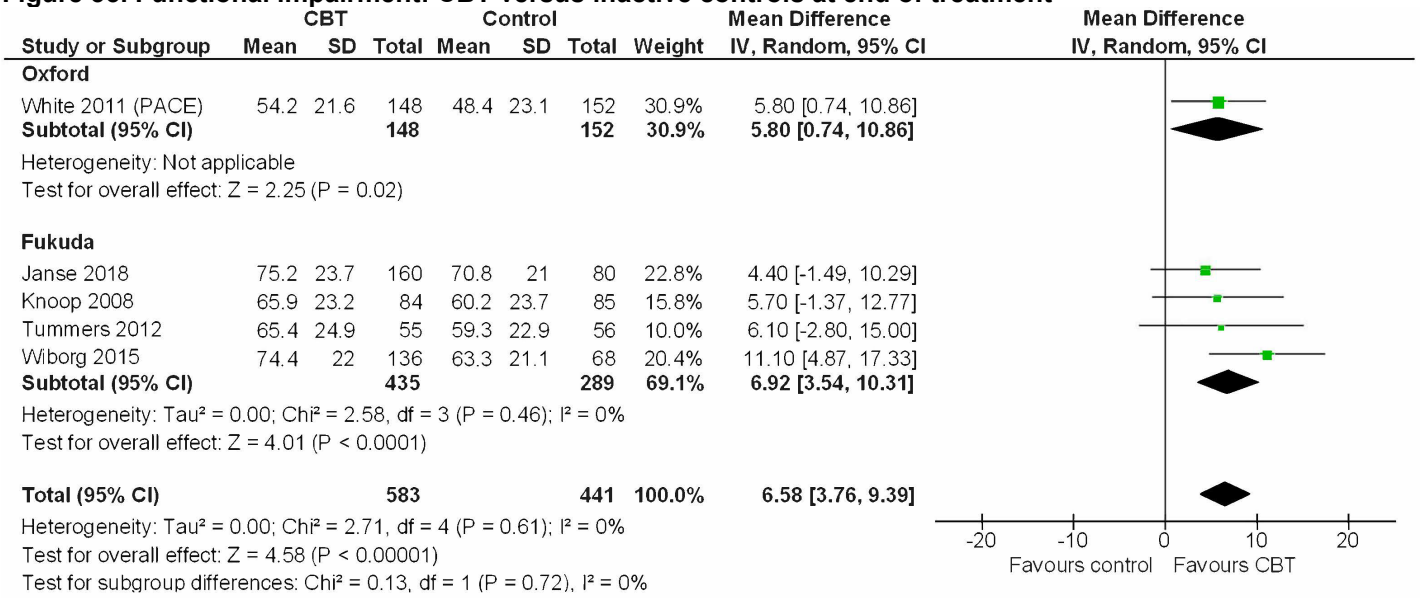
CBT was also associated with decreased severity of functional impairment versus controls at post-intervention follow-up (3 trials, N=489, SMD 0.37, 95% CI 0.08 to 0.66, I^2=50%;[38,73,74] **Figure 34**). Although statistical heterogeneity was present, results from all trials favored CBT. One trial evaluated function with the SF-36 physical function subscale (mean difference 7.40 on a 0 to 100 scale, 95% CI 1.88 to 12.92),[38] one trial used the SF-36 physical component summary (mean difference 1.50 on a 0 to 100 scale, 95% CI -1.66 to 4.66),[73] and one trial reported the percentage interference with activity (mean difference 13.00 on a 0 to 100 scale [reversed so higher score indicates decreased functional impairment], 95% CI 5.04 to 20.96).[74]

Three trials evaluated improvement in function as a dichotomous outcome.[38,73,74] Functional improvement was defined as SF-36 physical component summary score improved ≥15% from baseline,[73] Karnofsky Performance Scale (KPS) improved ≥10 points from baseline,[74] or SF-36 physical function subscale score improved ≥8 points from baseline.[38] The latter trial (PACE) used a different definition in the main publication than the study protocol,[39] which defined functional improvement as an SF-36 physical function score of ≥75 or ≥50% improvement from baseline. Using the definition from the main PACE publication, there was no difference between CBT versus attention control, usual care, or specialist care in likelihood of functional improvement (3 trials, N=488, RR 1.06, 95% CI 0.83 to 1.35, I^2=47%; **Figure 35**). Results were

similar when using data based on the PACE protocol definition[40] (3 trials, N=509, RR 1.03, 95% CI 0.86 to 1.22, I^2=0%).

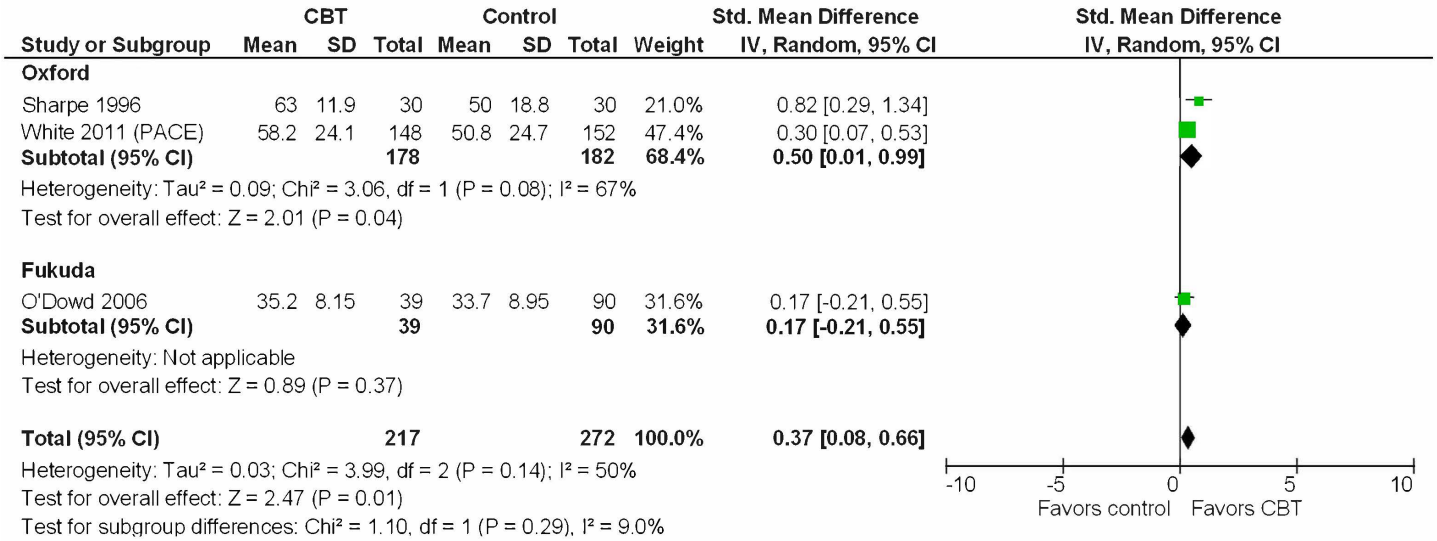
The PACE trial also evaluated longer-term, post-trial outcomes (median duration from randomization 31 months).[63] For SF-36 physical function, both CBT (mean change 3.3, 95% CI 0.02 to 6.7) and usual specialist care (mean change 7.1, 95% CI 4.0 to 10.3) were associated with improved scores at long-term follow-up compared with the end of the trial. At long-term follow-up, there was no difference between graded exercise vs. usual specialist medical care in SF-36 physical function (mean difference 2.8, 95% CI -3.2 to 8.8), based on mixed model analyses.

Figure 33. Functional impairment: CBT versus inactive controls at end of treatment



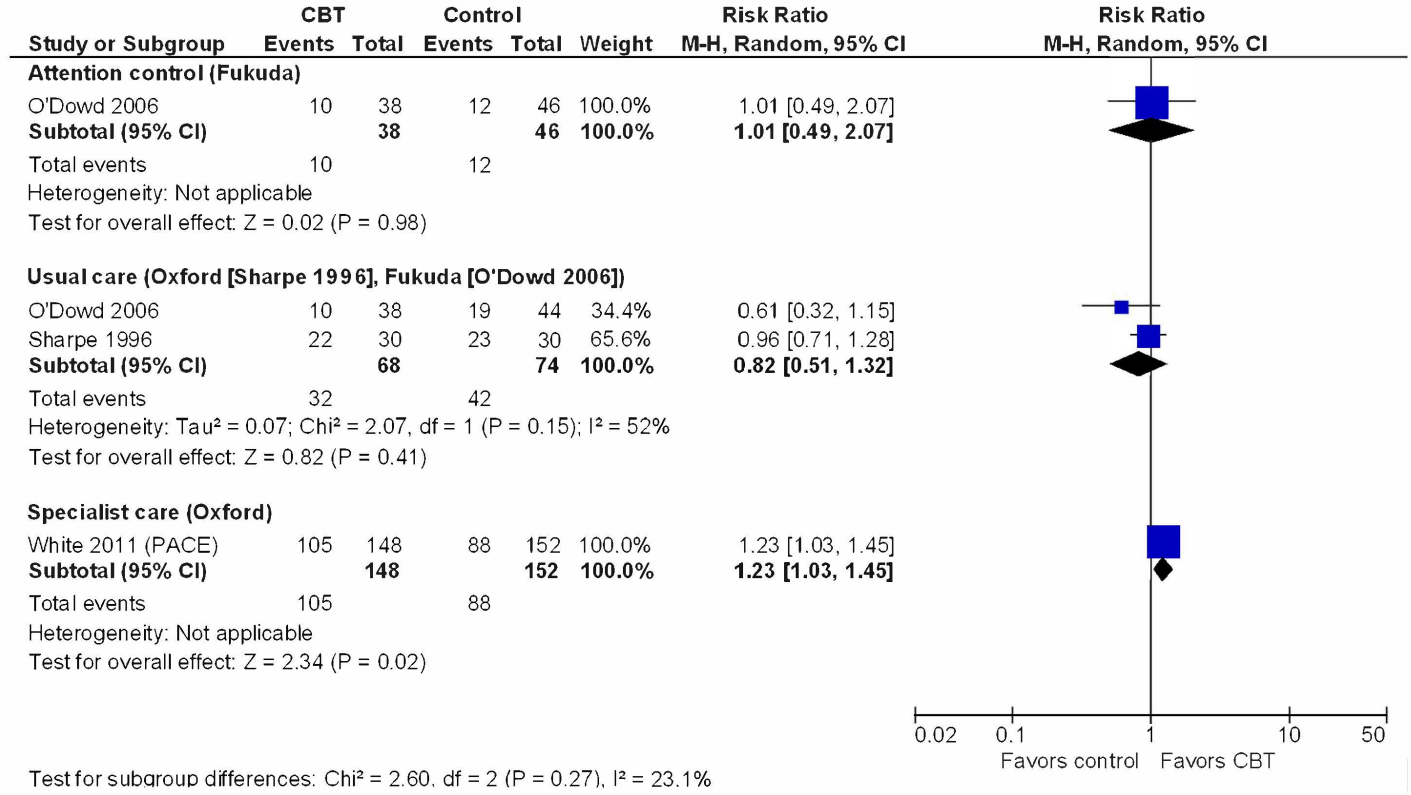
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 34. Functional impairment: CBT versus inactive controls at post-intervention follow-up



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 35. Likelihood of functional improvement: CBT versus inactive controls



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

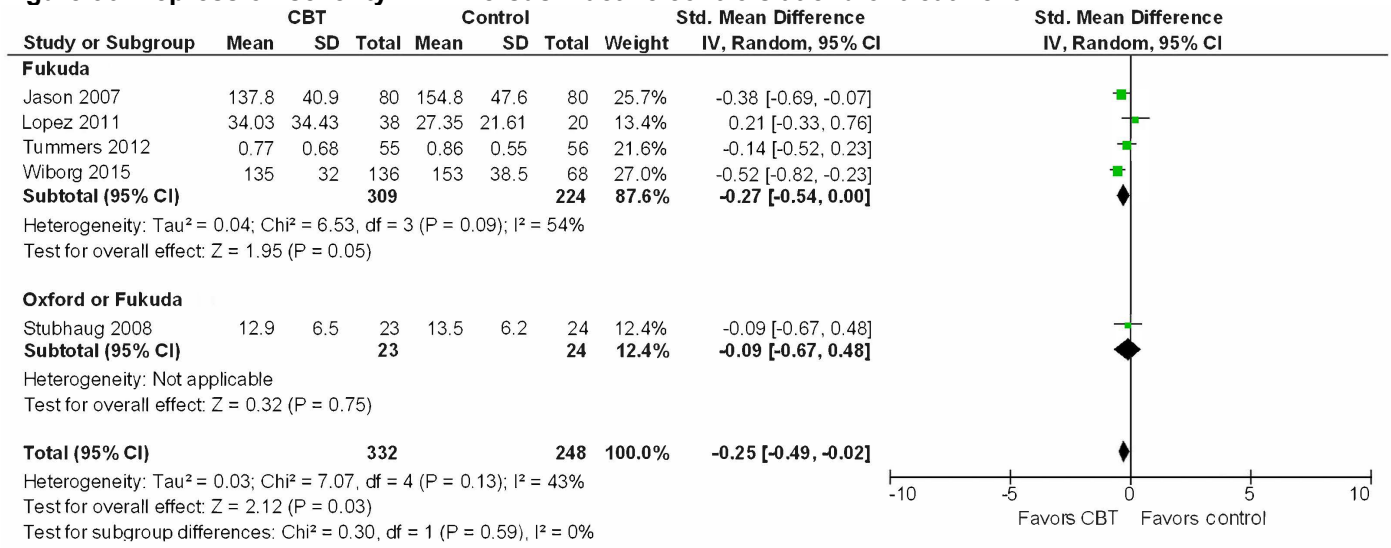
Depression and anxiety

CBT was associated with decreased depression severity versus wait list, placebo medication, or an attention control at the end of the intervention (5 trials, N=660, SMD -0.25, 95% CI -0.49 to -0.02, I^2=43%;[70,72,75-77] **Figure 36**). In two trials (N=444) that used the 90 to 450 Symptom Checklist-90 scale (SCL-90) the mean difference was -17.66 (95% CI -25.66 to -9.65, I^2=0%).[70,77] The other trials each used a different scale for depression severity (**Table 9**). Excluding a high risk of bias trial[72] had little effect on the pooled estimate (4 trials, N=602, SMD -0.35, 95% CI -0.54 to -0.17, I2=11%). CBT was also associated with decreased depression severity versus usual care, usual specialist care, or an attention control at post-intervention follow-up (3 trials, N=483, mean difference -1.24, 95% CI -2.01 to -0.47 on the 0 to 21 HADS depression scale, I^2=12%;[38,73,74] **Figure 37**).

CBT was associated with decreased anxiety severity versus usual care, usual specialist care, or an attention control at post-intervention follow-up (3 trials, N=481, mean difference -1.22, 95% CI -1.94 to -0.49 on the 0 to 21 HADS anxiety scale, I^2=0%;[38,73,74] **Figure 38**). No trial evaluated anxiety at the end of the intervention.

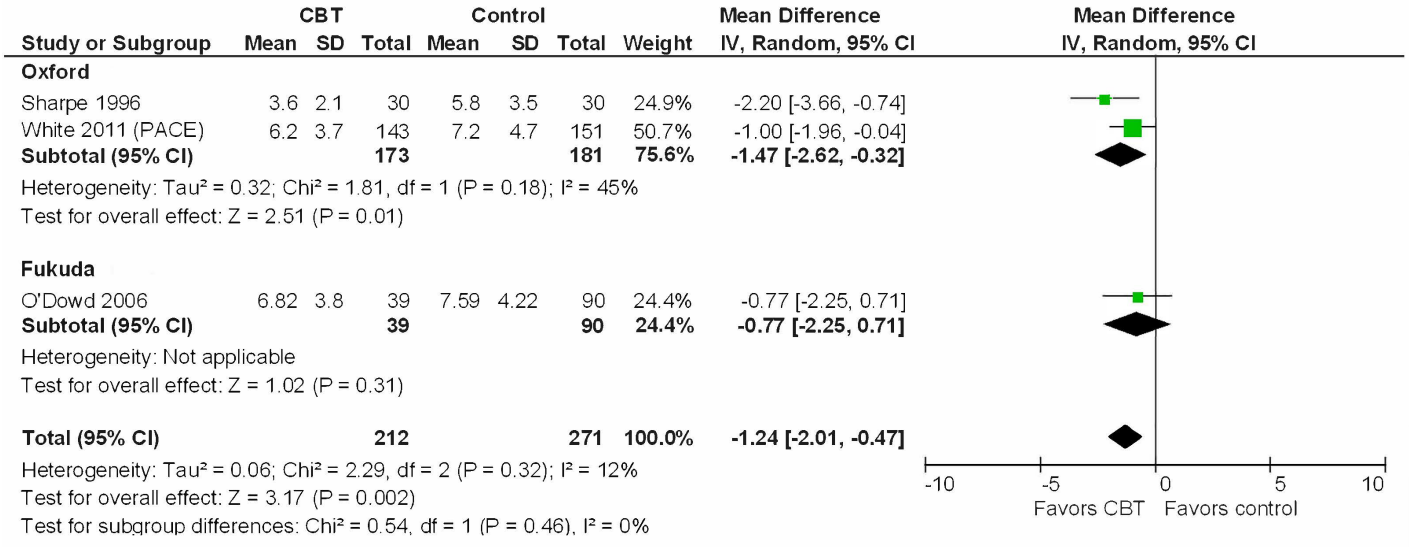
For both depression and anxiety, results were consistent in stratified analyses based on control type, ME/CFS criteria, or CBT type, with no statistically significant subgroup differences (**Table 9**). However, stratified analyses were based on small numbers of trials.

Figure 36. Depression severity: CBT versus inactive controls at end of treatment



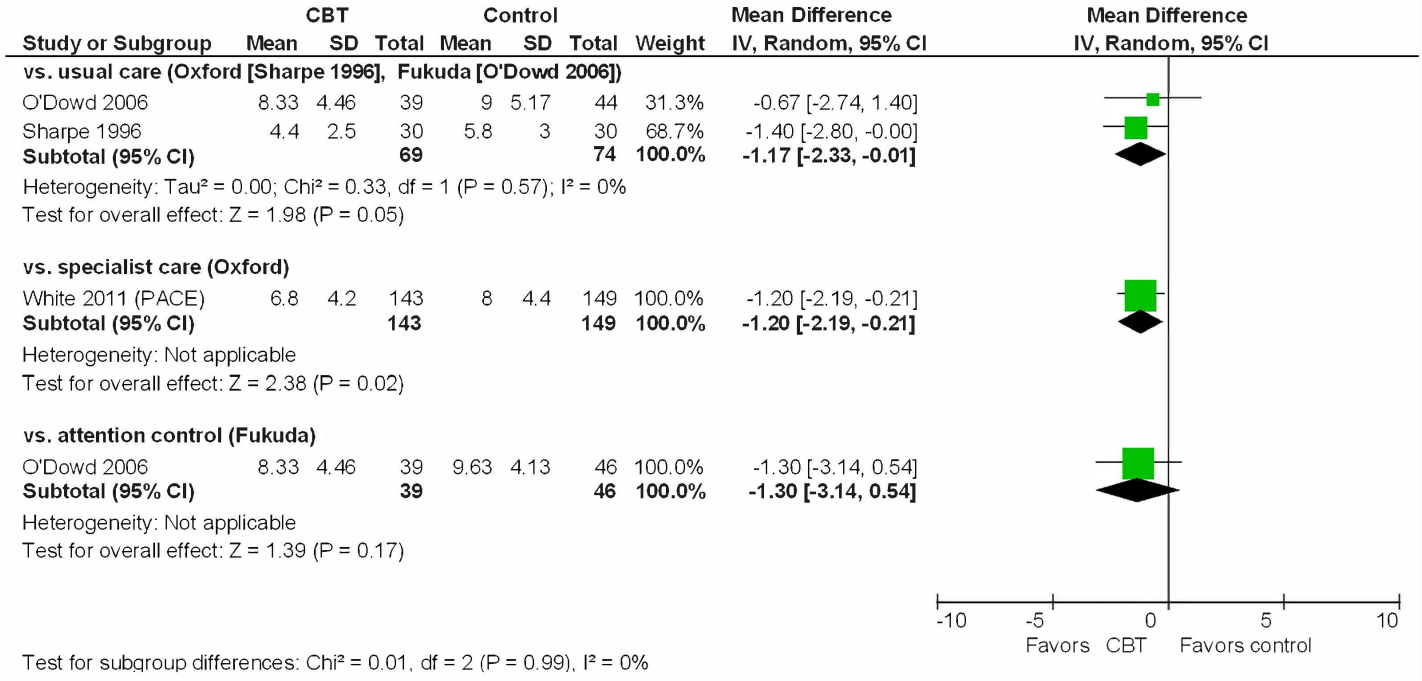
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; IV = instrumental variable; SD = standard deviation; Std = standard

Figure 37. Depression severity: CBT versus inactive controls at post-intervention follow-up



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 38. Anxiety severity: CBT versus inactive controls at post-intervention follow-up



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Sleep

One trial found CBT associated with improved sleep quality versus usual specialist care at post-intervention follow-up (N=292, mean difference -1.20 on the 0 to 20 Jenkins Sleep Questionnaire, 95% CI -2.19 to -0.21).[38] No trial evaluated effects of CBT versus controls on sleep quality at the end of the intervention.

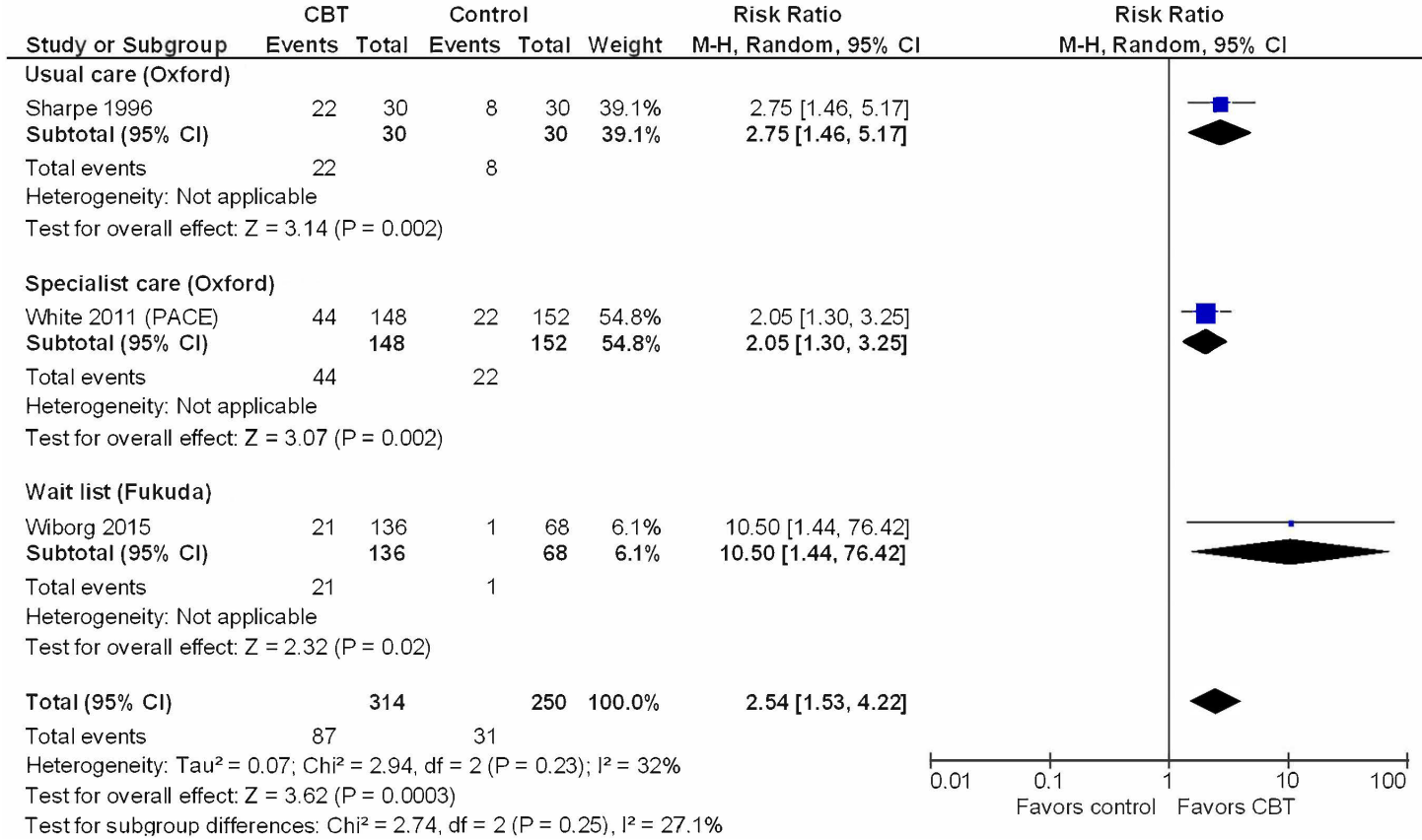
Pain

A post-hoc analysis of the PACE trial found CBT associated with decreased severity of muscle pain (mean difference -0.38, 95% CI -0.69 to -0.08 on a 0 to 4 scale).[64] The effect on joint pain was not statistically significant (mean difference -0.25, 95% CI -0.58 to 0.08).

Recovery

Three trials evaluated effects of CBT versus usual care, usual specialist care, or wait list on likelihood of recovery.[38,74,77] One trial[74] defined recovery as a KPS final score of ≥80 and one trial[77] defined recovery as a Checklist Individual Strength severity subscale score <27, SF-36 physical function score ≥80, and Sickness Impact Profile total score <203. The third trial, PACE, used a different definition for recovery in the main publication than described in the original protocol (see Results, exercise for details).[38] Based on the published results from PACE (proportion of patients meeting definition for recovery 22% vs. 7%), CBT was associated with increased likelihood of recovery versus usual care, usual specialist care, or wait list (3 trials, N=564, RR 2.54, 95% 1.53 to 4.22, I^2=32%; ARD 21%, 95% CI 8% to 34%;[38,74,77] **Figure 39**). Replacing the data from PACE with results based on the original protocol definition for recovery (proportion meeting definition 7% vs. 3%)[39,65] resulted in a similar pooled estimate (3 trials, N=585, RR 2.88, 95% CI 1.62 to 5.11, I^2=9%; ARD 17%, 95% CI 2% to 32%).

Figure 39. Likelihood of recovery: CBT versus inactive controls



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Overall improvement

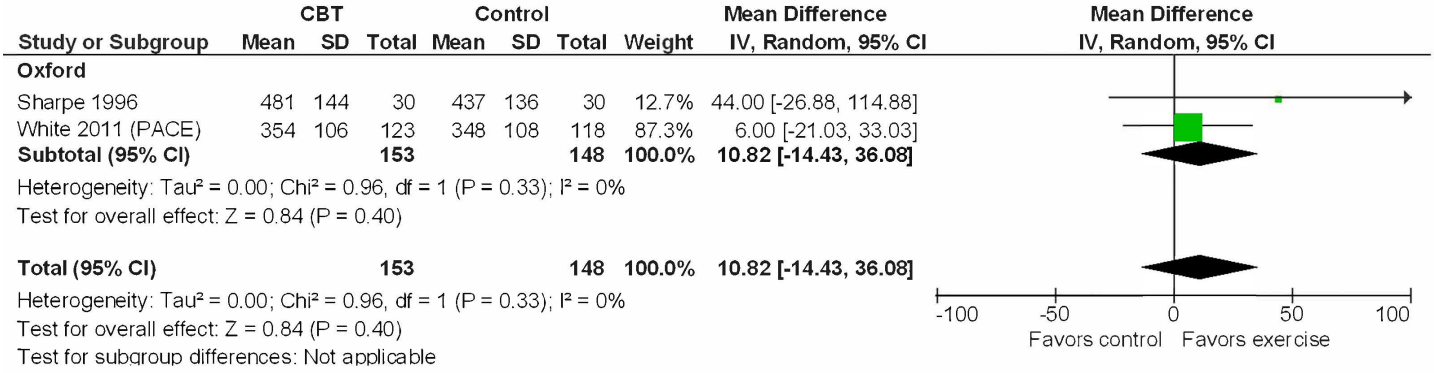
In the PACE trial, a composite outcome for overall improvement, based on an 11-item Chalder fatigue scale (0 to 11) score ≤3 or >50% improvement from baseline and SF-36 physical function score ≥75 or >50% improvement from baseline at 52 weeks, was described as the primary outcome in the study protocol[39] but not reported in the main publication. In a subsequent publication, the authors reported that CBT plus usual specialist care was associated with greater likelihood of overall improvement than usual specialist care alone, using the protocol definition[40] (20% vs. 10%, RR 1.99, 95% CI 1.14 to 3.48).

One other trial found CBT associated with improved overall efficacy (defined as Checklist Individual Strength fatigue severity subscale score <36 and reliable change index >1.96, SF-36 physical function score ≥65, and Sickness Impact Profile overall impairment score <700) versus wait list (38% vs. 2.9%, RR 13.00, 95% CI 3.26 to 51.78).[77]

6-minute walk test

Two trials found no difference between CBT versus usual care or usual specialist care in the 6-minute walk test at post-intervention follow-up (2 trials, N=301, mean difference 10.82 meters, 95% CI -14.43 to 36.08, I^2=0%;[38,74] **Figure 40**). As previously noted, in PACE results for the 6-minute walk test were available for fewer patients (72.2%) than for other outcomes (e.g., work and social adjustment scale, 91.9%).

Figure 40. 6-minute walk test: CBT versus inactive controls at post-intervention follow-up



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Harms

Data on harms were limited to the PACE trial (N=321), which found no differences between CBT versus usual specialist care alone in likelihood of serious adverse events (RR 0.99, 95% 0.36 to 2.77) or physical function worsening (RR 0.83, 95% CI 0.26 to 2.66).[38] Only one case of withdrawal due to adverse events occurred. In PACE, CBT was associated with decreased likelihood of post-exertional malaise (49% vs. 63%, RR 0.78, 95% CI 0.64 to 0.95).

Table 9. Cognitive behavioral therapy vs. inactive controls

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
Fatigue, end of intervention	7 (1129)	SMD -0.61 (-0.83 to -0.40)	64%	--
By control type: vs. wait list	4 (724)	SMD -0.77 (-0.99 to -0.55)	50%	0.05
vs. usual specialist care	1 (300)	SMD -0.34 (-0.57 to -0.11)	--	--
vs. attention control	1 (58)	SMD -0.31 (-0.85 to 0.24)	--	--
vs. placebo medication	1 (47)	SMD -0.40 (-0.98 to 0.18)	--	--
On original scale: Chalder (11-item, 0 to 33)	2 (347)	MD -2.53 (-4.06 to -1.00)	0%	--
Checklist Individual Strength, fatigue severity (8 to 56)	4 (724)	MD -9.20 (-12.09 to -6.31)	66%	--
Profile of Mood States fatigue/inertia (0 to 28)	1 (58)	MD -2.24 (-6.09 to 1.61)	--	--
By ME/CFS criteria: Oxford	1 (300)	SMD -0.34 (-0.57 to -0.11)	--	0.07
Fukuda	5 (782)	SMD -0.72 (-0.94 to -0.49)	53%	--
Oxford or Fukuda	1 (47)	SMD -0.40 (-0.98 to 0.18)	--	--
By CBT type: Individual, face-to-face	1 (300)	SMD -0.34 (-0.57 to -0.11)	--	0.14
Individual, self-guided	2 (280)	SMD -0.73 (-0.97 to -0.48)	0%	--
Individual, web-based	1 (240)	SMD -0.57 (-0.84 to -0.29)	--	--
Group, face-to-face	3 (309)	SMD -0.63 (-1.18 to -0.09)	75%	--
Excluding high risk of bias trial	6 (1071)	SMD -0.65 (-0.88 to -0.41)	68%	--
Using difference in change from baseline	6 (1049)	SMD -0.69 (-0.98 to -0.40)	79%	--
Fatigue, post-intervention	3 (489)	SMD -0.57 (-0.89 to -0.25)	57%	--
By control type: vs. usual care	1 (143)	SMD -0.61 (-1.47 to 0.24)	--	0.95
vs. usual specialist care	1 (300)	SMD -0.48 (-0.71 to -0.25)	--	--
vs. attention control	1 (85)	SMD -0.52 (-0.96 to -0.09)	--	--
On original scale: Chalder (11-item, 0 to 33)	2 (429)	MD -3.29 (-4.71 to -1.86)	0%	--
0 to 10 Likert	1 (60)	MD -2.00 (-2.94 to -1.06)	--	--

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
By ME/CFS criteria Oxford	2 (360)	SMD -0.72 (-1.28 to -0.15)	74%	0.31
Fukuda	1 (129)	SMD -0.36 (-0.74 to 0.02)	--	--
By CBT type: Individual, face-to-face	2 (360)	SMD -0.72 (-1.28 to -0.15)	74%	0.31
Group, face-to-face	1 (129)	SMD -0.36 (-0.74 to 0.02)	--	--
Using difference in change from baseline	3 (489)	SMD -0.58 (-0.87 to -0.29)	47%	--
Fatigue improvement (dichotomous)	4 (784)	RR 3.00 (0.95 to 9.49)	93%	--
By control type vs. wait list	3 (484)	RR 4.40 (2.70 to 7.16)	0%	<0.00001
vs. usual specialist care	1 (300)	RR 1.17 (1.01 to 1.36)	--	--
By ME/CFS criteria Oxford	1 (300)	RR 1.17 (1.01 to 1.36)	--	--
Fukuda	3 (484)	RR 4.40 (2.70 to 7.16)	0%	--
By CBT type Individual, face-to-face	1 (300)	RR 1.17 (1.01 to 1.36)	--	<0.00001
Individual, self-guided	2 (280)	RR 3.78 (2.03 to 7.04)	0%	--
Group, face-to-face	1 (204)	RR 13.00 (3.26 to 51.78)	--	--
Using original PACE definition	4 (805)	RR 3.30 (1.96 to 5.55)	51%	--
SF-36 physical function (0 to 100), end of intervention	5 (1024)	MD 6.58 (3.76 to 9.39)	0%	--
By control type: vs. wait list	4 (724)	MD 6.92 (3.54 to 10.31)	0%	0.72
vs. usual specialist care	1 (300)	MD 5.80 (0.74 to 10.86)	--	--
By ME/CFS criteria: Oxford	1 (300)	MD 5.80 (0.74 to 10.86)	--	0.72
Fukuda	4 (724)	MD 6.92 (3.54 to 10.31)	0%	--
By CBT type Individual, face-to-face	1 (300)	MD 5.80 (0.74 to 10.86)	--	0.44
Individual, self-guided	2 (280)	MD 5.85 (0.32 to 11.39)	0%	--
Individual, web-based	1 (160)	MD 4.40 (-1.49 to 10.29)	--	--
Group, face-to-face	1 (204)	MD 11.10 (4.87 to 17.33)	--	--
Using difference in change from baseline	5 (1024)	MD 9.24 (4.68 to 13.79)	64%	--
Function, post-intervention	3 (489)	SMD 0.37 (0.08 to 0.66)	50%	--
By control type: vs. usual care	1 (143)	SMD 0.40 (-0.37 to 1.18)	--	0.97
vs. usual specialist care	1 (300)	SMD 0.30 (0.07 to 0.53)	--	--
vs. attention control	1 (85)	SMD 0.33 (-0.10 to 0.76)	--	--
On original scale: SF-36 physical function (0 to 100)	1 (300)	MD 7.40 (1.88 to 12.92)	--	--
SF-36 physical component summary (0 to 100)	1 (129)	MD 1.50 (-1.66 to 4.66)	--	--
Percentage interference with activity (0 to 100, reversed so that higher score indicates better function)	1 (60)	MD 13.00 (5.04 to 20.96)	--	--
By ME/CFS criteria Oxford	2 (360)	SMD 0.50 (0.01 to 0.99)	67%	0.29
Fukuda	1 (129)	SMD 0.17 (-0.21 to 0.55)	--	--
By CBT type Individual, face-to-face	2 (360)	SMD 0.50 (0.01 to 0.99)	67%	0.29
Group, face-to-face	1 (129)	SMD 0.17 (-0.21 to 0.55)	--	--
Using difference in change from baseline	3 (489)	SMD 0.37 (-0.07 to 0.81)	77%	--
Functional improvement	3 (488)	RR 1.06 (0.83 to 1.35)	47%	--
By control type: vs. usual care	2 (142)	RR 0.82 (0.51 to 1.32)	52%	0.27
vs. usual specialist care	1 (300)	RR 1.23 (1.03 to 1.45)	--	--
vs. attention control	1 (84)	RR 1.01 (0.49 to 2.07)	--	--
By ME/CFS criteria: Oxford	2 (360)	RR 1.11 (0.88 to 1.41)	52%	0.25
Fukuda	1 (128)	RR 0.76 (0.42 to 1.40)	--	--

Outcome	Number of studies (N)	Estimate (95% CI)	I^2	p for subgroup difference
By CBT type Individual, face-to-face	2 (360)	RR 1.11 (0.88 to 1.41)	52%	0.25
Group, face-to-face	1 (128)	RR 0.76 (0.42 to 1.40)	--	--
Using original PACE definition	3 (509)	RR 1.03 (0.86 to 1.22)	0%	--
Depression, end of intervention	5 (660)	SMD -0.26 (-0.49 to -0.03)	45%	--
By control type: vs. wait list	3 (555)	SMD -0.38 (-0.58 to -0.18)	18%	0.10

Abbreviations: CFS = chronic fatigue syndrome; CI = confidence interval; CBT = cognitive behavioral therapy; MD = mean difference; ME = myalgic encephalomyelitis; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference

CBT Versus Active Interventions

Four trials (N=876) compared CBT versus active interventions in adults (**Table 7 and 8, Evidence Table Appendix E2**).[38,54,69,75] The active interventions were exercise (2 trials),[38,54] relaxation (2 trials),[54,69] cognitive therapy (1 trial),[54] adaptive pacing (1 trial),[38] and mirtazapine (1 trial).[75] All the trials except for one (mirtazapine)[75] were included in the prior AHRQ report. The duration of the CBT intervention ranged from 12 weeks to 6 months. Three trials[38,69,75] evaluated patients at the end of the intervention and two trials[38,54] evaluated patients 26 to 29 weeks following the end of the intervention. All of the trials were rated medium risk of bias. Results stratified by the active comparator are summarized in **Table 10** and shown in **Figures 41 to 56**.

CBT versus exercise

See Results for exercise (pages 41 to 56).

CBT versus relaxation

Two trials compared CBT versus relaxation in patients who met the Fukuda case definition.[54] or both the Oxford and Fukuda case definitions.[69] The duration of therapy was 4 to 6 months. One trial[69] evaluated outcomes at the end of therapy and both trials evaluated outcomes 6 months following the completion of therapy. There was no difference between CBT versus relaxation in fatigue severity at the end of the intervention (1 trial, N=60, mean difference -0.3, 95% CI -2.4 to 1.8 on the 11-item 0 to 11 Chalder scale, **Figure 41**) or at post-intervention follow-up (2 trials, N=117, SMD -0.49, 95% CI -1.03 to 0.04, I^2=52%, **Figure 42**). CBT was associated with decreased severity of functional impairment at the end of therapy (1 trial, mean difference 21.6, 95% CI 7.8 to 35.4 on the 0 to 100 SF-36 physical function subscale, **Figure 43**) but the difference was not statistically significant at post-intervention follow-up (2 trials, N=117, mean difference 15.45, 95% CI -19.60 to 50.49, I^2=91%, **Figure 44**). Statistical heterogeneity was present at post-intervention follow-up for both fatigue (SMD -0.76, 95% CI -1.29 to -0.24 in one trial[69] and SMD -0.22, 95% CI -0.74 to 0.30 in the other trial)[54] and function (mean difference 33.2, 95% CI 19.31 to 47.09 on the SF-36 physical function subscale in one trial[69] and -2.6, 95% CI -17.7 to 12.5 in the other trial).[54]

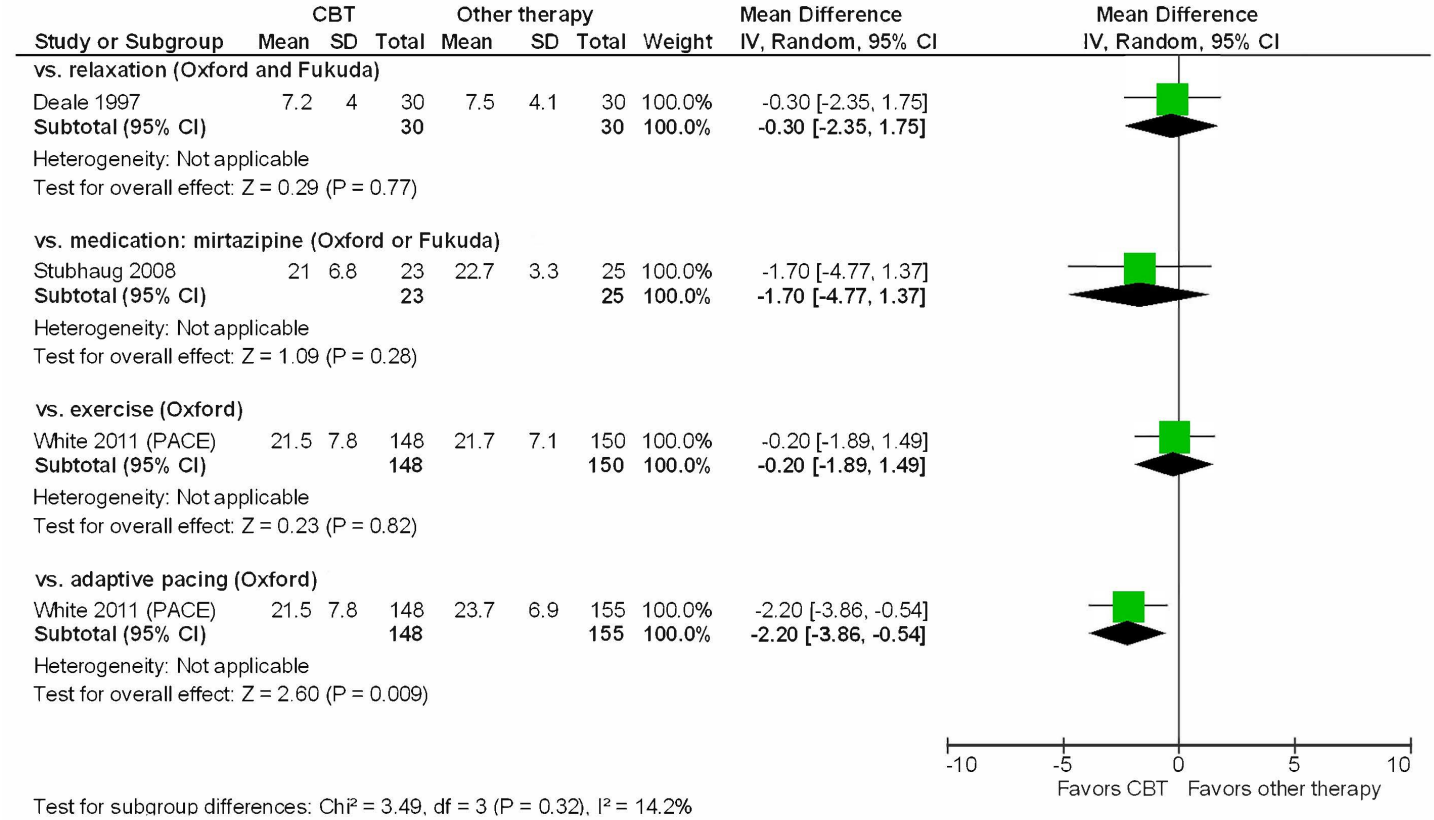
There were no differences between CBT versus relaxation in depression at the end of therapy (1 trial, N=60, mean difference -3.0, 95% CI -6.3 to 0.3 on the 0 to 63 Beck Depression Inventory, **Figure 45**)[69] or at post-intervention follow-up (2 trials, N=117, mean difference -1.4, 95% CI -4.7 to 1.9 on the Beck Depression Inventory, **Figure 46**).[54,69] One trial found no differences between CBT versus relaxation in anxiety **(Figure 47)**, pain **(Figure 48)**, sleep

quality **(Figure 49)**, the 6-minute walk test **(Figure 50)**, or severity of sore throat, tender lymph nodes, impaired memory, or headache symptoms at post-intervention follow-up.[54]

CBT was associated with increased likelihood of functional improvement versus relaxation (2 trials, N=110, RR 1.78, 95% CI 0.41 to 7.86, I^2=79%, **Figure 51**).[54,69] However, statistical heterogeneity was present (RR 0.80, 95% CI 0.28 to 2.34 in one trial[69] and RR 3.66, 95% CI 1.60 to 8.35 in the other trial)[54] Functional improvement was defined as SF-36 physical function score improved by ≥50 or end score ≥83 in one trial and as improvement in SF-36 physical function score greater than the age adjusted reliable change index and within 1 standard deviation of the normative value in the other trial.[54] CBT was also associated with increased likelihood of recovery (2 trials, N=110, RR 4.41, 95% CI 1.79 to 10.83, I^2=0%, **Figure 52**).[54,69] Recovery was defined as an 11-item Chalder fatigue score <4 in one trial[69] and as no longer meeting CFS criteria based on physician diagnosis in the other trial.[54] One trial found CBT associated with increased likelihood of having a global improvement rating of better or much better (70% vs. 31%, RR 2.29, 95% CI 1.22 to 4.28).[69] One trial found no difference between CBT versus relaxation in severity of post-exertional malaise (mean difference 3.4, 95% CI -14.6 to 21.4 on the 0 to 100 CFS Questionnaire).[54] The trials did not report serious adverse events, withdrawals due to adverse events, or other harms.

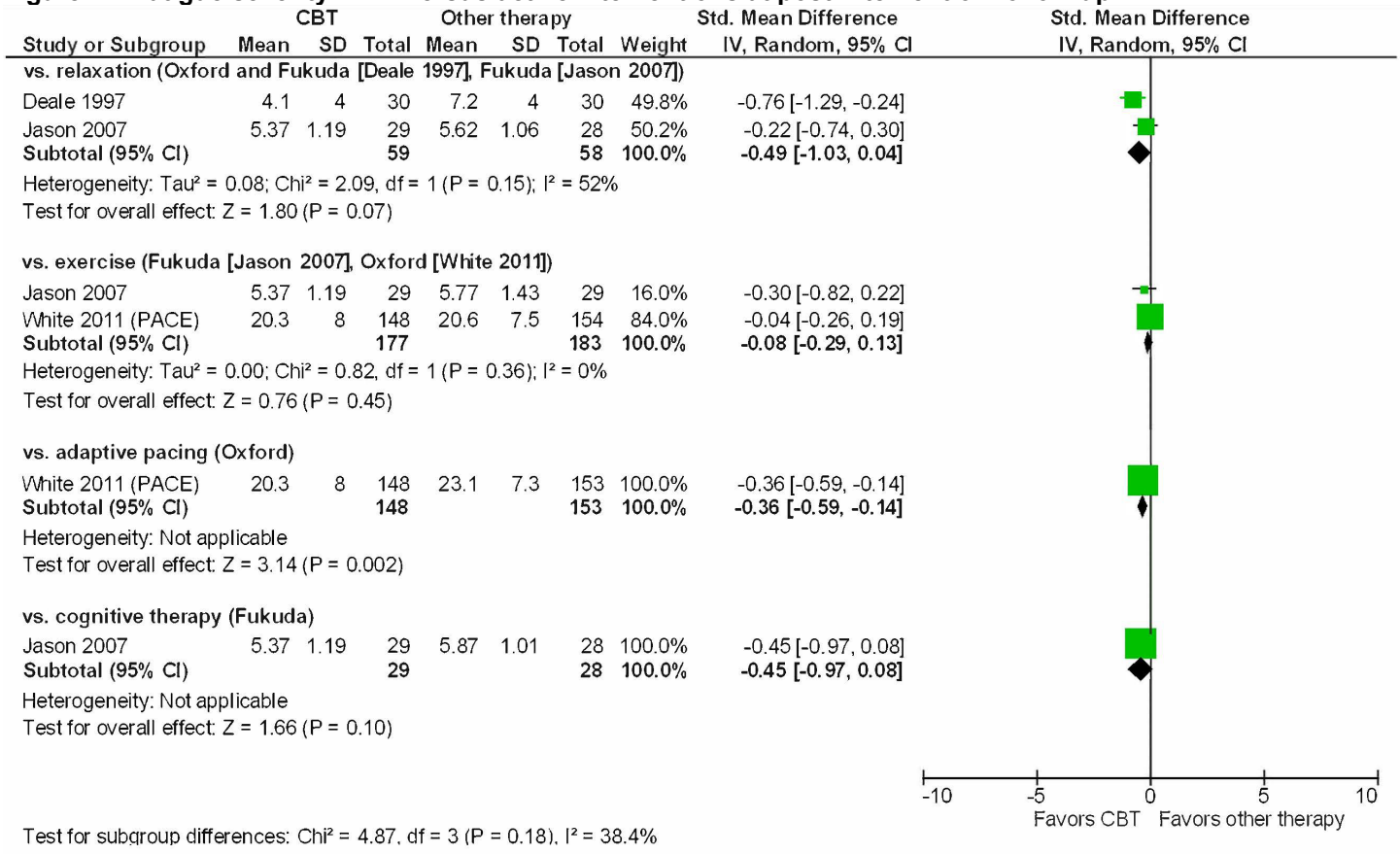
One of the trials (N=53) evaluated long-term outcomes at 5 years (4 years after trial completion).[78] Around 56% of patients in both groups received additional post-trial CFS treatments. CBT was associated with increased likelihood of self-rated global improvement of "much better" or "very much better" (68% vs. 36%, RR 1.90, 95% CI 1.08 to 3.35). There were no differences in likelihood of functional improvement (SF-36 physical function >83, 48% vs. 32%, RR 1.49, 95% CI 0.76 to 2.93), fatigue improvement (11-item 0 to 11 Chalder <4, 28% vs. 25%, RR 1.12, 95% CI 0.46 to 2.75), and psychological distress (General Health Questionnaire <4, 48% vs. 54%, RR 0.90, 95% CI 0.53 to 1.53), though some estimates were imprecise.

Figure 41. Fatigue severity: CBT versus active interventions at end of treatment



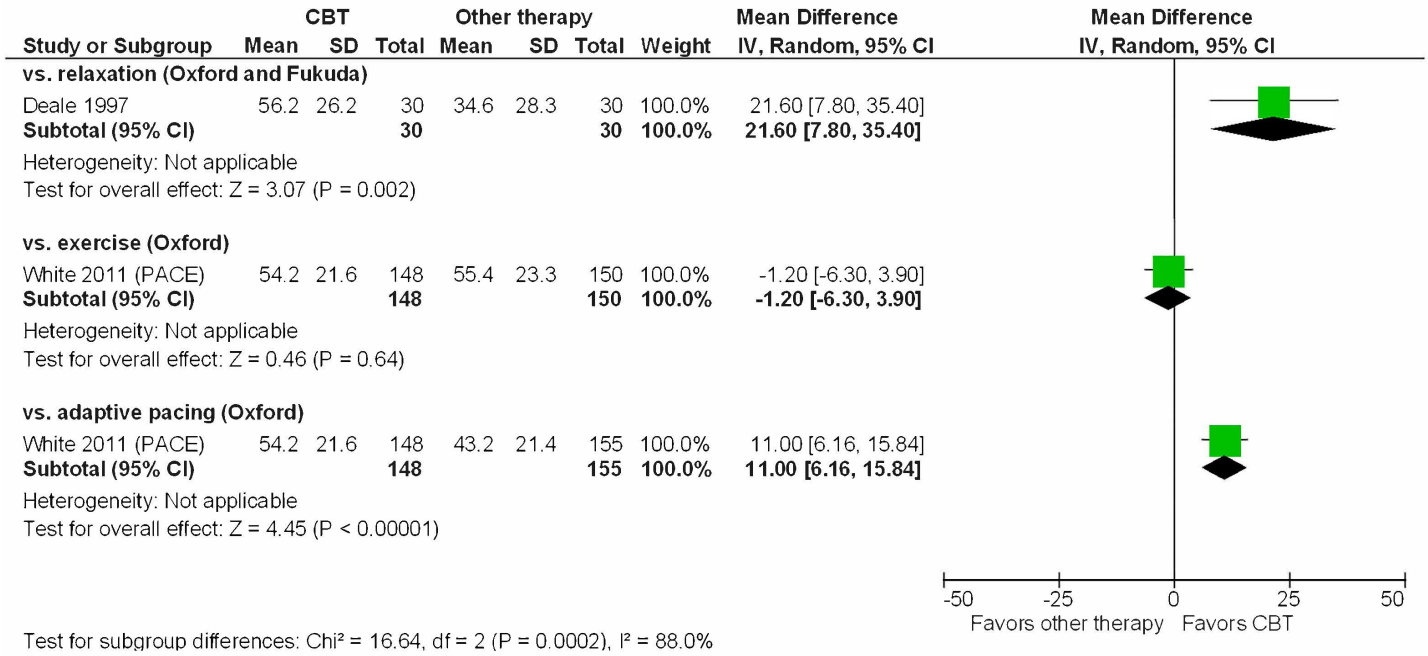
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 42. Fatigue severity: CBT versus active interventions at post-intervention follow-up



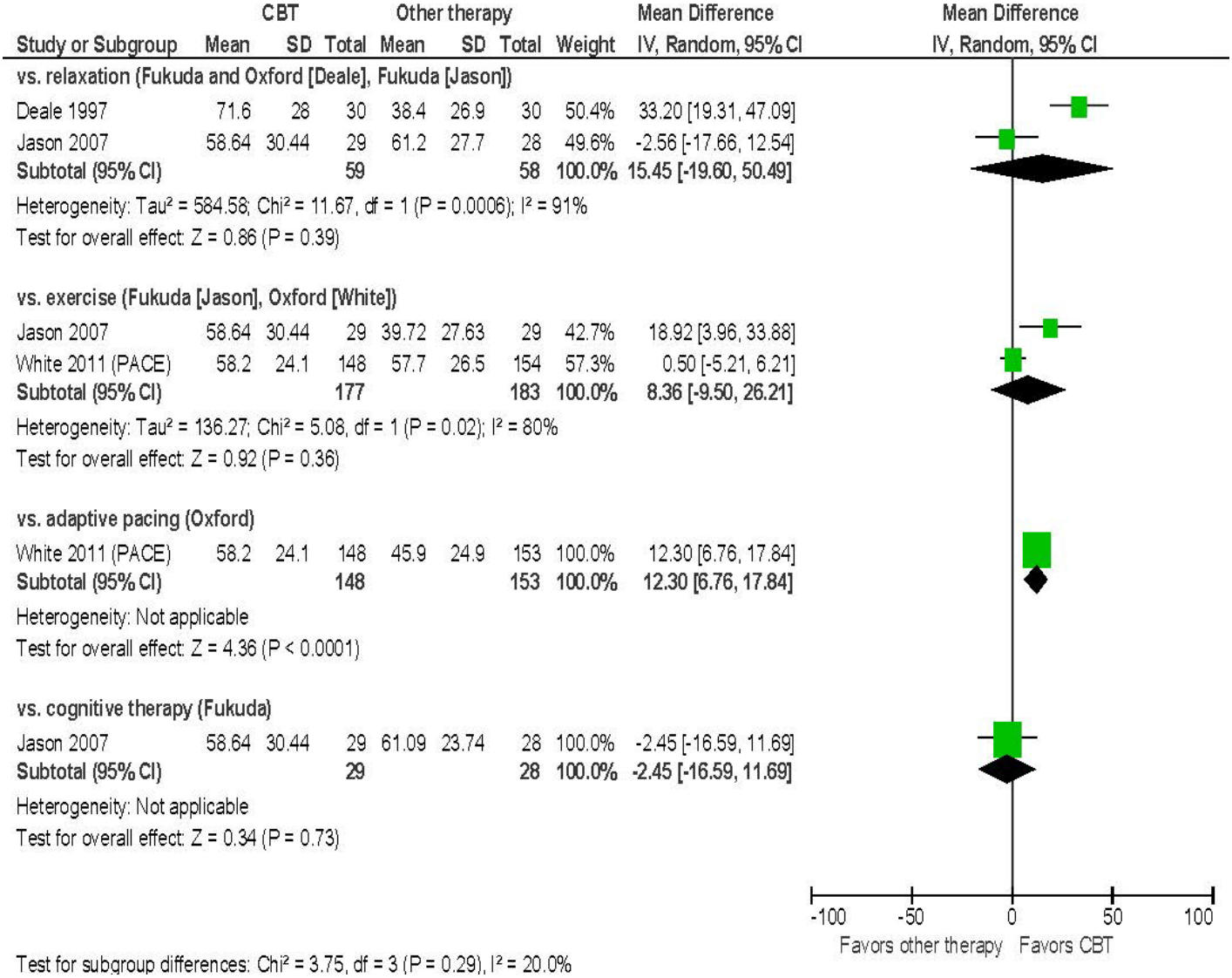
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation; Std. = standard

Figure 43. Functional impairment: CBT versus active interventions at end of treatment



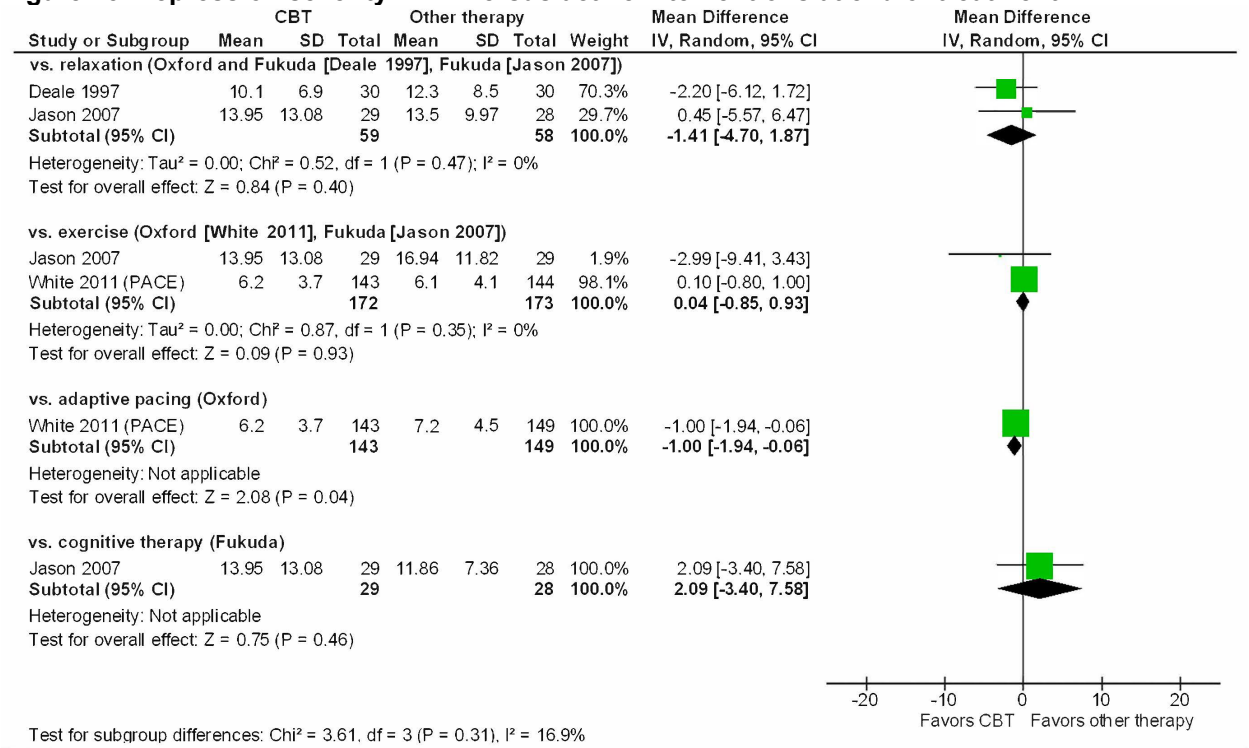
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 44. Functional impairment: CBT versus active interventions at post-intervention follow-up



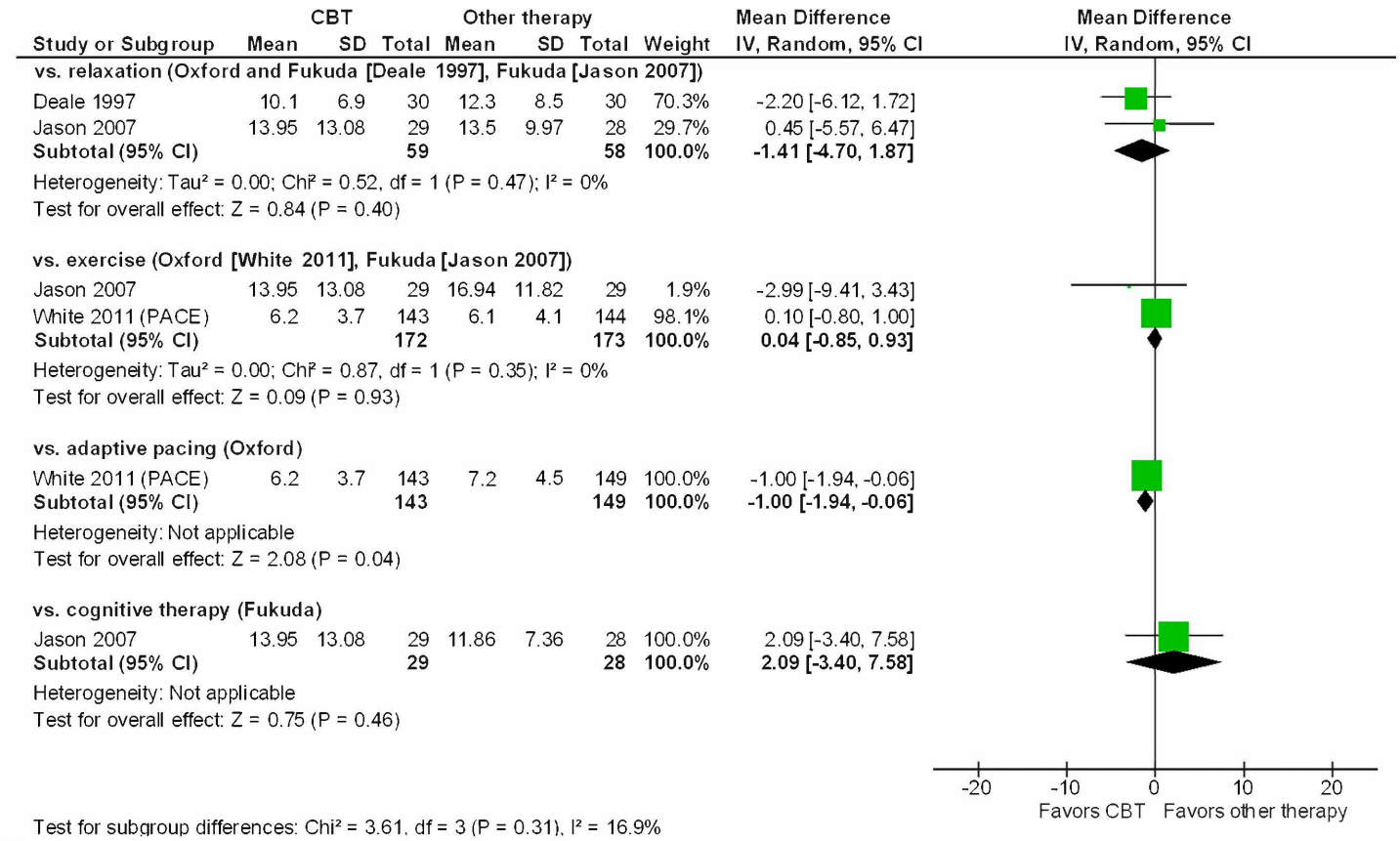
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 45. Depression severity: CBT versus active interventions at end of treatment



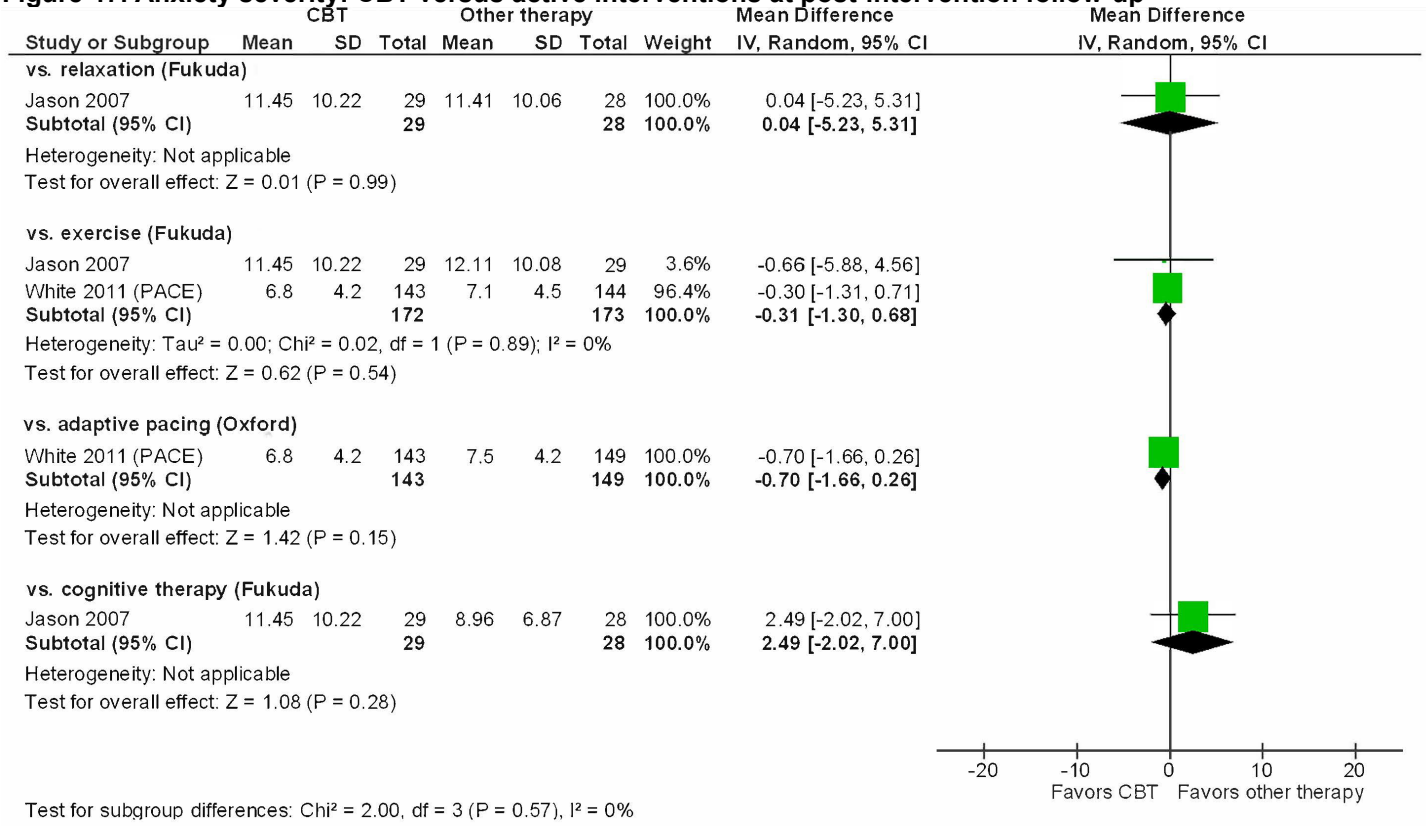
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; IV = instrumental variable; SD = standard deviation

Figure 46. Depression severity: CBT versus active interventions at post-intervention follow-up



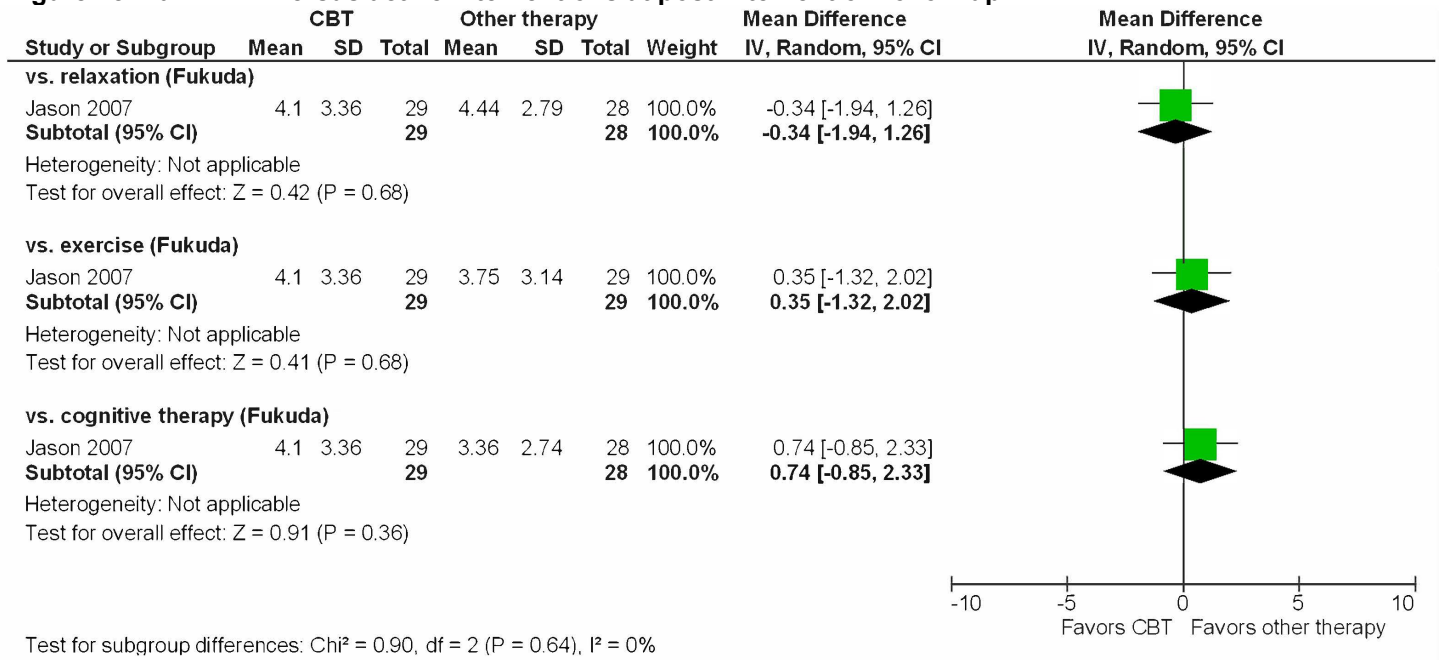
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 47. Anxiety severity: CBT versus active interventions at post-intervention follow-up



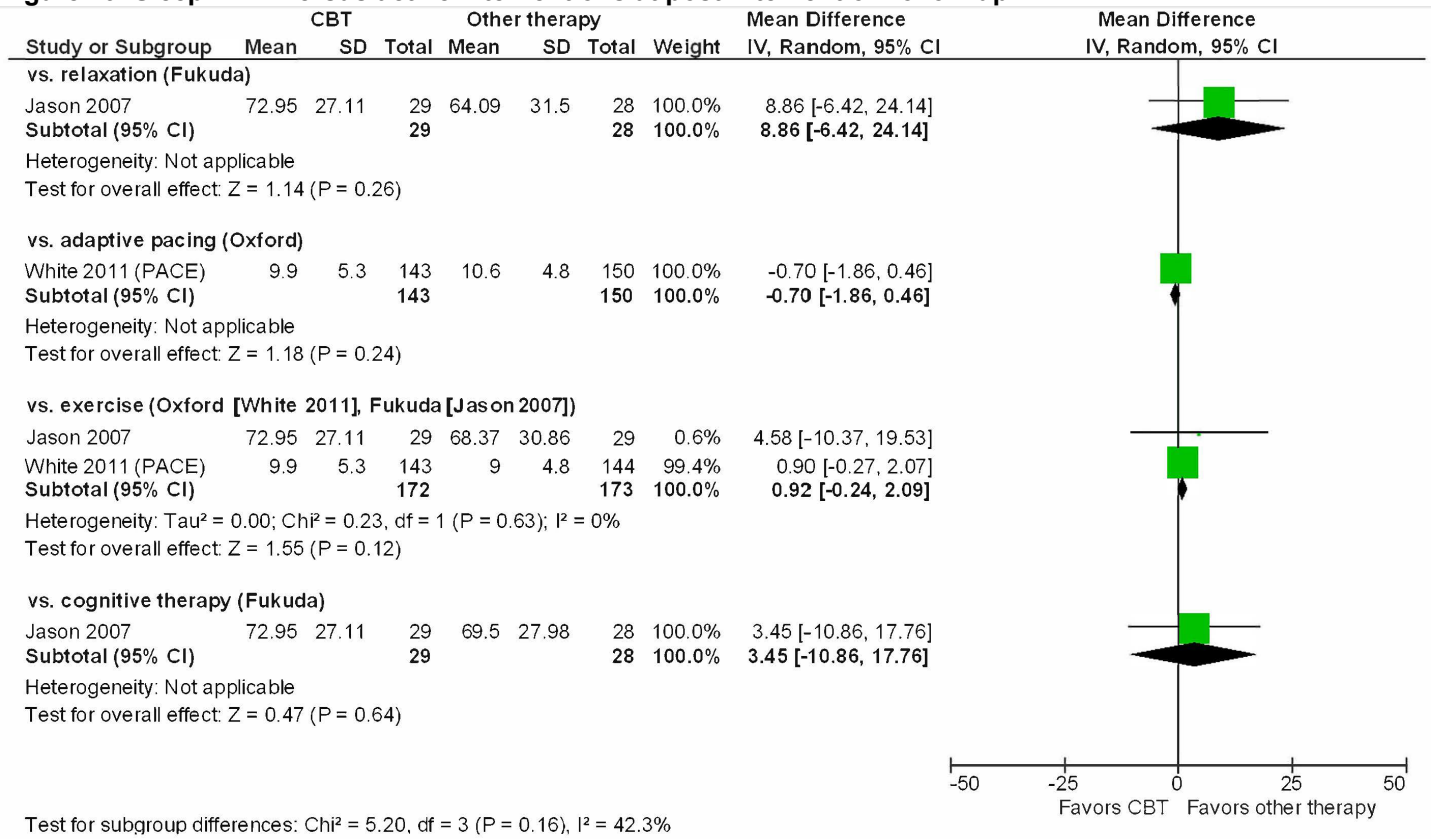
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 48. Pain: CBT versus active interventions at post-intervention follow-up



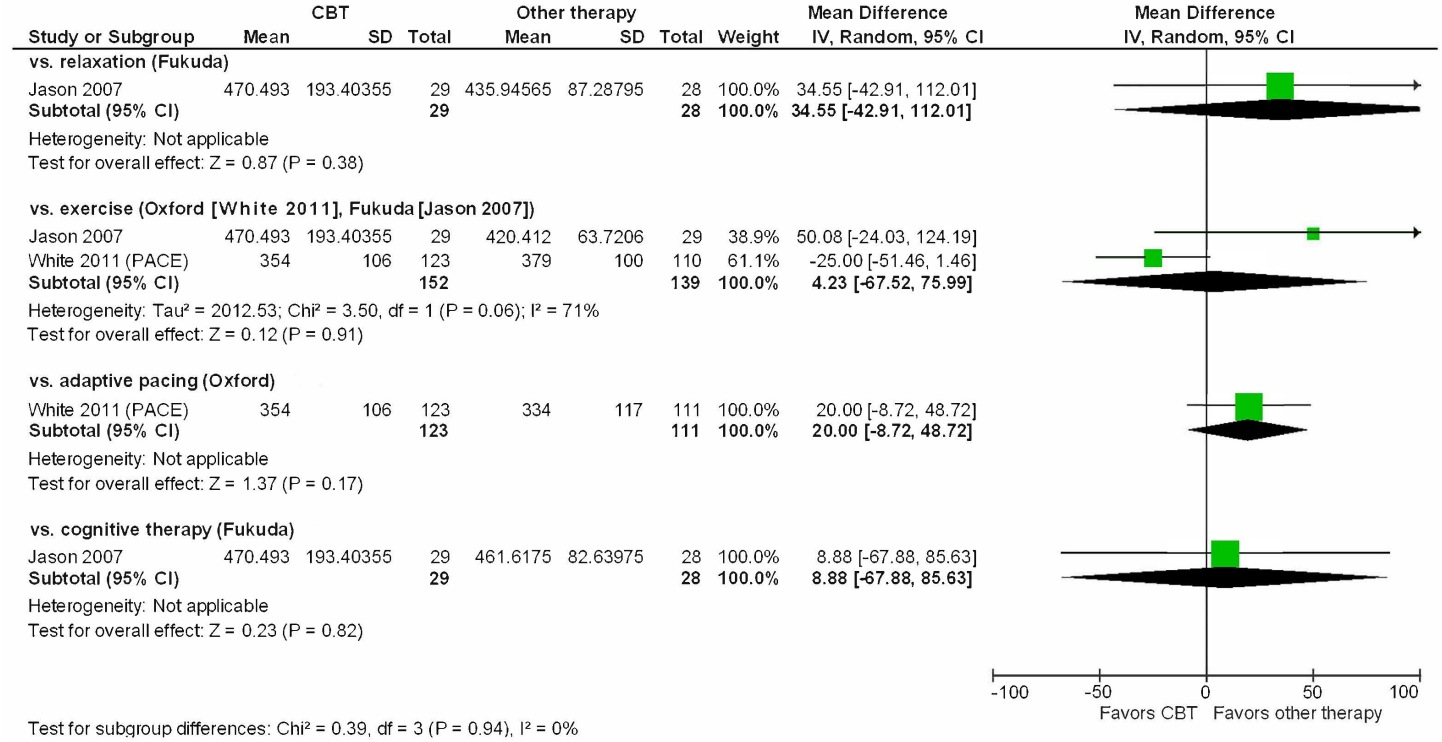
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; IV = instrumental variable; SD = standard deviation

Figure 49. Sleep: CBT versus active interventions at post-intervention follow-up



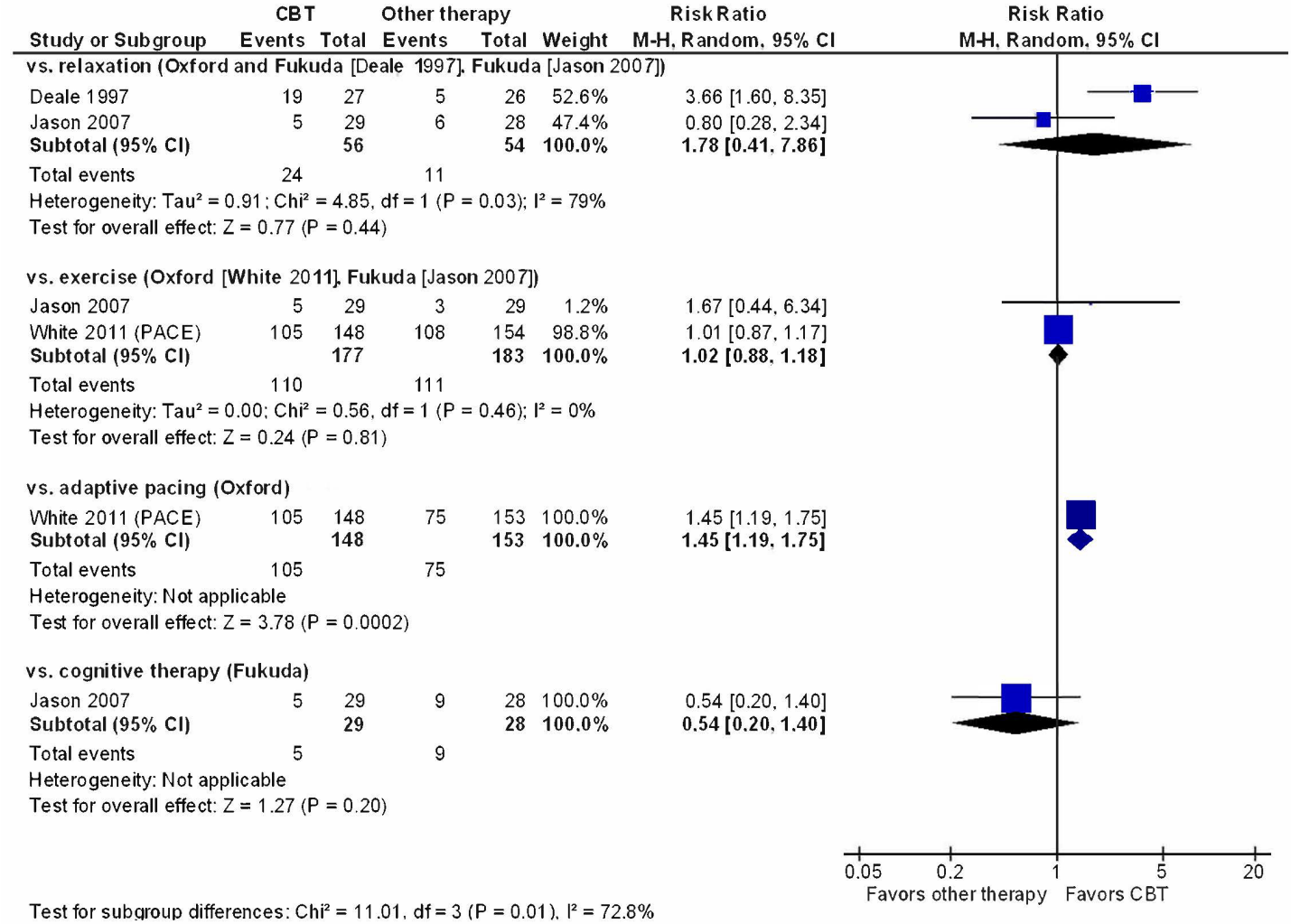
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 50. 6-minute walk test: CBT versus active interventions at end of treatment



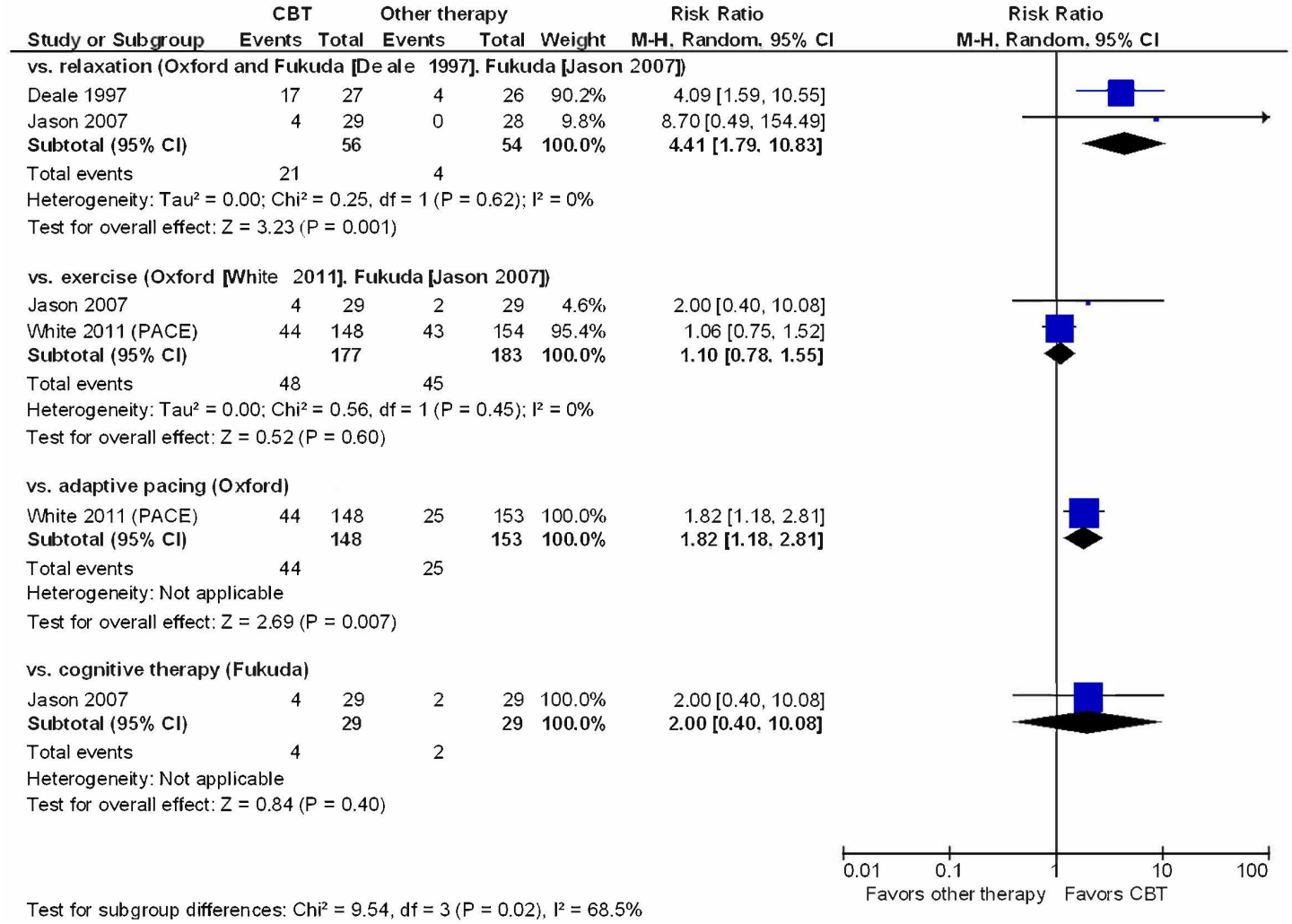
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation; IV = instrumental variable; SD = standard deviation

Figure 51. Likelihood of functional improvement: CBT versus active interventions



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 52. Likelihood of recovery: CBT versus active interventions



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

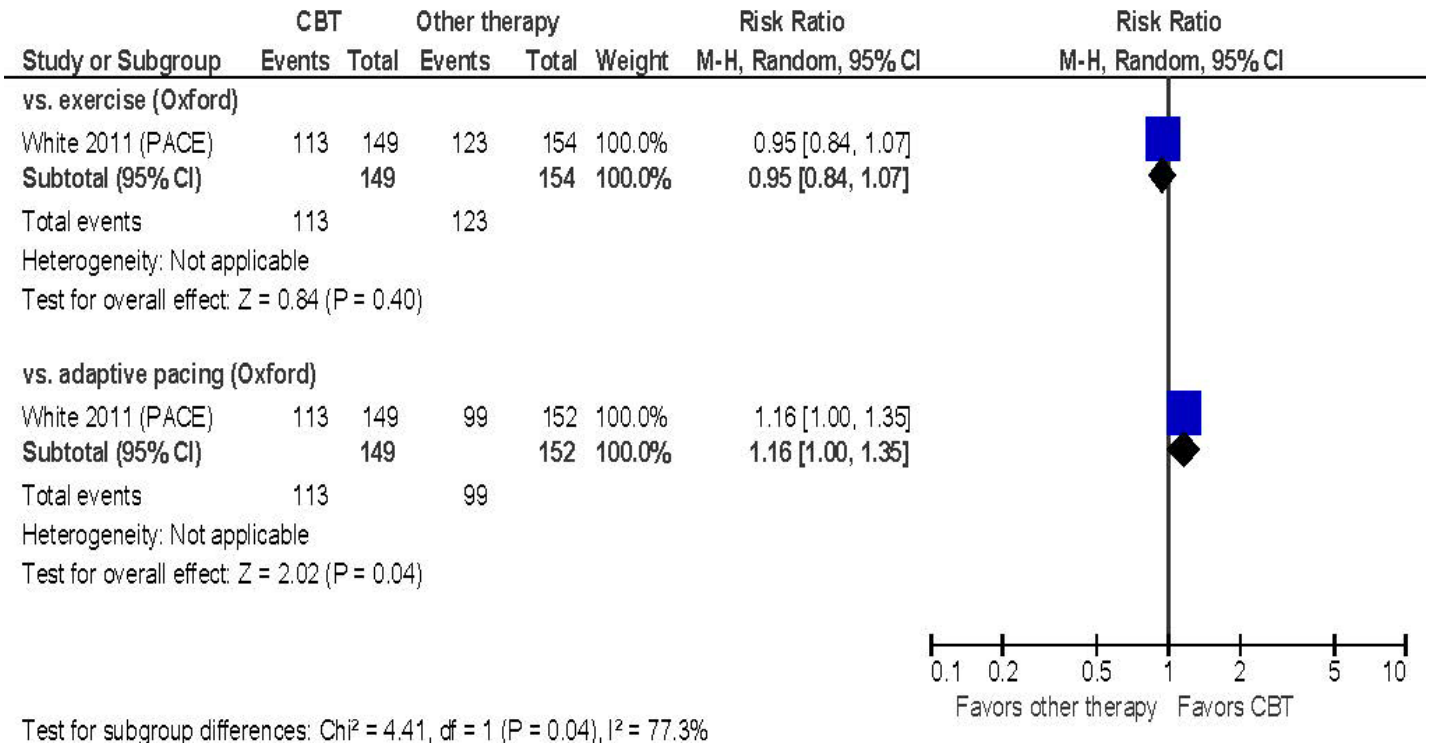
CBT versus adaptive pacing

The PACE trial included a comparison of CBT versus adaptive pacing (n=320). Outcomes were reported at the end of therapy at 24 weeks and 28 weeks later.[38] CBT was associated with decreased fatigue severity versus adaptive pacing at the end of the intervention (mean difference -2.2, 95% CI -3.9 to -0.5 on the 11-item 0 to 33 Chalder scale, **Figure 41**) and at post-intervention follow-up (mean difference -2.8, 95% CI -4.5 to -1.1, **Figure 42**). CBT was also associated with decreased severity of functional impairment at the end of the intervention (mean difference 11.00, 95% CI 6.16 to 15.84 on the SF-36 physical function subscale, **Figure 43**) and at post-intervention follow-up (mean difference 12.30, 95% CI 6.76 to 17.84, **Figure 44**), as well as decreased depression severity at post-intervention follow-up (mean difference -1.0, 95% CI -2.0 to -0.05 on the 0 to 21 HADS depression scale, **Figure 46**). Effects on sleep quality (mean difference -0.7, 95% CI -1.9 to 0.5 on the 0 to 20 Jenkins sleep scale, **Figure 49**), the 6-minute walk test (mean difference 20 meters, 95% CI -9 to 49, **Figure 50**) and anxiety (mean difference -0.77, 95% CI -1.66 to 0.26 on the 0 to 21 HADS anxiety scale, **Figure 47**) favored CBT at post-intervention follow-up, but differences were small and not statistically significant. In a post-hoc analysis, CBT was associated with decreased severity of muscle pain (mean difference -0.34, 95% CI -0.65 to -0.02) and joint pain (mean difference -0.35, 95% CI -0.68 to -0.02) versus specialist care (each assessed on a 0 to 4 scale).[64]

CBT was associated with increased likelihood of improvement in fatigue (76% vs. 65%, RR 1.16, 95% CI 1.00 to 1.35, **Figure 53**), improvement in function (71% vs. 49%, RR 1.45, 95% CI 1.19 to 1.75, **Figure 51**), and recovery (30% vs. 16%, RR 1.82, 95% CI 1.18 to 2.81, **Figure 52**) versus adaptive pacing, based on the definitions in the main PACE publication. Results were similar using the original protocol definitions[39] for improvement in fatigue (26% vs. 14%, RR 1.80, 95% CI 1.14 to 2.85), functional improvement (49% vs. 40%, RR 1.22, 95% CI 0.95 to 1.56), and overall improvement (composite of improvement in fatigue and function, 20% vs. 9%, RR 2.11, 95% CI 1.19 to 3.74),[40] though the estimate for functional improvement was attenuated and no longer statistically significant. CBT was also associated with lower likelihood of post-exertional malaise (49% vs. 63%, RR 0.78, 95% CI 0.64 to 0.95).[38] There were no differences between CBT versus adaptive pacing in risk of serious adverse events **(Figure 54)**, withdrawal due to adverse events **(Figure 55)**, or worsening of function **(Figure 56)**, but estimates were imprecise (**Table 10**).

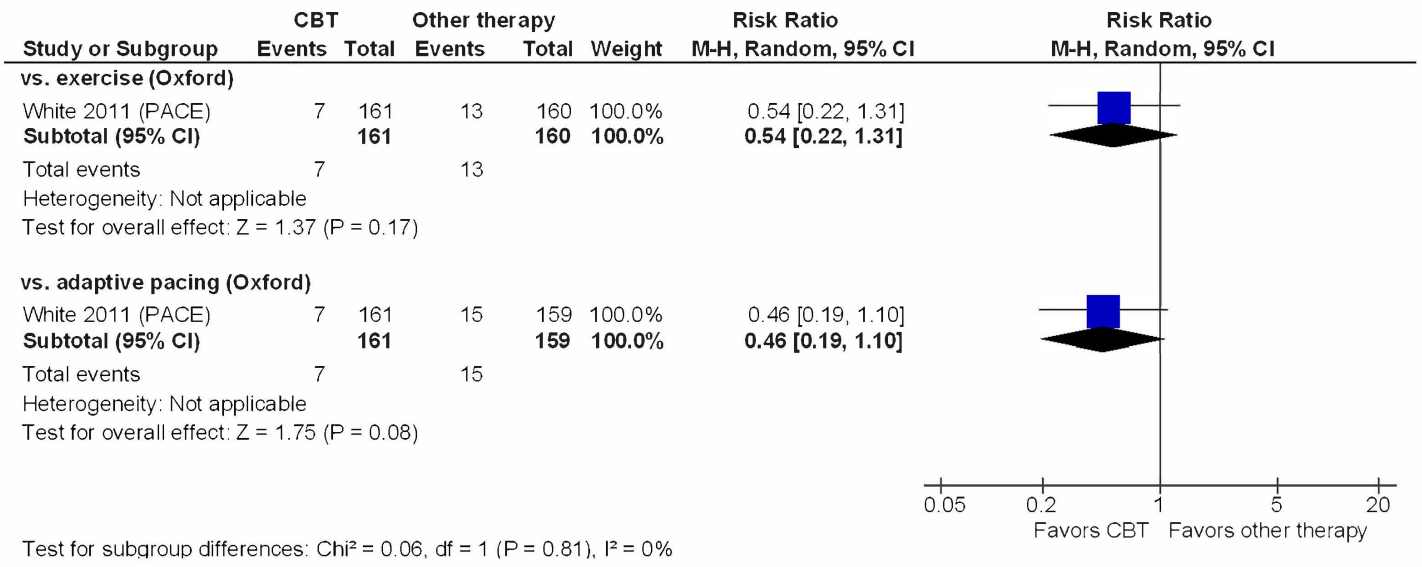
The PACE trial also evaluated longer-term, post-trial outcomes (median duration from randomization 31 months).[63] About 31% of patients in the CBT group and 50% in the adaptive pacing group received non-randomly allocated therapies between the end of the trial at 1 year and long-term follow-up. Fatigue severity was decreased at long-term follow-up compared with end-of-trial scores in the CBT (mean change -2.2, 95% CI -3.7 to -0.6 on the 11-item, 0 to 33 Chalder scale) and adaptive pacing (mean change -3.0, 95% CI -4.4 to -1.6) groups. SF-36 physical function also improved between the end of the trial and long-term follow-up in the CBT (mean change 3.3, 95% CI 0.02 to 6.7 on a 0 to 100 scale) and adaptive pacing groups (mean change 8.5, 95% CI 4.5 to 12.5), though the change was smaller in the CBT group. At long-term follow-up, mixed model analysis showed no difference between CBT versus adaptive pacing in fatigue severity (mean difference -1.6, 95% CI -3.6 to 0.3), but CBT was associated with decreased functional impairment (mean difference 6.4, 95% CI 0.4 to 12.4).

Figure 53. Likelihood of fatigue improvement: CBT versus active interventions



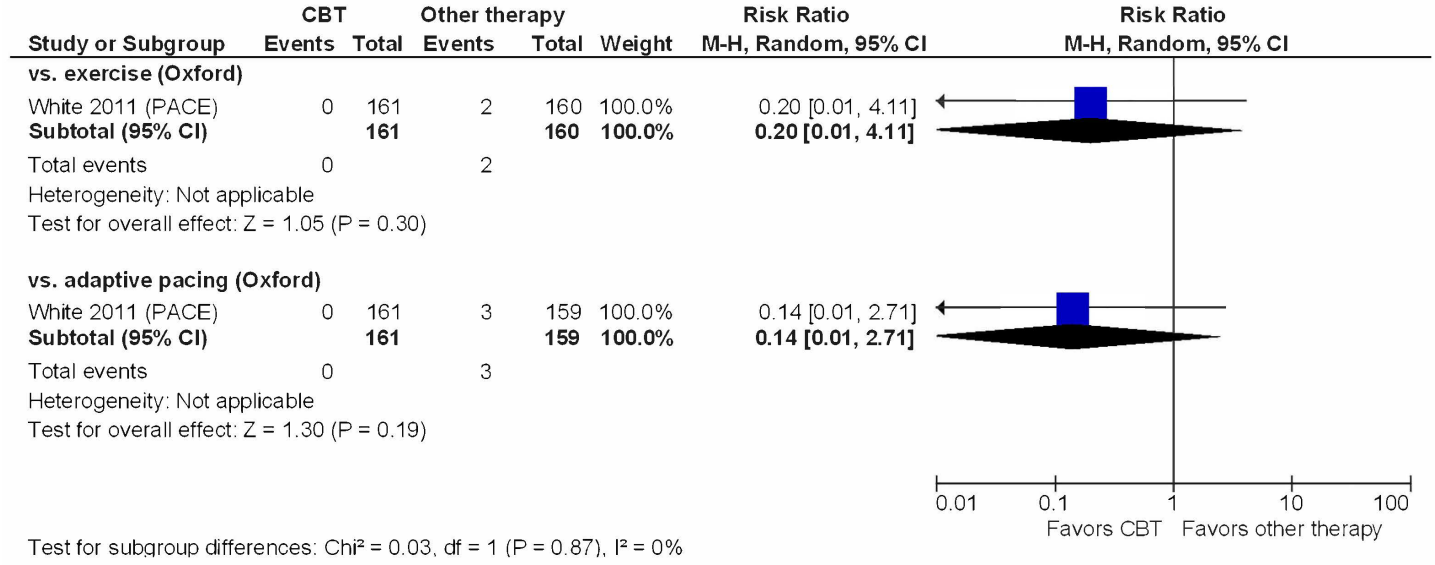
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 54. Likelihood of serious adverse event: CBT versus active interventions



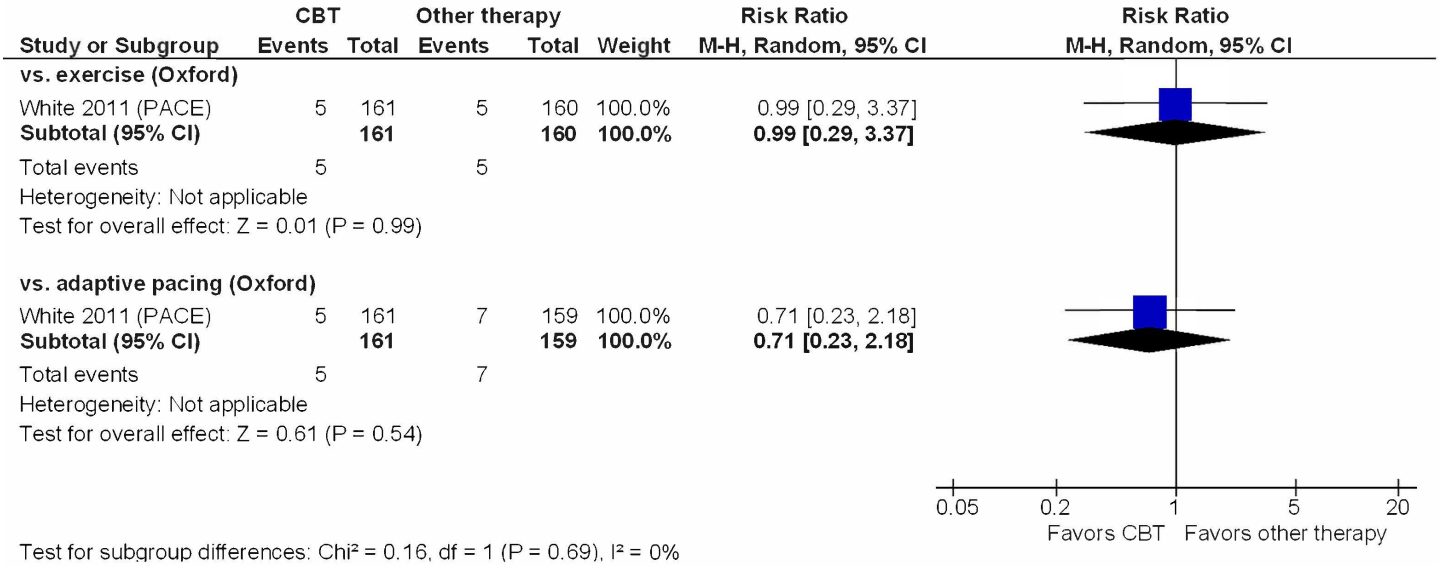
Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 55. Likelihood of withdrawal due to adverse event: CBT versus active interventions



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

Figure 56. Likelihood of worsening function: CBT versus active interventions



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; M-H = Mantel-Haenszel test; PACE = pacing, graded activity, cognitive behavior therapy; a randomized Evaluation

CBT versus cognitive therapy

One trial included a comparison of CBT versus cognitive therapy (N=57).[54] Outcomes were evaluated 6 months following completion of 6 months of treatment. There were no differences between CBT versus cognitive therapy in severity of fatigue **(Figure 42)**, functional impairment **(Figure 44)**, depression **(Figure 46)**, anxiety **(Figure 47)** or pain **(Figure 48)**; sleep quality **(Figure 49)**; or distance on the 6-minute walk test (**Figure 50, Table 10**). CBT was associated with higher severity of sore throat (mean difference 16.4, 95% CI 1.5 to 31.3 on the 0 to 100

CFS Questionnaire) and tender lymph node symptoms (mean difference 15.9, 95% CI 0.3 to 31.5). There were no differences in severity of post-exertional malaise (mean difference 1.5, 95% CI -17.2 to 20.3), impaired memory (mean difference -3.1, 95% CI -17.3 to 11.1), or headache (mean difference 11.6, 95% CI -8.6 to 32.0) symptoms. There was also no difference in likelihood of functional improvement **(Figure 51)** or recovery **(Figure 52)**, but estimates were imprecise.

CBT versus medication (mirtazapine)

One trial (N=48) compared CBT versus mirtazapine (an antidepressant).[75] Outcomes were evaluated at the completion of 12 weeks of treatment. Effects on fatigue severity favored mirtazapine, but the difference was not statistically significant (mean difference -1.7, 95% CI -4.8 to 1.4 on the 11-item, 0 to 33 Chalder scale). There was no difference between CBT versus mirtazapine in depression severity (mean difference 0.30, 95% CI -2.6 to 3.2 on the 0 to 52 Hamilton Rating Scale for Depression). The trial did not evaluate harms.

Table 10. CBT versus active interventions: summary of stratified results

Outcome	Number of studies (N)	Estimate (95% CI)	I^2
Fatigue, end of intervention			
CBT vs. exercise	1 (298)	SMD -0.03 (-0.25 to 0.20)	--
CBT vs. relaxation	1 (60)	SMD -0.07 (-0.58 to 0.43)	--
CBT vs. adaptive pacing	1 (303)	SMD -0.30 (-0.52 to -0.07)	--
CBT vs. mirtazapine	1 (48)	SMD -0.32 (-0.89 to 0.25)	--
Fatigue, post-intervention			
CBT vs. exercise	2 (360)	SMD -0.08 (-0.29 to 0.13)	0%
CBT vs. relaxation	2 (117)	SMD -0.49 (-1.03 to 0.04)	52%
CBT vs. cognitive therapy	1 (57)	SMD -0.45 (-0.97 to 0.08)	--
CBT vs. adaptive pacing	1 (301)	SMD -0.36 (-0.59 to -0.14)	--
SF-36 physical function (0 to 100), end of intervention			
CBT vs. exercise	1 (298)	MD -1.20 (-6.30 to 3.90)	--
CBT vs. relaxation	1 (60)	MD 21.60 (7.80 to 35.40)	-
CBT vs. adaptive pacing	1 (303)	MD 11.00 (6.16 to 15.84)	--
SF-36 physical function (0 to 100), post-intervention			
CBT vs. exercise	2 (360)	MD 8.36 (-9.50 to 26.21)	80%
CBT vs. relaxation	2 (117)	MD 15.45 (-19.60 to 50.49)	91%
CBT vs. cognitive therapy	1 (57)	MD -2.45 (-16.59 to 11.69)	--
CBT vs. adaptive pacing	1 (301)	MD 12.30 (6.76 to 17.84)	--
Depression, end of intervention			
CBT vs. relaxation	1 (60)	SMD -0.45 (-0.96 to 0.06)	--
CBT vs. mirtazapine	1 (48)	SMD 0.06 (-0.51 to 0.63)	--
Depression, post-intervention			
CBT vs. exercise	2 (345)	SMD -0.02 (-0.23 to 0.19)	0%
CBT vs. relaxation	2 (117)	SMD -0.12 (-0.49 to 0.24)	0%
CBT vs. cognitive therapy	1 (57)	SMD 0.19 (-0.33 to 0.71)	--
CBT vs. adaptive pacing	1 (292)	SMD -0.24 (-0.47 to -0.01)	--
Anxiety, post intervention			
CBT vs. exercise	2 (345)	SMD 0.07 (-0.28 to 0.14)	0%
CBT vs. relaxation	1 (57)	SMD 0.00 (-0.52 to 0.52)	--
CBT vs. cognitive therapy	1 (57)	SMD 0.28 (-0.24 to 0.80)	--
CBT vs. adaptive pacing	1 (292)	SMD -0.17 (-0.40 to 0.06)	--
Sleep, post intervention			
CBT vs. exercise	2 (345)	SMD -0.17 (-0.04 to 0.39)	0%
CBT vs. relaxation	1 (57)	SMD 0.30 (-0.22 to 0.82)	--
CBT vs. cognitive therapy	1 (57)	SMD 0.12 (-0.40 to 0.64)	--
CBT vs. adaptive pacing	1 (293)	SMD -0.14 (-0.37 to 0.09)	--
Brief Pain Inventory (0 to 10), post intervention			

Outcome	Number of studies (N)	Estimate (95% CI)	I^2
CBT vs. exercise	1 (58)	MD 0.35 (-1.32 to 2.02)	--
CBT vs. relaxation	1 (57)	MD -0.34 (-1.94 to 1.26)	--
CBT vs. cognitive therapy	1 (57)	MD 0.74 (-0.85 to 2.33)	--
6-minute walk test (meters), end of intervention			
CBT vs. exercise	2 (291)	MD 4.23 (-67.52 to 75.99)	71%
CBT vs. relaxation	1 (57)	MD 34.55 (-42.91 to 112.01)	--
CBT vs. cognitive therapy	1 (57)	MD 8.88 (-67.88 to 85.63)	--
CBT vs. adaptive pacing	1 (234)	MD 20.00 (-8.72 to 48.72)	--
Recovery			
CBT vs. exercise	2 (360)	RR 1.10 (0.78 to 1.55)	0%
CBT vs. relaxation	2 (110)	RR 4.41 (1.79 to 10.83)	0%
CBT vs. cognitive therapy	1 (58)	RR 2.00 (0.40 to 10.08)	--
CBT vs. adaptive pacing	1 (301)	RR 1.82 (1.18 to 2.81)	--
Fatigue improvement			
CBT vs. exercise	1 (303)	RR 0.95 (0.84 to 1.07)	--
CBT vs. adaptive pacing	1 (301)	RR 1.16 (1.00 to 1.35)	--
Functional improvement			
CBT vs. exercise	2 (360)	RR 1.02 (0.88 to 1.18)	0%
CBT vs. relaxation	2 (110)	RR 1.78 (0.41 to 7.86)	79%
CBT vs. cognitive therapy	1 (57)	RR 0.54 (0.20 to 1.40)	--
CBT vs. adaptive pacing	1 (301)	RR 1.45 (1.19 to 1.75)	--
Serious adverse events			
CBT vs. exercise	1 (321)	RR 0.54 (0.22 to 1.31)	--
CBT vs. adaptive pacing	1 (320)	RR 0.46 (0.19 to 1.10)	--
Withdrawal due to worsening			
CBT vs. exercise	1 (321)	RR 0.20 (0.01 to 4.11)	--
CBT vs. adaptive pacing	1 (320)	RR 0.14 (0.01 to 2.71)	--
Physical function worsening			
CBT vs. exercise	1 (321)	RR 0.99 (0.29 to 3.37)	--
CBT vs. adaptive pacing	1 (320)	RR 0.71 (0.23 to 2.18)	--
Post-exertional malaise			
CBT vs. exercise	1 (321)	RR 1.10 (0.87 to 1.40)	--
CBT vs. adaptive pacing	1 (319)	RR 0.78 (0.64 to 0.95)	--

Abbreviations: CI = confidence interval; CBT = cognitive behavioral therapy; MD = mean difference; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference

One Method of CBT Delivery Versus Another

Face-to-face versus telephone CBT

One trial (N=58) compared face-to-face versus telephone CBT (with face-to-face assessment and discharge appointment).[68] The duration of therapy was 6 months and outcomes were evaluated at the end of the intervention though 12 months following the completion of therapy. The trial was rated medium risk of bias. There were no differences between face-to-face versus telephone CBT in fatigue or function at the end of the intervention or at post-intervention follow-up or in likelihood of a global improvement rating of much better or very much better.

Other Behavioral Approaches in Adults

Illness management and peer counseling versus wait list

One medium risk of bias trial (N=47) included in the prior report compared an illness management and peer counseling intervention (8 biweekly 2-hour group sessions over 4 months, 1 month break, and 7 months of one-on-one peer counseling) versus wait list in patients who met

the Fukuda case definition (**Tables 11 and 12, Evidence Table Appendix E2, Risk of Bias Table Appendix F**).[79] The intervention involved 8 biweekly 2-hour group sessions over 4 months, followed by a 1 month break, then 7 months of individual peer counseling. At baseline, the mean Chronic Fatigue Syndrome Symptoms Rating Form score was 14.6 on a 0 to 100 scale. At the end of the intervention, there was no difference in severity of symptoms (mean difference -0.9, 95% CI -2.8 to 1.0 on the 0 to 100 Chronic Fatigue Syndrome Symptoms Rating Form), the Quality of Life Index overall quality of life scale (mean difference 1.1, 95% CI -1.2 to 3.4 on a 0 to 30 scale) or the Quality of Life Index health and functioning, social and economic, psychological and spiritual, and family subscales (differences ranged from 0.1 to 0.5, each on a 0 to 30 scale). The trial did not report harms.

Mindfulness-based cognitive therapy versus usual care or wait list

Two small trials (N=18 and 35) evaluated mindfulness-based cognitive therapy versus usual care or wait list in patients who met the Fukuda case definition or either the Fukuda or Oxford case definition[80,81] (**Tables 11 and 12, Evidence Table Appendix E2, Risk of Bias Table Appendix F**). Neither trial was included in the prior AHRQ report. At baseline, mean scores on the 11-item 0 to 33 Chalder scale were 24.3 and 23.4 and on SF-36 physical function were 41.5 and 58.3. Mindfulness training occurred in weekly group classes for 8 weeks in both trials; in one trial[80] there was 1 follow-up class at 4 months. Both trials were rated high risk of bias (**Risk of Bias Table Appendix F**).

For outcomes assessed at the end of the intervention, both trials found mindfulness training associated with reduced fatigue severity, though the difference was only statistically significant in one trial (mean differences -1.82, p=0.08 and -3.9, p=0.01 on the 11-item 0 to 33 Chalder scale). There were no statistically significant effects on SF-36 physical function (mean differences 3.50, p=0.58 and 12.2, p=0.12 on a 0 to 100 scale). Both trials found mindfulness-based cognitive therapy associated with decreased depression severity versus wait list or usual care, but the difference was statistically significant in only one trial (mean differences -1.17, p=0.28 and -3.6, p=0.04 on the 0 to 21 HADS depression scale). Although one trial found mindfulness-based cognitive therapy associated with decreased anxiety severity, the difference was small (mean differences -1.6, p=0.17 and -0.41, p=0.01 on the 0 to 21 HADS anxiety scale).

One of the trials also evaluated outcomes 2 months following the completion of therapy.[80] Effects of mindfulness based cognitive therapy persisted at post-intervention follow-up (mean difference -3.7, p=0.03). There were no statistically significant differences between mindfulness-based cognitive therapy versus wait list in function, depression, or anxiety.

One trial reported that there were "no substantive adverse events" and the other trial did not report harms.

Self-management versus usual care

Two trials (N=124 and 125) compared self-management interventions versus usual care in patients with ME/CFS[82,83] (**Tables 11 and 12, Evidence Table Appendix E2**). Neither trial was included in the prior AHRQ report. In one trial,[82] the self-management intervention was delivered using a booklet and audio compact discs (CDs) and in the other trial[83] the self-management intervention was conducted by a peer counselor and occupational therapist in group sessions. Both trials were rated medium risk of bias (**Risk of Bias Table Appendix F**).

One trial (N=124) compared two self-management interventions versus usual care in patients who met the Fukuda case definition.[82] At baseline, mean scores were 6.5 on the Fatigue Severity

Scale (1 to 7 scale) and 37.9 for SF-36 physical function (0 to 100 scale). Both self-management interventions were delivered using a booklet and audio CDs and differed in the way that fatigue and compliance was monitored (web diaries and actigraphy [activity monitor] versus paper diaries and a step counter). Outcomes were assessed at the end of 3 months of therapy and at 12 months. Effects of the two self-management interventions were similar and we combined the results. Self-management was associated with decreased fatigue severity versus usual care at the end of therapy (mean difference -0.40, 95% CI -0.66 to -0.14) and at post-intervention follow-up (mean difference -0.37, 95% CI -0.66 to -0.08). Self-management was also associated with increased likelihood of improvement in fatigue, defined as Fatigue Severity Scale score at 12 months >2 standard deviations below the baseline sample mean (26% vs. 8.7%, RR 2.98, 95% CI 1.09 to 8.15). There was no difference in function at end of therapy (mean difference 6.24, 95% CI -1.58 to 14.06 on SF-36 physical function) or at post-intervention follow-up (mean difference 2.06, 95% CI -6.62 to 10.74). The self-management interventions were associated with reduced depression severity (differences of 4.7 to 4.9 points on the 0 to 63 Beck Depression Inventory) but no difference in anxiety severity. There was no difference in the likelihood of post-exertional malaise after more than 24 hours (33% vs. 34%, RR 0.96, 95% CI 0.66 to 1.41) or in the 6-minute walk test (data not provided).

The other trial (N=125) compared a group self-management intervention (8 biweekly, 2.5 hour sessions) versus usual care in patients who met the Fukuda and 2003 Canadian case definitions.[83] At 6 months follow-up (2 months after completion of therapy) self-management was associated with less improvement in fatigue severity versus usual care (difference in change from baseline 2.5, p=0.04), with no differences between groups in SF-36 physical function, the SF-36 physical component summary, or the SF-36 mental component summary. At 12 months, there were no differences in fatigue or SF-36 measures at 12 months follow-up. The trial did not report harms.

Table 11. RCTs of behavioral approaches in adults: study characteristics

Author, year Country Risk of Bias	Study n (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
Friedberg, 2016[82] United States Medium	n: 124 Age: 48.4 % Female: 88	Criteria: Fukuda Duration: 14.5 years	Fatigue Scale: Fatigue Severity Scale 9-item (1 to 7) Baseline: 6.5 (SD 0.48) Post-exertional fatigue or malaise: not reported	Major depression: Current or past depression with melancholic or psychotic features excluded Baseline depression: Beck Depression Inventory (0 to 63): 19.2 (SD 10.8) Baseline function: SF-36 physical function (0 to 100): 40.6 (SD 20.8)	A: Self-management with web diaries and actigraphs B: Self-management with paper diaries and step counters C: Usual care Frequency: not described Session length: not described Duration of treatment: 12 weeks Duration of follow-up: 12 months
Pinxterhuis, 2017[83] Norway Medium	n: 125 Age: 43.9 % Female: 86	Criteria: Fukuda or Canadian Criteria (2003) Duration: Median 3 years	Fatigue Scale: Fatigue Severity Scale (9 to 63) Baseline: 57.3 (SD 5.1) Post-exertional fatigue or malaise: Not reported	Major depression: Not reported Baseline depression: Not reported Baseline function: SF-36 physical function (0 to 100): 46.0 (SD 19.2)	A: Group based self-management B: Usual care Frequency: 8 sessions every other week Session length: 2.5 hours Duration of treatment: 16 weeks Duration of follow-up: 1 year
Rimes, 2013[80] United Kingdom High	n: 35 Age: 43.5 % Female: 83	Criteria: Fukuda or Oxford Duration: Mean 7.2 years	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 24.3 (SD 4.5) Post-exertional fatigue or malaise: Not reported	Major depression: Excluded for current major depression (29% on antidepressants at baseline) Baseline depression: HADS depression (0 to 21): 7.3 (SD 4.5) Baseline function: SF-36 physical function (0 to 100): 58.3 (SD 23.2)	A: Mindfulness-based cognitive therapy B: Wait list Frequency: 8 weekly sessions and 1 follow-up class at 4 months Session length: 2.25 hours Duration of treatment: 2 months Duration of follow-up: 4 months
Surawy, 2005[81] United Kingdom High	n: 17 Age: Not reported (range 18 to 65) % Female: 56	Criteria: Fukuda Duration: Not reported	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 23.4 (SD 7.8) Post-exertional fatigue or malaise: Not reported	Major depression: Excluded Baseline depression: HADS depression (0 to 21): 9.7 (SD 4.0) Baseline function: SF-36 physical function (0 to 100): 41,5 (SD 24.8)	A: Group mindfulness training B: Usual care Frequency: 8 sessions over 8 weeks Session length: not described Duration of treatment: 8 weeks Duration of follow-up: 8 weeks

Author, year Country Risk of Bias	Study n (analyzed) Age, mean years % Female	ME/CFS criterion ME/CFS duration	Fatigue Scale Baseline fatigue	Baseline Depression Baseline Function	Intervention Frequency, duration, and intensity Duration of treatment Duration of follow-up
Taylor, 2004[79] United States Medium	n: 47 Age: Not reported % Female: 96	Criteria: Fukuda Duration: Not reported	Fatigue Scale: Chronic Fatigue Syndrome Symptom Rating Scale (0 to 100) Baseline: 14.6 (SD 2.9) Post-exertional fatigue or malaise: Not reported	Major depression: Not reported Baseline depression: Not reported Baseline function: Not reported	A: Illness management and peer counseling B: Wait list Frequency: Biweekly illness management sessions over 4 months, then 7 months of peer counseling (frequency not reported) Session length: 2 hours for illness management sessions, not described for peer counseling Duration of treatment: 11 months Duration of follow-up: 11 months

Abbreviations: CFS = chronic fatigue syndrome; HADS = Hospital Anxiety and Depression Scale; ME = myalgic encephalomyelitis; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey

Table 12. RCTs of behavioral approaches in adults: study results

Author, Year ME/CFS Criterion	Intervention A: intervention (n) B: control (n) Duration of treatment Duration of follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Friedberg, 2016[82] Fukuda	A: Self-management with web diaries and actigraphs (45) B: Self-management with paper diaries and step counters (44) C: Usual care (48) Duration of treatment: 12 weeks Duration of follow-up: 12 months	Fatigue Severity Scale 9-item (1 to 7), mean (SE): 3 months: 6.12 (0.11) vs. 5.92 (0.11) vs. 6.42 (0.10), FSM:ACT vs. FSM:CTR p<0.05, other comparisons p>0.05 12 months: 6.00 (0.13) vs. 6.10 (0.13) vs. 6.42 (0.12), all comparisons p>0.05	Beck Depression Inventory (0 to 63), mean (SE): 3 months: 14.40 (1.65) vs. 14.98 (1.65) vs. 19.36 (1.55), all comparisons p>0.05 12 months: 13.08 (1.48) vs. 14.42 (1.48) vs. 18.64 (1.39), Usual care vs. both other arms p<0.05, intervention arms vs. each other p>0.05	Overall function: SF-36 physical function (0 to100 scale) mean, (SE): 3 months: 43.25 (3.20) vs. 43.75 (3.32) vs. 37.26 (3.13), all comparisons p>0.05 12 months: 46.50 (3.68) vs. 45.75 (3.68) vs. 44.07 (3.47), all comparisons p>0.05
Pinxterhuis, 2017[83] Fukuda or Canadian Criteria (2003)	A: Group based self-management (73) B: Usual care (73) Duration of treatment: 16 weeks Duration of follow-up: 1 year	Fatigue Severity Scale (9 to 63), mean (SD): 6 months: 56.0 (6.8) vs. 55.5 (8.2); p=0.039; Mean change from baseline (95% CI): -0.2 (-1.7, 1.3) vs. -2.7 (-4.7, -0.7) 12 months: 56.4 (6.9) vs. 57.1 (6.7); p=NS; Mean change from baseline (95% CI): 0.4 (-1.4, 2.2) vs. -1.4 (-3.0, 0.1)	Not reported	SF-36 physical function (0 to 100), mean (SD): 6 months: 47.5 (21.2) vs. 50.5 (23.7); p=NS; Mean change from baseline (95% CI): 0.6 (-2.9, 4.0) vs. 4.3 (-0.4, 8.9) 12 months: 48.9 (17.7) vs. 46.3 (22.3); p=NS; Mean change from baseline (95% CI): 0.8 (-4.2, 5.7) vs. -0.3 (-5.4, 4.9)
Rimes, 2013[80] Fukuda or Oxford	A: Mindfulness-based cognitive therapy (18) B: Wait list (19) Duration of treatment: 2 months Duration of follow-up: 4 months	Modified Chalder Fatigue Scale 11-item (0 to 33), mean (SD): 2-month follow-up: 21.3 (6.2) vs. 25.0 (6.1)	HADS depression (0 to 21), mean (SD): 2-month follow-up: 5.6 (2.9) vs. 7.7 (4.6); p=0.153	PF-10 (0 to 100), mean (SD): 2-month follow-up: 65.6 (26.3) vs. 55.9 (23.3) Work and Social Adjustment Scale (0 to 40), mean (SD): 2-month follow-up: 20.0 (10.4) vs. 25.8 (6.7)
Surawy, 2005[81] Fukuda	A: Group mindfulness training (9) B: Usual care (9) Duration of treatment: 8 weeks Duration of follow-up: 8 weeks	Chalder Fatigue Scale 14-item (0 to 42), mean (SD):18.56 (8.13) vs. 20.38 (8.26), p=0.08	HADS depression (0 to 21) mean (SD): 8.33 (1.66) vs. 9.50 (3.96), p=0.28	SF-36 physical function (0 to 100), mean (SD): 40.00 (16.78) vs. 35.50 (27.00), p=0.58
Taylor, 2004[79] Fukuda	A: Illness management and peer counseling (23) B: Wait list (24) Duration of treatment: 11 months Duration of follow-up: 11 months	Not reported	Not reported	Not reported

Abbreviations: ACT = anaerobic activity therapy; CI = confidence interval; CFS = chronic fatigue syndrome; FSM:CTR = fatigue self-management with paper diaries and step counters; HADS = Hospital Anxiety and Depression Scale; ME = myalgic encephalomyelitis; NS = not significant; PF = physical function; RCT = randomized controlled trial; SD = standard deviation; SE = standard error; SF-36 = 36-item Short Form Health Survey

CBT in Adolescents

Five trials evaluated CBT in adolescents with ME/CFS.[84-88] Three trials[85-87] compared CBT versus inactive controls, one trial compared CBT plus biofeedback versus biofeedback alone,[84] and one trial[88] compared CBT versus pacing (**Tables 13 and 14; Evidence Table Appendix E2**). Three trials[84,86,87] used the Fukuda case definition, one trial[88] used the Oxford case definition, and one trial[85] used the Fukuda or Oxford case definitions. The duration of ME/CFS symptoms ranged from a mean or median of 26 weeks to 2 years. Sample sizes ranged from 13 to 127 (total N=438) and the mean age ranged from 12 to 16 years. The duration of treatment ranged from 6 to 18 months. All trials assessed outcomes at the end of treatment and one trial also assessed outcomes 12 months following the end of the intervention. Four trials were rated medium risk of bias and two trials[84,88] were rated high risk of bias (**Risk of Bias Table Appendix F**).

Table 13. RCTs of CBT and behavioral approaches in adolescents: study characteristics

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Al-Haggar, 2016[84] Egypt High	n: 92 Age: 12.6 % Female: 73	Criteria: Fukuda Duration: 26.4 weeks	Fatigue Scale: Fatigue Activity Scale (reported as %) Baseline: 53.5 (SD 3.9) Post-exertional fatigue or malaise: not reported	Major depression: Excluded Baseline depression: not reported Baseline function: not reported	A: CBT plus biofeedback B: Biofeedback Frequency: 40 to 60 sessions over 18 months once to twice weekly Session length: not reported Duration of treatment: 18 months Duration of follow-up: 18 months
Chalder, 2010[85] United Kingdom Medium	n: 63 Age: 15 median % Female: 68	Criteria: Oxford or Fukuda Duration: 24 months	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 23.6 (SD 5.4) Post-exertional fatigue or malaise: Not reported	Major depression: Excluded Baseline depression: not reported Baseline function: SF-36 physical function (0 to 100): 46.5 (SD 25.6)	A: Family-focused CBT B: Psycho-education (4 sessions over 6 months) Frequency: 13 sessions biweekly Session length: 1 hour Duration of treatment: 6 months Duration of follow-up: 18 months
Crawley, 2018[89] United Kingdom Medium	n: 81 Age: 14.6 % Female: 76	Criteria: NICE Duration: 12 months	Fatigue Scale: Chalder (11-item, 0 to 33) Baseline: 25.0 (SD 4.2) Post-exertional fatigue or malaise: Not reported	Major depression: not reported Baseline depression: HADS depression (0 to 21): 7.8 (SD 3.8) Baseline function: SF-36 physical function (0 to 100): 54.5 (SD 20.2)	A: Osteopathy, life coaching, and neurolinguistic programming intervention (Lightning Process) plus specialist medical care B: Specialist medical care Frequency: three 4-hour sessions plus 2 follow-up sessions Session length: 4 hours for initial three sessions Duration of treatment: 3 days Duration of follow-up: 12 months

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Nijhof, 2012[86] (FITNET) The Netherlands Medium	n: 127 Age: 15.8 % Female: 82	Criteria: Fukuda Duration: Median 16 vs. 19 months	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 51.4 (SD 4.5) Post-exertional fatigue or malaise: not reported	Major depression: Excluded for primary depression Baseline depression: Children's Depression Inventory (0 to 54): 11.3 (SD 5.2) Baseline function: Child Health Questionnaire-CF87 physical functioning (0 to 100): 58.8 (SD 18.0)	A: Web-based CBT B: Usual care Frequency: Weekly then biweekly therapist contact; 21 interactive modules Session length: Not described Duration of treatment: 6 months Duration of follow-up: 6 months
Stulemeijer, 2005[87] The Netherlands Medium	n: 62 Age: 15.6 % Female: 90	Criteria: Fukuda Duration: Median 16 vs. 18 months	Fatigue Scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline: 52.1 (SD 4.0) Post-exertional fatigue or malaise: Not reported Pervasively passive: 25% Relatively active: 72%	Major depression: Excluded for psychiatric comorbidity Baseline depression: Depression scale not assessed Baseline function: SF-36 physical function (0 to 100): 43.7 (SD 16.8)	A: CBT based on activity pattern, with parental involvement B: Wait list Frequency: 10 sessions over 5 months Session length: Not described Duration of treatment: 5 months Duration of follow-up: 5 months
Wright, 2005[88] United Kingdom High	n: 13 Age: Mean not reported, range 8.9 to 16.9 years % Female: Not reported	Criteria: Oxford Duration: Median 12 vs. 14.5 months	Fatigue Scale: Chalder (14-item, 0 to 42) Baseline: 20.4 (SD 7.9) Post-exertional fatigue or malaise: Not reported	Major depression: Proportion with depression not reported Baseline depression: Birleson Depression Scale (0 to 36): 15.0 (SD 5.6) Baseline function: Young Persons' Functional Ability Scale (0 to 100): 59.2 (SD 20.8)	A: Cognitive therapy and education (STAIRway to Health) B: Pacing Frequency: Weekly for 1 month, every 2 weeks for 3 months, every 3 weeks for 2 months, and every 4 weeks for 6 months Session length: Not described Duration of treatment: 12 months Duration of follow-up: 12 months

Abbreviations CBT = cognitive behavioral therapy; CFS = chronic fatigue syndrome; FITNET = fatigue in teenagers on the internet; HADS = Hospital Anxiety and Depression Scale; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence; RCT = randomized controlled trial; SD = standard deviation; SE = standard error; SF-36 = 36-item Short Form Health Survey

Table 14. RCTs of CBT and behavioral approaches in adolescents: study results

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Al-Haggar, 2016[84] Fukuda	A: CBT plus biofeedback (50) B: Biofeedback (46) Duration of treatment: 18 months Duration of follow-up: 18 months	Fatigue severity, mean (SD) Checklist Individual Strength, fatigue severity subscale (8 to 56): 32.2 (3.8) vs. 46.5 (14.2), p=0.02	Not reported	School attendance, mean (SD) hours per month: 92.8 (18.4) vs. 66.6 (22.8), p=0.004
Chalder, 2010[85] Oxford or Fukuda	A: Family-focused CBT (32) B: Psycho-education (4 sessions over 6 months) (27) Duration of treatment: 6 months Duration of follow-up: 18 months	Chalder fatigue Likert score at 6-month follow-up (11-item, 0 to 33), mean (SD): 13.3 (5.9) vs. 14.2 (8.4), mean difference: 0.24, 95% CI -3.61 to 4.10	Not reported	No significant effects of group x time (6 and 24 months) in fatigue, SF-36 physical function, global functioning, satisfaction, or recovery
Crawley, 2019[89] NICE	A: Osteopathy, life coaching, and neurolinguistic programming intervention (Lightning Process) plus specialist medical care (51) B: Specialist medical care (49) Duration of treatment: 3 days Duration of follow-up: 12 months	Chalder Fatigue Scale 11-item (0 to 33) 6 months, mean: 14.4 vs. 19.8, adjusted difference in means: -4.7 (95% CI, -7.9 to 1.6), p=0.003 Fatigue, Mean Chalder Fatigue Scale 11-item (0 to 33) 12 months: 12.3 vs, 15.7, adjusted difference in means: -3.2 (95% CI, -6.3 to 0.10), p=0.045	HADS-Depression, mean: 6 months: 4.2 vs. 5.9, p=0.141 12 months: 2.8 vs. 4.6, p=0.033	Overall Function, SF-36 physical function (0 to 100) at 6 months, mean: 81.7 vs. 70.2, adjusted (based on age, gender and baseline outcome) difference in means: 12.5 (95% CI, 4.5 to 20.5), p=0.003
Nijhof, 2012[86] (FITNET) Fukuda	A: Web-based CBT (68) B: Usual care (67) Duration of treatment: 6 months Duration of follow-up: 6 months	Fatigue severity at 6 months, Checklist Individual Strength, fatigue severity subscale (8 to 56), cutoff score <40: 85% (57/67) vs. 27% (17/64), RR 3.2 (95% CI, 2.1 to 4.9), NNT 1.7, p<0.0001	Not reported	Physical functioning: Child Health Questionnaire-CF87 physical functioning (0 to 100) cutoff score of 85% or more) at 6 months: 78% (52/67) vs. 20% (13/64), RR 3.8 (95% CI, 2.3 to 6.3), NNT 1.8, p<0.0001
Stulemeijer, 2005[87] Fukuda	A: CBT based on activity pattern, with parental involvement (36) B: Wait list (35) Duration of treatment: 5 months Duration of follow-up: 5 months	Fatigue severity subscale of the checklist of individual strength at 5 months, mean (SD): 30.2 (16.8) vs. 44.0 (13.4), treatment effect 17.3 (95% CI, 6.2 to 28.4), p=0.003	Not reported	SF-36 physical function (0 to 100) at 5 months, mean (SD): 69.4 (28.0) vs. 55.3 (21.1), treatment effect 14.5 (95% CI, 7.4 to 21.6), p=0.001

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Wright, 2005[88] Oxford	A: Cognitive therapy and education (STAIRway to Health) (7) B: Pacing (6) Duration of treatment: 12 months Duration of follow-up: 12 months	Fatigue score (Chalder 0 to 42 14 item version): 25.2 (219.8 to 9.49); F= 0.67; p= 0.44	Not reported	Young Person Functional Ability Scale (0 to 100): 17.0 (217.0 to 51.0) F=1.3; p= 0.28

Abbreviations: CBT = cognitive behavioral therapy; CFS = chronic fatigue syndrome; CHQ-CF = child health questionnaire-child form; CI = confidence interval; FITNET = fatigue in teenagers on the internet; ME = myalgic encephalomyelitis; NICE = National Institute for Health and Care Excellence; NNT = number needed to treat; RCT = randomized controlled trial; RR = relative risk; SD = standard deviation; SE = standard error; SF-36 = 36-item Short Form Health Survey

CBT Versus Inactive Controls in Adolescents

Three trials compared individual CBT versus usual care, wait list, or an attention control (psycho-education) in adolescents with ME/CFS.[85-87] The mode of administration was face-to-face in two trials and via web in one trial. All three trials noted a family focus or involvement of patients in CBT. The duration of therapy was 5 to 6 months. All trials evaluated outcomes at the end of the intervention and one trial evaluated outcomes 1 year following the completion of therapy. Two of the trials also reported longer-term, post-trial follow-up based on the original randomized groups.[85,86] All of the trials were rated medium risk of bias (**Risk of Bias Table Appendix F**). Results are summarized in **Table 15** and shown in **Figures 57 to 59**.

Fatigue

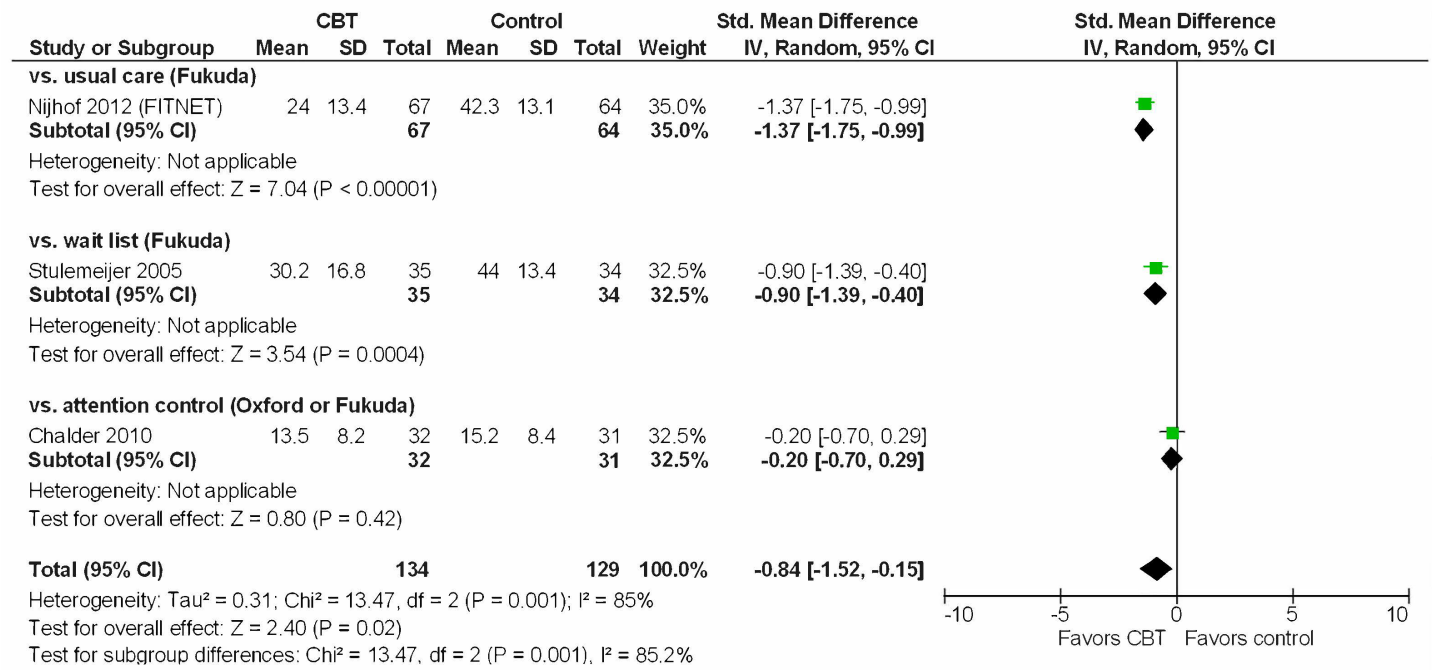
CBT was associated with decreased fatigue severity versus usual care, wait list, or an attention control (3 trials, N=263, SMD -0.84, 95% CI -1.52 to -0.15, I^2=85%;[85-87] **Table 15, Figure 57**). Although the estimate in all trials favored CBT, statistical heterogeneity was large. One trial of CBT versus an attention control (psychoeducation) reported a small and non-statistically significant effect on fatigue severity (SMD -0.20, 95% CI -0.70 to 0.29).[85] The other two trials, which compared CBT versus usual care or wait list, each reported larger effects, with a statistically significant pooled estimate (2 trials, N=200, mean difference -16.9, 95% CI -21.0 to -12.9 on the 8 to 56 Checklist Individual Strength fatigue severity subscale, I^2=8%).[86,87]

The trial of CBT versus an attention control found no difference in fatigue severity at 1 year post-intervention follow-up (N=63, mean difference -1.9, 95% CI -5.3 to 1.5 on the 0 to 33 11-item Chalder scale).[85]

CBT was associated with increased likelihood of improvement in fatigue versus usual care or wait list (2 trials, N=200, RR 3.13, 95% CI 2.18 to 4.49, I^2=0%; ARD 51%, 95% CI 32% to 69%).[86,87] Improvement in fatigue was defined as an improvement in the Checklist Individual Strength fatigue severity subscale score <40 in one trial[86] and as a score ≤35.7 and reliable change index >1.96 in the other trial.[87]

One study (N=44) that reported long-term, post-trial follow-up at 24 months (12 months after trial completion) found no difference in severity of fatigue (data not provided).[90]

Figure 57. Fatigue severity in adolescents: CBT versus inactive controls



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; FITNET = fatigue in teenagers on the internet; IV = instrumental variable; SD = standard deviation; Std = standard

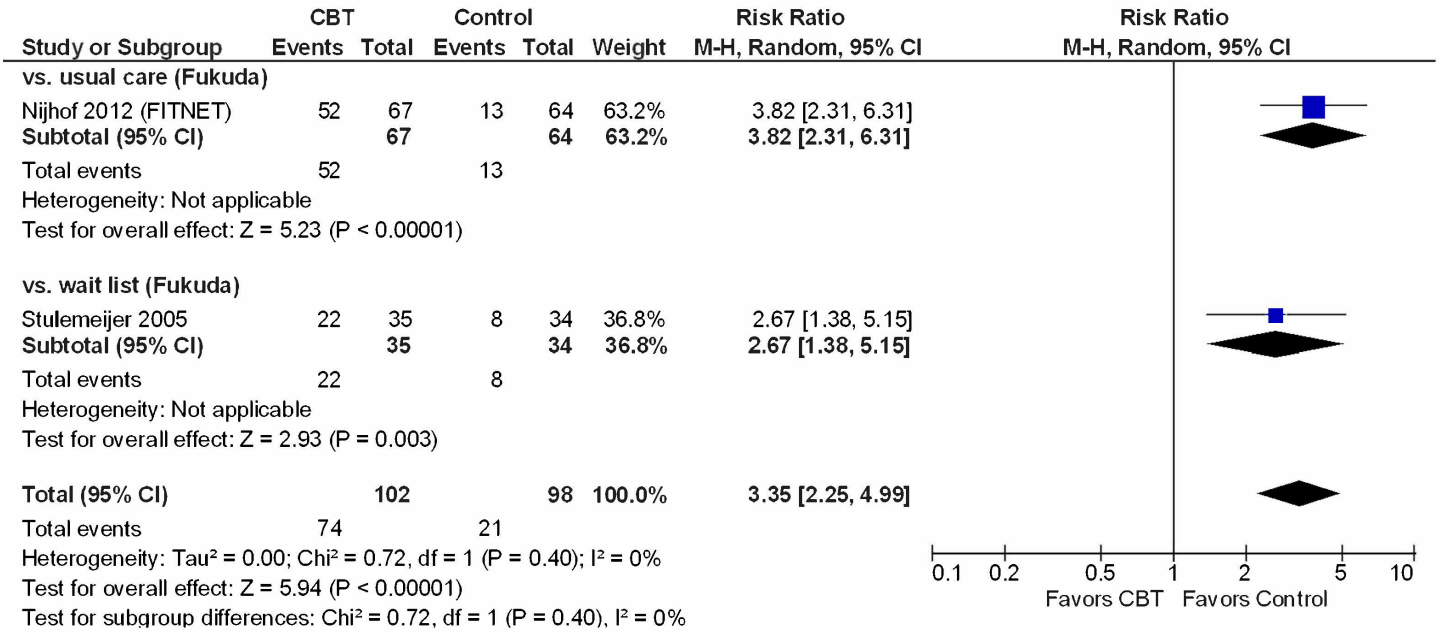
Function

There was no statistically significant difference between CBT versus usual care, wait list, or an attention control in severity of functional impairment, though the estimate favored CBT (3 trials, N=263, SMD 0.49, 95% CI -0.34 to 1.32, I^2=90%).[85-87] Statistical heterogeneity was large, and the pooled estimate was imprecise. The trial[85] that compared CBT versus an attention control (psychoeducation) did not report a positive effect on severity of functional impairment (SMD -0.28, 95% CI -0.77 to 0.22) while the other two trials[86,87] found CBT associated with less severe functional impairment versus usual care or wait list (SMD 1.16, 95% CI 0.79 to 1.53 and SMD 0.56, 95% CI 0.08 to 1.04). In these trials, differences were 14 to 18 points on the 0 to 100 SF-36 or Child Health Questionnaire-CF87 physical function subscales.

The trial of CBT versus an attention control also found no difference in severity of functional impairment at 1 year post-intervention follow-up (N=63, mean difference 6.1, 95% CI -9.2 to 21.4 on the 0 to 100 SF36 physical function subscale).[85] There was also no difference in patients originally randomized to this trial at long-term (24 month) post-trial follow-up (mean difference 5.6, 95% CI -12.1 to 23.3).[90]

CBT was associated with increased likelihood of improvement in function versus usual care or wait list (2 trials, N=200, RR 3.35, 95% CI 2.25 to 4.99, I^2=0%; ARD 50%, 95% CI 33% to 68%;[86,87] **Figure 58**). One trial[86] defined improvement in fatigue as a Child Health Questionnaire-CF87 score ≥85 and the other trial[87] as an SF-36 physical function score increase ≥50 or score ≥75.

Figure 58. Likelihood of functional improvement in adolescents: CBT versus inactive controls

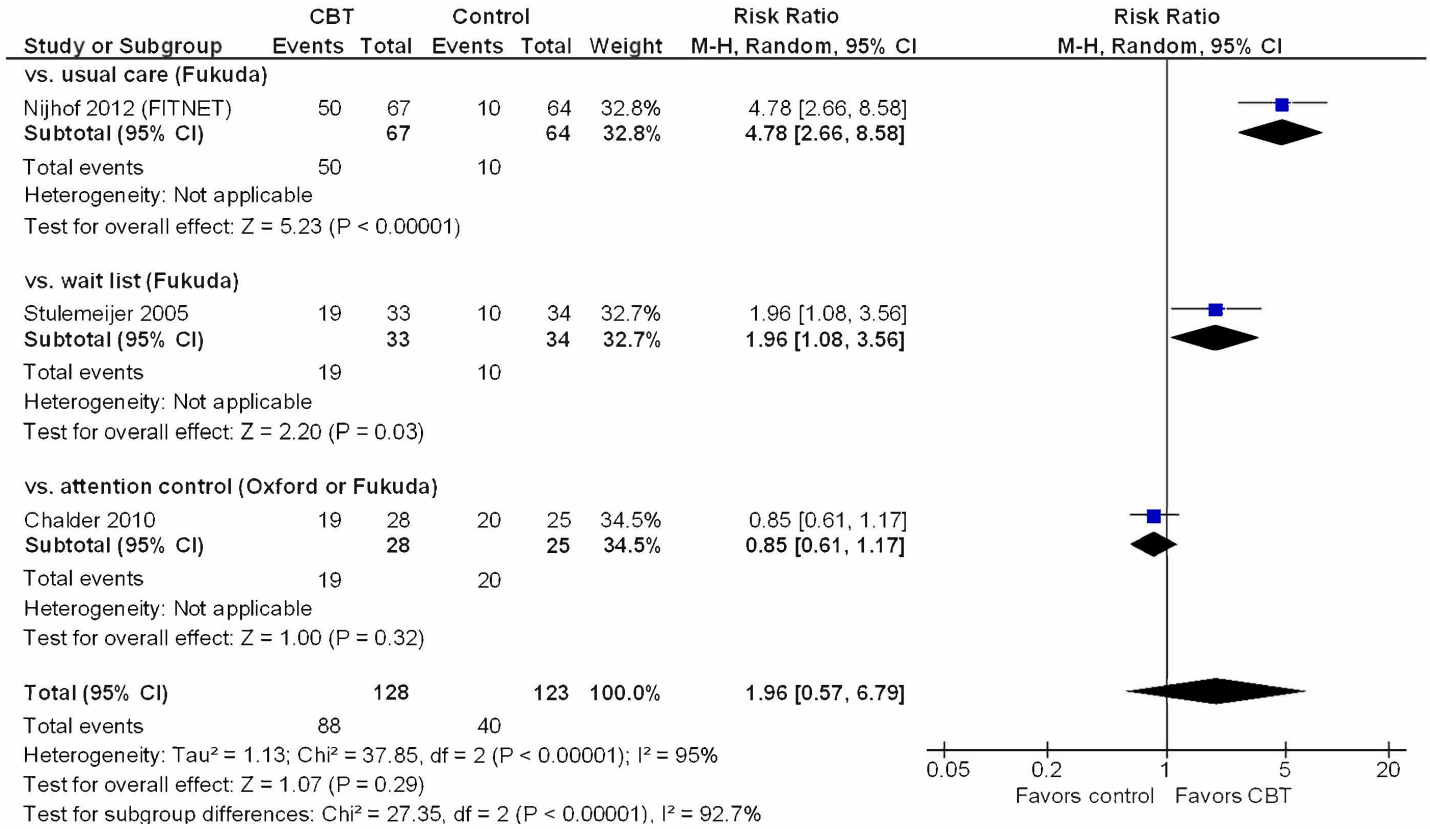


Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; FITNET = fatigue in teenagers on the internet; M-H = Mantel-Haenszel test

School attendance

There was no statistically significant difference between CBT versus usual care, wait list, or an attention control in likelihood of school attendance (3 trials, N=251, RR 1.96, 95% CI 0.57 to 6.79, I^2=95%; **Figure 59**).[85-87] Although the estimate favored CBT, statistical heterogeneity was large and the estimate was imprecise. The trial of CBT versus an attention control showed no effect on likelihood of school attendance (N=53, RR 0.85, 95% CI 0.61 to 1.17);[85] results were similar at long-term (24 month) post-trial follow-up.[90] In the other two trials CBT was associated with increased likelihood of school attendance versus usual care of wait list (2 trials, N=198, RR 3.06, 95% CI 1.25 to 7.49; I^2=78%; ARD 45%, 95% CI 14% to 75%).

Figure 59. Likelihood of school attendance: CBT versus inactive controls



Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; FITNET = fatigue in teenagers on the internet; M-H = Mantel-Haenszel test

Overall improvement

There was no statistically significant difference between CBT versus usual care, wait list, or an attention control in likelihood of overall improvement (3 trials, N=256, RR 1.66, 95% CI 0.67 to 4.10, I^2=95%).[85-87] Although the estimate favored CBT, statistical heterogeneity was substantial and the estimate was imprecise. The trial that compared CBT versus an attention control showed no effect on likelihood of overall improvement, defined as a child-reported global improvement "good outcome" (N=56, RR 0.99, 95% CI 0.83 to 1.19).[85] The other two trials each found CBT associated with increased likelihood of overall improvement, defined as self-rating of "I have completely recovered" or "I feel much better but still experience some symptoms" (2 trials, N=200, RR 2.18, 95% CI 1.21 to 3.93, I^2=73%).[86,87]

Recovery

One trial found web-based CBT associated with increased likelihood of recovery versus usual care at the end of treatment (N=131, 78% vs. 20%, RR 3.82, 95% CI 2.31 to 6.31).[86] Recovery was defined as school absence <10%, Checklist Individual Strength, fatigue severity subscale <40, CHQ-CF87 ≥85, and overall assessment of "I have completely recovered" or "I feel much better but still experience some symptoms." However, there was no difference in this trial in likelihood of recovery at long-term (2.7 year) post-trial follow-up (64% vs. 53%, RR 1.20, 95% CI 0.81 to 1.78).[91] Another trial found no differences between CBT versus an attention control in likelihood of recovery (defined as 11-item 0 to 33 Chalder score ≤18 and

school attendance ≥70%) at 6-month post-intervention follow-up (68% vs. 69%)[85] or at long-term, 24 month post-trial follow-up (79% vs. 64%, p=0.34).[90]

Additional Fukuda 1994 criteria symptoms

One trial (N=69) found CBT associated with improvement in severity of unrefreshing sleep (mean difference -1.2, 95% CI -1.8 to -0.6), muscle pain (mean difference -1.1, 95% CI -1.6 to -0.6), and impaired concentration (mean difference -1.1, 95% CI -1.5 to -0.65) versus wait list (all assessed on a 4 point Likert scale).[87] There were no differences in severity of headache, impaired memory, multi-joint pain, or sensitive lymph nodes.

Harms

Harms were not well-reported. Two trials reported no serious adverse events.[86,89] One trial found CBT associated with decreased tiredness after exercise versus wait list (N=69, mean difference -1.0, 95% CI -1.5 to -0.5 on a 4 point Likert scale).[87]

Table 15. CBT versus inactive controls in adolescents: summary of results

Outcome	Number of Studies (N)	Estimate (95% CI)	I^2
Fatigue, end of intervention	3 (263)	SMD -0.84 (-1.52 to -0.15)	85%
11-item Chalder fatigue scale (0 to 33), post-intervention	1 (63)	MD -1.9 (-5.3 to 1.5)	--
Fatigue improvement (dichotomous)	2 (200)	RR 3.13 (2.18 to 4.49)	0%
Function, end of intervention	3 (263)	SMD 0.49 (-0.34 to 1.32)	90%
SF-36 physical function (0 to 100), post-intervention	1 (63)	MD 6.1 (-9.2 to 21.4)	--
Functional improvement	2 (200)	RR 3.35 (2.25 to 4.99)	0%
School attendance	3 (251)	RR 1.96 (0.57 to 6.79)	95%
Overall improvement	3 (256)	RR 1.66 (0.67 to 4.10)	95%
Recovery	1 (131)	RR 3.82 (2.31 to 6.31)	--

Abbreviations: CI = confidence interval; CBT = cognitive behavioral therapy; MD = mean difference; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference

Cognitive Behavioral Therapy Plus Biofeedback Versus Biofeedback

One trial (N=92) compared CBT plus biofeedback versus biofeedback alone in adolescents **(Tables 13 and 14)**.[84] The duration of treatment was 18 months and outcomes were assessed at the end of therapy. The trial was rated high risk of bias; attrition was high (40%) and persons who withdrew were excluded from the analysis. CBT plus biofeedback was associated with less severe fatigue (mean difference 12.2, 95% CI 7.4 to 14.8 on the 0 to 100 Fatigue Activity Scale) and greater school attendance (mean difference 23 hours monthly, 95% CI 20.6 to 26.8). CBT plus biofeedback was also associated with less severity of unrefreshing sleep (mean difference -1.20, 95% CI -1.62 to -0.78) and myalgia (mean difference -0.80, 95% CI -1.23 to -0.37), with no difference in joint pains or tender glands (each symptom assessed on a 4-point Likert scale). Harms were not reported.

Cognitive Therapy and Education Versus Pacing

One small (N=17) pilot trial compared a cognitive therapy and education program (STAIRway to Health) versus pacing in adolescents. The duration of treatment was 12 months and outcomes were assessed at the end of therapy **(Tables 13 and 14)**.[88] There were no differences in fatigue, function, anxiety or depression, though estimates were imprecise. Global health ratings favored the cognitive therapy and education program (mean difference -1.8, 95% CI -2.7 to -0.9 on a 1 to 5 scale).

Other Behavioral Approaches in Adolescents

Osteopathy, life coaching and neurolinguistic programming intervention plus usual specialist care versus usual specialist care

One medium risk of bias trial (N=81) compared an osteopathy, life coaching and neurolinguistic programming intervention ("Lightning Process") plus usual specialist care versus usual specialist care alone in adolescents (mean age 14.6 years) who met the NICE case definition (**Tables 13 and 14, Evidence Table Appendix E2, Risk of Bias Table Appendix F**).[89] The intervention consisted of three 4-hour sessions plus two follow-up sessions and outcomes were assessed through 12 months. At the end of follow-up, the intervention was associated with improved function (adjusted mean difference 12.9, 95% CI 3.6 to 22.1 on the 0 to 100 SF-36 physical function subscale); results were similar at 6-month follow-up. There were no statistically significant effects of the intervention on fatigue, pain, anxiety, depression, or quality of life at 6 or 12 months. There was also no difference between groups in school attendance at 6 months, though the intervention was associated with higher attendance at 12 months (adjusted mean difference 1.0, 95% CI 0.2 to 1.8 for days of attendance in the previous week).

Predictors of treatment response in trials of exercise therapy and CBT

Two trials evaluated how application of different ME/CFS case definitions impacted outcomes.[38,52] GETSET, which enrolled patients who met the NICE case definition, found no differences in results when analyses were restricted to patients who also met the Fukuda or Oxford case definitions.[52] PACE, which enrolled patients who met the Oxford case definition,[38] found no interactions between whether patients also met the 2003 International CFS criteria definition[92] or the 1994 London case definition[24] and effects of interventions on fatigue or function. Since the alternative definitions were a subset of patients who met the original criteria, the applicability of these analyses when alternative case definitions are applied independently is unclear. PACE also found no interactions between presence of a primary depressive or anxiety disorder and effects of interventions.

Evidence on the interaction between severity of baseline functional impairment and effects of exercise of CBT was limited and inconsistent. The GETSET trial found an interaction between worse physical function at baseline and larger effects of exercise on function at follow-up, though there was no interaction between baseline physical function and fatigue.[52] However, two trials found lower baseline function associated with poorer response to exercise or CBT.[71,93] One other trial found that effects of CBT versus wait list on fatigue and function were slightly greater in the subgroup of patients with baseline SF-36 physical function score ≤70 compared to the whole study population, but did not report results in the subgroup with a score >70 or perform statistical testing for a subgroup effect.[76]

Three trials found no interaction between baseline depression and effects of exercise or CBT.[52,69,93] One trial also found no interaction between receipt of antidepressant therapy, sleep disturbance, duration of illness, or initial illness beliefs and effects of exercise.[93]

Other Therapies

Medications

Nineteen randomized controlled trials (RCTs) evaluated pharmacological treatment of ME/CFS.[59,75,94-111] The pharmacological therapy was an immune modulating drug (rintatolimod, IgG, rituximab, anakinra, and alfa-interferon) in nine trials,[94,95,99-101,104-107] an antidepressant in four trials,[59,75,103,110] and a corticosteroid in two trials;[96,111] valganciclovir,[98] galantamine,[108] clonidine,[102] and methylphenidate plus a nutritional supplement aimed at modulating mitochondrial function were each evaluated in one trial (**Tables 16 and 17; Evidence Table Appendix E2, Risk of Bias Table Appendix F**).[109] Eleven of these were not included in the prior AHRQ report (6 trials of immune modulators, 3 trials of antidepressants, 1 of clonidine, and 1 of methylphenidate). None of the studied medications have been FDA-approved for treating ME/CFS. Intravenous rintatolimod is not FDA approved for any indication, although it was reviewed by the FDA for ME/CFS and failed to receive approval in 2012. All of the studies compared the study drug versus matching placebo; one study of fluoxetine also randomized patients to GET (see exercise Results for comparison of graded exercise versus fluoxetine).[59] The median duration of study treatment was 12 weeks (range 4 to 42 weeks). The timing of outcome assessment ranged from the end of treatment to 30 weeks after the end of treatment. Eight of the trials were conducted in the United States, six in Europe, and two in Australia. All of the included studies evaluated the effect of the study drug on ME/CFS symptoms; one trial of an antidepressant stratified randomization based on depression status at baseline to measure the impact on depression symptoms (see Question 3a).[103]

Ten RCTs enrolled patients based on the Fukuda case definition,[94,96-98,102,104,106-111] three used the prior CDC case definition (Holmes, 1988),[99,100,105] one used a combination of the Fukuda and Holmes criteria,[101] two used the Oxford case definition,[59,103] and one used the Canadian Consensus criteria (Carruthers 2003).[95] Seventeen trials enrolled adults, weighted mean age 37.5 years (range of mean enrolled age 31 to 49),[59,75,94-101,103-106,108-111] and two trials enrolled only adolescents, both with a mean age of 15 years.[102,107] The proportion female ranged from 47% to 100% and the sample size ranged from 26 to 423 (N=1,150). Most trials did not report race or ethnicity, but when reported, most participants were White. The mean duration of illness ranged from 18 months to 13 years in 12 trials that provided this information.[59,96-107,111] Methods for measuring severity of baseline fatigue and functional status varied and some trials did not report baseline values (**Table 16**). The most common methods for measuring baseline fatigue were the Chalder 14-item 0 to 42 scale (two trials, mean score at baseline 32 in one trial and not reported in the other)[59,108] or 11-item 0 to 33 scale (1 trial, mean 19.2)[102] and the POMS, fatigue subscale in two trials (mean 18.8 and 18.7 out of 24).[96,104] The most common methods for measuring baseline functional status were the SF-36 physical function subscale (2 trials, mean 53.9 and 59.5 on a 0 to 100 scale)[105,110] and the KPS (3 trials, mean ranged from 51 to 70.3 on a 0 to 100 scale);[100,101,104] the other trials used different methods to evaluate function,[106,108] or did not report baseline functional status. One antidepressant trial excluded patients with major depression at baseline[110] and in another antidepressant trial 50% of patients had major depression at baseline.[103] Three non-antidepressant medication trials reported that 3% to 24% had major depression at baseline;[59,96,100] the other trials did not describe depression status.

Three trials (anakinra,[106] clonidine,[102] and rituximab[95]) were rated low risk of bias and one trial (alfa-2a interferon)[99] was rated high risk of bias, primarily due to poor reporting of methods or attrition[99] (**Evidence Table Appendix E2**). The other trials were rated moderate risk of bias.

Eight of the trials reported funding and/or drug and matching placebo provided by pharmaceutical companies, five trials without industry support reported funding from foundations or other sources, and three trials did not report sources of support (**Risk of Bias Table Appendix F**).

Table 16. Medication RCTs: study characteristics

Author, Year Country Risk of Bias	Study N (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency Duration of Treatment Duration of Follow-up
Arnold, 2015[110] United States Medium	n=57 Age: 44 % Female: 87	Criteria: Fukuda Duration: >6 months	Fatigue Scale: CDC Symptom Inventory Baseline: 40.0 (SD 133)	Major depression: Excluded Baseline depression: HADS depression (0-21): 9.27 (SD 3.9 Function: SF-36 physical function (0 to 100): 59.5 (SD 19.8)	A: Duloxetine 120 mg/d B: Placebo Duration of treatment: 12 weeks (4 weeks at maximum dose) Duration of follow-up: 12 weeks
Blacker, 2004[108] United Kingdom Medium	n=423 Age: 38 % Female: 68	Criteria: Fukuda Duration: <7 years	Fatigue Scale: Chalder Baseline: NR	Major Depression: NR Baseline Depression: NR Function: FIQ 13.47 (SD NR)	A: Galantamine 2.5 mg B: Galantamine 5 mg C: Galantamine 7.5 mg D: Galantamine 10 mg E: Placebo Duration of treatment: 4 months (16 weeks, 8 weeks at full dose) Duration of follow-up: 4 weeks
Blockmans, 2003[111] Belgium Medium	n=80 Age: 38 % Female: 91	Criteria: Fukuda Duration: mean 30 months	Fatigue Scale: # criteria for CFS Baseline: 6 (SD 2)	Major Depression: NR Baseline Depression: HADS depression (0 to 21): 9.6 (SD 3.5) Function: SF-36 Physical Function (0 to 100): 27.3 (SD 12.3)	A: Hydrocortisone 5 mg daily + 9-alpha fludrocortisone 50 µg daily B: Placebo Duration of treatment: 3-month treatment; 3-month placebo crossover Duration of follow-up: end of 3-month crossover
Fluge, 2011[94] Norway Medium	N=30 Age: 34.4 % Female: 70	Criteria: Fukuda Duration: mean 6.6 years	Fatigue Scale: VAS (0 to 10) Baseline: 8	Major Depression: NR Baseline Depression: NR Function: SF-36 physical function (%, lower score denotes increasing symptoms): 34.5 (SD 6.5)	A. Rituximab 500 mg/m^2, maximum 1,000 mg B. Placebo Duration of treatment: 2 weeks Duration of follow-up: 12 months
Fluge, 2019[95] Norway Low	N=152 Age: 36.7 % Female: 82	Criteria: Canadian consensus (Carruthers, 2003) Duration: mean 8 years	Fatigue Scale: Scale not named (0 to 6) Baseline: 3.0	Major Depression: NR Baseline Depression: 8.5% Function: SF-36 Physical Function (%, lower score denotes increasing symptoms): 33.8	A. Rituximab 500 mg/m^2, maximum 1,000 mg B. Placebo Duration of treatment: 12 months Duration of follow-up: 24 months

Author, Year Country Risk of Bias	Study N (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency Duration of Treatment Duration of Follow-up
McKenzie, 1998[96] United States Medium	N=70 Age: 38 % Female: 80	Criteria: Fukuda and Holmes Duration: 54 months	Fatigue Scale: POMS fatigue subscale (0 to 28) Baseline: 18.7 (SD 5.2)	Major Depression: 3% Baseline Depression: HADS depression (0 to 21): 9.6 (SD 3.5) Function: NR	A. Hydrocortisone 20-30 mg every morning, 5 mg every evening B: Placebo Duration of treatment: 12 weeks Duration of follow-up: 12 weeks
Montoya, 2018[109] United States Medium	N=128 Age: 49 % Female: 63	Criteria: Fukuda Duration: 13.1 years	Fatigue Scale: MFI-20 (20 to 100) Baseline: 78.6 (SD NR)	Major Depression: NR Baseline Depression: NR Function: NR	A: Methylphenidate 20mg daily + Mitochondrial nutritional supplement: 4 tablets twice daily. B: Placebo Duration of treatment: 12 weeks (10 weeks at full dose) Duration of follow-up: 12 weeks
Montoya, 2013[98] United States Medium	N=30 Age: 43 % Female: 72	Criteria: Fukuda Duration: 53% <10 years, 47% >10 years	Fatigue Scale: MFI-20 (20 to 100) Baseline: 79.5 (SD 13.40)	Major Depression: NR Baseline Depression: NR Function: NR	A: Valganciclovir 900 mg BID for 21 days, then 900 mg daily B: Placebo Duration of treatment: 6 months Duration of follow-up: 12 months
Peterson, 1990[105] United States Medium	N=28 Age: 38 % Female: 73	Criteria: Holmes Duration: 3.8 years	Fatigue Scale: # of CFS criteria Baseline: 8.8 (SD NR)	Major Depression: NR Baseline Depression: NR Function: SF-36 physical function (o to 100): 53.9 (SD 22.7)	A: IgG 1 g/kg IV every 30 days B: Placebo Duration of treatment: 6 months Duration of follow-up: 6 months
Roenik 2017[106] The Netherlands Low	n=50 Age: 31 % Female:	Criteria: Fukuda Duration: 41 months	Fatigue Scale: Mean fatigue severity CIS-fatigue score (ranges from 8 to 56, higher scores indicate worse fatigue) Baseline: 51.5	Major Depression: NR Baseline Depression: MR Function: Sickness Impact Profile (SIP-8) (0 to 5,799): 1647 vs. 1706	A: Anakinra 100 mg SQ daily B: Placebo Duration of treatment: 4 weeks Duration of follow-up: 20 weeks after treatment
Rowe, 1997[107] Australia Medium	n=70 Age: 15 % Female: 100	Criteria: Fukuda Duration: 41 months	Fatigue Scale: CIS-fatigue score (ranges from 8 to 56, higher scores indicate worse fatigue) Baseline: 52 (SD NR)	Major Depression: NR Baseline Depression: NR Function: Mean functional impairment Sickness Impact Profile (SIP-8) (0 to 5,799): 1677 (SD NR)	A: IgG 1 gm/kg IV every 30 days B: Placebo Duration of treatment: 12 weeks Duration of follow-up: 25 weeks follow-up

Author, Year Country Risk of Bias	Study N (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency Duration of Treatment Duration of Follow-up
See, 1996[99] United States High	n=26 Age: 37 % Female: 80	Criteria: Holmes Duration: 4.6 years	Fatigue Scale: NR Baseline: NR	Major Depression: NR Baseline Depression: NR Function: NR	A: Alfa-2a Interferon 3 mu SQ 3 times per week B: Placebo Duration of treatment: 12 weeks Duration of follow-up: 12 weeks
Strayer, 1994[100] United States Medium	n=84 Age: NR % Female: 75	Criteria: Holmes and Fukuda, 1994) Duration: 5.25 years	Fatigue Scale: NR Baseline: NR	Major Depression: 24% Baseline Depression: NR Function: KPS (0 to 100): 51 (SD NR)	A: Rintatolimod 200 mg IV twice weekly 4 times, then 400 mg twice weekly B: Placebo Duration of treatment: 6 months Duration of follow-up: 6 months
Strayer, 2012[101] United States Medium	n=240 Age: 44 % Female: 71	Criteria: Holmes and Fukuda Duration: 9.7 years	Fatigue Scale: NR Baseline: NR	Major Depression: NR Baseline Depression: NR Function: NR	A: Rintatolimod 400 mg IV twice weekly B: Placebo Duration of treatment: 40 weeks Duration of follow-up: 40 weeks
Sulheim, 2014[102] Norway Low	n=96 Age: 15 % Female: 72	Criteria: Fukuda Duration: 18 months	Chalder Fatigue Scale 11-item (0 to 33): Baseline: 19.2 (SD NR)	Major Depression: NR Baseline Depression: NR Function: NR	A: Clonidine 25-50 mcg based on weight. B: Placebo Duration of treatment: 9 weeks treatment Duration of follow-up:30 weeks follow-up
Vercoulen, 1996[103] The Netherlands Medium	n=96 Age: 39 % Female: 47%	Criteria: Oxford Duration: 6 years	Fatigue Scale: Subjective fatigue, daily observed fatigue score, measured 4 times a day on a 4-point scale, and combined, with higher scores indicating worse fatigue: Baseline: 9.4 (SD NR)	Major Depression: 50% Baseline Depression: Beck Depression Inventory (0 to 63): 22.5 in depressed group; 7,5 in non-depressed Function: NR	A: Fluoxetine 20 mg daily B: Placebo Duration of treatment: 8 weeks treatment Duration of follow-up: 10 weeks follow-up
Vollmer-Conna, 1997[104] Australia Medium	n=99 Age: 40 % Female: 76	Criteria: Fukuda Duration: 6.25	Fatigue Scale: POMS fatigue subscale (0 to 28): Baseline: 18.8 (SD 6.3)	Major Depression: NR Baseline Depression: POMS Depression (0 to 60): 15.7 (SD 12.1) Function: KPS (0 to 100): 70.3 (SD 10)	A: IgG 0.5 gm/kg B: IgG 1.0 mg/kg C: IgG 2.0 mg/kg D: Placebo IV every 30 days Duration of treatment: 12 weeks Duration of follow-up: 12 weeks

Author, Year Country Risk of Bias	Study N (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency Duration of Treatment Duration of Follow-up
Wearden, 1998[59] United Kingdom Medium	n=68 Age: 39 % Female: 71	Criteria: Oxford Duration: 28 months	Fatigue Scale: Chalder Fatigue Scale 14-item (0 to 42) Baseline: 34 (SD NR)	Major Depression: 10% Baseline Depression: HADS depression score (0 to 21): 8.8 (SD 3.5) Function: NR	A: Fluoxetine 20 mg daily B: Placebo Duration of treatment: 26 weeks Duration of follow-up: 26 weeks

Abbreviations: BDI = Beck Depression Inventory; BID = twice daily; CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; FIQ = Fibromyalgia Impact Questionnaire; HADS-D = Hospital Anxiety and Depression Scale-depression; IgG = immunoglobulin G; IV = intravenous; KPS = Karnofsky Performance Scale; ME = myalgic encephalomyelitis; MFI-20 = Multidimensional Fatigue Inventory 20-item; NR = not reported; POMS = profile of mood states; q = every; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey; SIP-8 = Sickness Impact Profile; SQ = subcutaneous

Table 17. Medication RCTs: study results

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Arnold, 2015[110] Fukuda	A. Duloxetine 120 mg/d (30) B. Placebo (30) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	MFI-20 general fatigue subscale (4 to 20), observed mean change (SD): -3.3 (4.2) vs. -1.8 (2.8), model-based difference between groups: -1.0 (95% CI, -2.8 to 0.7), p=0.23	SF-36 physical function (0 to 100): 14.3 (22.6) vs. 7.5 (27.4); difference: 6.8, 95% CI -8.5 to 22.0, p=0.38	HADS-Depression, change from baseline: -1.6 (2.9) vs. -1.9 (3.0), p=0.67 HADS-Anxiety: -3.2 (2.2) vs. 2.0 (3.2), p=0.24 Brief pain inventory (0 to 10): Average pain severity, mean (SD): -1.6 (1.5) vs. -0.8 (2.3): 0.73 (95% CI, 0.54 to 1.00), p=0.05 Average pain interference, mean (SD): -1.9 (1.3) vs. -1.1 (2.8): 0.70 (95% CI, 0.51 to 0.96), p=0.03 CDC Symptom Inventory, CFS Questions: mean change (SD): -9.7 (13.1) vs. -8.2 (14.6), between-group difference at endpoint: -1.5 (95% CI, -9.9 to 6.9), p=0.72

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Blacker, 2004[108] Fukuda	A: Galantamine 7.5 mg (89) B: Galantamine 15 mg (86) C: Galantamine 22.5 mg (91) D: Galantamine 30 mg (86) E: Placebo (82) Duration of treatment: 4 months Duration of follow-up: 4 weeks	Chalder Fatigue Scale (mean change from baseline) Physical: 9.25 vs. 8.77 vs. 11.02 vs. 9.99 vs. 9.86, no significant differences Mental: 6.46 vs. 5.89 vs. 7.74 vs. 6.60 vs. 6.80, no significant differences	Not reported	Pittsburgh Sleep Quality Index Total score (0-21, higher score indicates worse sleep): -1.60 vs. -2.28 vs. -1.43 vs. -1.73 vs. -2.02 , no significant differences
Blockmans, 2003[111] Fukuda	A. Hydrocortisone 5 mg daily + 9-alpha fludrocortisone 50 µg daily (50) B. Placebo Duration of treatment: 3 months + 3 months crossover (50) Duration of follow-up: end of 3-month crossover	VAS(0 to 10), mean (SD): 6.6 (2.0) vs. 6.7 (2.1), p=0.76 Short fatigue questionnaire score: 8 (5) vs. 7 (5), p=0.69	SF-36 physical function (0 to 100): 31.7 (18.2) vs. 30.4 (18.1), p=0.34	Depression: HADS depression (0 to 21): 8 (5) vs. 9 (4), p=0.04, but not significant after Bonferroni correction Anxiety: HADS anxiety (0 to 21): 9 (4) vs. 10 (4), p=0.28
Fluge, 2011[94] Fukuda	A: Rituximab 500 mg/m^2, maximum 1,000 mg (15) B: Placebo (15) Duration of treatment: 2 weeks Duration of follow-up: 12 months	Fatigue: Major clinical responses: 9 (60%) vs. 7 (7%), p=0.002 Moderate clinical responses: 1 (7%) vs. 1 (7%) Overall, 95% CI: 10 (67%) (95% CI, 41% to 85%) vs. 2 (13%) (95% CI, 4% to 38%), p=0.003 Response duration: weeks, mean (range): 25 (8 to >44), n=10 vs. 41 (34 to >48), n=2 Difference between groups in self-reported fatigue score at 40 to 52 weeks: 0.63 (95% CI, -0.09 to 1.34), adjusted p value: 0.25 Difference in physician-assessed fatigue score at 12 months after intervention: 0.62 (95% CI, -0.09 to 1.34), adjusted p-value: 0.17	SF-36 physical function, (percent, lower score denotes increasing symptoms), max change %, mean (SD): 39 (33) vs. 11 (22)	NR

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Fluge, 2019[95] Canadian consensus (Carruthers, 2003)	A: Rituximab 500 mg/m^2, maximum 1,000 mg (77) B: Placebo (75) Duration of treatment: 12 months Duration of follow-up: 24 months	Fatigue: Fatigue score (0 to 6), at 16 to 20 months: 3.12 vs. 3.18, mean difference: -0.06 (95% CI, -0.51 to 0.39), p=0.79 Fatigue Severity Scale (9 to 63), mean at 18 months: 55.98 vs. 56.05, mean difference: -0.07 (95% CI, --3.21 to 3.08), p=0.68	Overall Function: SF-36 physical function(0 to 100) at 18 months: 45.67 vs. 45.23, mean difference: 0.42 (95% CI, -8.12 to 8.96), p=0.52 Function level, % at 16 to 20 months: 25.25 vs. 25.93, mean difference: -0.68 (95% CI, -5.90 to 4.54), p=0.31	Mean steps per 24 hours, 17 to 21 months: 3,777 vs. 3,904, mean difference: -127 (95% CI, -1004 to 749), p=0.58
McKenzie, 1998 [96] Fukuda and Holmes	A. Hydrocortisone 20-30 mg every morning, 5 mg every evening (35) B. Placebo (35) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	POMS, mean change: fatigue subscale: -3.6 (5.3) vs. -1.8 (4.5); p=0.21 POMS vigor subscale: 1.2 (3.3) vs. 0.7 (3.3); p=0.45	Activity Scale, mean change: 0.3 (1.1) vs. 0.7 (1.4); p=0.32	Beck Depression Inventory (0-63, higher most severe) change: -2.1 (5.1) vs. -0.4 (4.1); p=0.17
Montoya, 2018[109] Fukuda	A. Methylphenidate 20mg/d + Mitochondrial nutritional supplement: 4 tablets twice daily. (67) B. Placebo (68) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	CIS total score (20 to 140): 95.3 vs 98.6, mean change from baseline: -16.9 (±23.52) vs. -13.8 (±22.15), (95% CI, -11.1 to 4.0), p=0.359 VAS fatigue change from baseline: -18.2 mm (±25.05) vs. -11.1 mm (±22.08), (95% CI, -11.5 to 2.3), p=0.189	NR	NR
Montoya, 2013[98] Fukuda	A. Valganciclovir 900 mg BID for 21 days, then 900 mg daily (20) B. Placebo (10) Duration of treatment: 6 months Duration of follow-up: 12 months	Fatigue Severity Scale 9-item (1 to 7) (change in score, negative indicates better health): -0.06 vs. 0.02; p=0.006 MFI-20 (20 to 100), change in score: -6.15 vs. -1.10; p=0.224	Self-reported physical function: 1.02 vs. 0.46; p=0.217 CDC Symptom Inventory: NS	NR
Peterson, 1990[105] Holmes	A. IgG 1 g/kg IV every 30 days (15) B. Placebo (15) Duration of treatment: 6 months Duration of follow-up: 6 months	NR	MOS-SF social function higher in placebo group: 5.2 (5.5) vs. 9.4 (7.9); p<0.05 MOS-SF physical function (0 to 100): 56.0 (23.2) vs. 51.8 (22.2); p=NS	NR

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Roerink, 2017[106] Fukuda	A. Anakinra 100 mg SQ daily (25) B. Placebo (25) Duration of treatment: 4 weeks Duration of follow-up: 20 weeks after treatment	CIS-fatigue score: 4 weeks: 46.7 vs. 45.1, p=0.59 24 weeks:45.3 vs. 44.0, p=0.69	SF-36 physical function (0 to 100): 4 weeks: 58.2 vs. 61.2, p=0.53 24 weeks: 60.8 vs. 64.8, p=0.47 Sickness Impact Profile (SIP-8) (0 to 5,799): 4 weeks: 1472.2 vs. 1353.7, p=0.47 24 weeks: 1351.5 vs. 1260.4, p=0.62	Psychological symptoms: SCL-90 (90 to 450): 4 weeks: 144.4 (136.6 to 152.2) vs. 139.9 (132.1 to 147.7), p=0.42 24 weeks: 143.5 (135.3 to 151.7) vs. 140.5 (132.3 to 148.7), p=0.63 Pain (VAS): 4 weeks: 7.4 (6.5 to 8.3) vs. 6.3 (5.4 to 7.2), p=0.104 24 weeks: 6.9 (5.9 to 7.9) vs. 6.6 (5.6 to 7.6), p=0.63
Rowe, 1997[107] Fukuda	A. IgG 1 gm/kg IV every 30 days (36) B. Placebo (35) Duration of treatment: 12 weeks Duration of follow-up: 25 weeks follow-up	NR	Investigator scale; % of normal 3 months: Not improved (<25% improvement)/ Improved (>25% improvement) %: NS 6 months: Not improved (<25% improvement).%: 27.8 (10/36) vs. 55.9 (19/34); p=0.02 (RR 0.50, 95% CI 0.21 to 0.91) Improved (>25% improvement) %: 72.2 (26/36) vs. 44.1 (15/34) p=0.02 (RR 1.64, 95% CI 1.07 to 2.51) Returned to full function (not defined) at 6 months, %: 25 (9/36) vs. 11 (4/34), p<0.04	Depression and Anxiety: SCL-90-R (90 to 450): NS
See, 1996[99] Holmes	A. Alfa-2a Interferon 3 mu SQ 3 times per week (15) B. Placebo (15) Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	NR	NR	NR

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Strayer, 1994[100] Fukuda and Holmes	A. Rintatolimod 200 mg IV twice weekly 4 times, then 400 mg twice weekly (45) B. Placebo (47) Duration of treatment: 6 months Duration of follow-up: 6 months	NR	Exercise duration (% change from baseline): 10.3 vs. 2.1; p=0.007 Exercise work (% change from baseline): 11.8 vs. 5.8; p=0.011 ADL score (% change from baseline): 23.1 vs. 14.1; p=0.034 KPS score (% change from baseline): +20 vs. 0; p=0.023	NR
Strayer, 2012[101] Fukuda and Holmes	A. Rintatolimod 400 mg IV twice weekly (117) B. Placebo (117) Duration of treatment: 40 weeks Duration of follow-up: 40 weeks	NR	Cardiopulmonary exercise tolerance % change from baseline: 36.5% vs. 15.2%; p=0.047	NR
Sulheim, 2014[102] Fukuda	A. Clonidine 25 -50 mcg based on weight. (60) B. Placebo (60) Duration of treatment: 9 weeks Duration of follow-up: 30 weeks	Chalder Fatigue Scale 11-item (0 to 33) at 30 weeks: 11.1 vs. 13.5, difference 0.5, 95% CI: -14.7 to 15.7, p=0.95	Mean Functional Disability Inventory (0 to 60) at 30 weeks: 17.5 vs. 16.8, difference 0.2, 95% CI: -13.3 o 13.6, p=0.98	Pain Brief pain inventory (0 to 10): 8 weeks: 4.1 vs. 3.4, p=0.14 30 weeks: 3.8 vs. 3.3, p=0.32 Sleep (KSQ Insomnia Score): 8 weeks: 3.7 vs. 3.8, p=0.54 30 weeks: 3.6 vs. 3.6, p=0.74
Vercoulen, 1996[103] Oxford	A. Fluoxetine 20 mg/d (54) B. Placebo (53) Duration of treatment: 8 weeks Duration of follow-up: 10 weeks	Daily observed fatigue score: NS	Self-reported change: NS	Depression: BDI (0 to 63, mean difference): -0.186 (95% CI, 0.35 to 0.02), p=NS

Author, Year ME/CFS Criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Function Outcomes	Other Outcomes (depression, sleep, pain, etc.)
Vollmer-Conna, 1997[104] Fukuda	A: IgG 0.5 gm/kg (22) B: IgG 1.0 mg.kg (28) C: IgG 2.0 mg/kg (23) D. Placebo (26) IV every 30 days Duration of treatment: 12 weeks Duration of follow-up: 12 weeks	POMS energy score: No significant difference between groups, data NR	KPS (0 to 100): median (1st to 3rd IQR): 80.0 (80 to 70) vs. 80.0 (80 to 70) vs. 75.0 (80 to 70) vs. 77.5 (80 to 70), difference in change between groups: p>0.13	NR
Wearden, 1998[59] Oxford	A. Fluoxetine 20 mg daily (35) B. Placebo (34) Duration of treatment: 26 weeks Duration of follow-up: 26 weeks	Chalder Fatigue Scale 14-item (0 to 42) (mean change from baseline): 12 weeks: -1.6 (-4.4 to 1.2) vs. -2.0 (-4.1 to 0.1) 26 weeks: -3.0 (-5.9 to -0.2) vs. -2.7 (-5.4 to 0.01) Chalder Fatigue Scale (cases of non-fatigue): 12 weeks: 1 (3/35) vs. 6 (2/34) 26 weeks: 6 (2/ 35) vs. 6 (2/34) Exercise improved Chalder Fatigue Scale scores, mean change: 12 weeks: 2.1 (95% CI -0.6 to 4.8), p=0.13 26 weeks: 2.9 (95% CI -0.2 to 6.1), p=0.07	Functional work capacity (mean change): 12 weeks: 0.4 (-1.2 to 2.0) vs. 0.4 (-0.9 to 1.7) 26 weeks: 1.0 (-0.9 to 3.0) vs. -0.1 (-1.7 to 1.6)	Depression: HADS depression (0 to 21): Week 12: -1.1 (95% CI -0.03 to -2.2; P=0.04) Week 26: -1.7 (-3.0 to -0.5) vs. -1.3 (-2.3 to -0.3)

Abbreviations: ADL = activities of daily living; BDI = Beck Depression Inventory; BID = twice daily; BPI = Brief Pain Inventory; CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; CI = confidence interval; HADS = Hospital Anxiety and Depression Scale; HADS-A = Hospital Anxiety and Depression Scale-Anxiety; HADS-D = Hospital Anxiety and Depression Scale-depression; IgG = immunoglobulin G; IV = intravenous; KPS = Karnofsky Performance Scale; KSQ = Karloinska Sleep Questionnaire; MDD = major depressive disorder; ME = myalgic encephalomyelitis; MFI-20 = Multidimensional Fatigue Inventory 20-item; MOS-SF = Medical Outcome Study-short form; NR = not reported; NS = not significant; POMS = profile of mood states; q = every; RCT = randomized controlled trial; RR = relative risk; SCL-90 = symptom checklist 90; SCL-90-R = symptom checklist 90-revised; SD = standard deviation; SF-36 = 36-item Short Form Health Survey; SIP-8 = Sickness Impact Profile; SQ = subcutaneous; TID = three times daily; VAS = visual analogue scale

Immune Modulators

Rintatolimod versus placebo. Two moderate risk of bias trials (N=332) evaluated rintatolimod, a synthetic derivative of inosinic acid with antiretroviral and immunomodulatory activities, in adults who met the Holmes case definition or both the Holmes and Fukuda case definitions.[100,101] Both studies evaluated exercise-related outcomes as the primary outcome; the prevalence of post-exertional fatigue at baseline was not reported. Effects on fatigue severity were not directly measured in either trial.

In the first trial (N=92), patients with a mean baseline KPS score of 51 (scale of 0-100, range 20 to 60) were randomized to intravenous rintatolimod (200 mg twice weekly for 4 weeks, then 400 mg twice weekly for a total of 24 weeks) versus placebo.[100] Rintatolimod was associated with greater improvement from baseline to week 24 in exercise duration (10.3 minutes vs. 2.1 minutes; p=0.007), exercise work (11.8 Kcal vs. 5.8 Kcal; p=0.011), activities of daily living (ADL) (23.1 vs. 14.1; p=0.034), and KPS (20 vs. 0; p=0.023). In subgroup analysis, there was no difference based on presence of markers of human herpes virus-6 (HHV-6) virus reactivation at baseline (based on mononuclear cell evaluation). There were no serious adverse events or withdrawals due to adverse events. Insomnia was reported more frequently in the placebo group, and dry skin in the rintatolimod group (p<0.05 for both outcomes, data otherwise not reported).

A second trial (N=240) randomized patients with KPS scores of 40 to 60 (mean not reported) to rintatolimod 400 mg twice weekly for 40 weeks versus placebo.[101] Rintatolimod was associated with greater mean percentage change in exercise tolerance (based on treadmill testing duration) at week 40 versus placebo (37% vs. 15%; p=0.047). Although other function outcomes were measured, they were not compared between groups (and reported data were not adequate to evaluate differences). In the rintatolimod and placebo groups, the mean values at endpoint were: KPS: 55 versus 50 (0 to 100 scale), ADL: 72.4 versus 69.4 (higher values better, but scale range unclear), SF-36 vitality subscale: 10 versus 10, and SF-36 general health perception subscale: 20 versus 25 (both 0 to 100 scales). More participants in the treatment group reported decreased use of medications for relief of CFS symptoms (68% vs. 55%; p=0.048). Adverse events occurred more frequently with rintatolimod than placebo with infusion-related headache the most common adverse event (64% vs. 20%, P<0.01). Other adverse events reported more often with rintatolimod were flu-like syndrome, chills, vasodilatation, and dyspnea (p<0.05). Serious adverse events and withdrawals due to adverse events were not reported.

Immunoglobulin G versus placebo. Three moderate risk of bias trials (N=197) evaluated intravenous IgG (administered as a monthly infusion) versus placebo in patients with ME/CFS.[104,105,107] Two trials enrolled adults (one using the Holmes case definition[105] and the other using the Fukuda case definition,[104] mean age 38 and 40 years, mean duration of ME/CFS 3.8 and 6.3 years) and one trial adolescents (Fukuda case definition, mean age 15 years, mean duration of ME/CFS 18 months).[107] The duration of treatment ranged from 12 to 24 weeks and the specific product was Gammimune N® or Gammagard®. All three trials used a 1 gm/kg dose, with one study also evaluating a 0.5 gm/kg and a 2gm/kg dose. The timing of outcome assessment ranged from the end of treatment to 6 months following completion of therapy. Baseline fatigue was 18.8 on the POMS fatigue subscale (0 to 28) in the one trial;[104] fatigue was not a reported outcome and baseline fatigue not reported in the other trials.[105,107] Regarding baseline function, one trial of adults reported mean SF-36 physical function score of 54 (0 to 100 scale), [105] and the other trial of adults reported a mean KPS score of 70 (0 to 100 scale).[104] The trial of adolescents used a non-validated measure of function, with a mean baseline score of 25%.[107]

In the two trials of adults, there were no statistically significant differences between IgG (any dose) versus placebo in severity of functional impairment at any follow-up timepoint.[104,105] One trial also found no effects on severity of fatigue or quality of life outcomes.[104]

In the trial of adolescents, using an unvalidated 0% to 100% scale, there was no difference between groups in mean function at end of treatment (3 months; 49.9% vs. 44.6%, RR 1.1, 95% CI 0.84 to 1.45) or at 3 months after end of treatment (64.1% vs. 52.1%, mean difference -12.0, 95% CI -26.1 to 2.12).[107] A subgroup analysis found no difference in effects based on the

duration of ME/CFS symptoms. The proportion of patients with at least 25% improvement in function was not different between groups at end of treatment (3 months, 52% vs. 31%, RR 1.67, 95% CI 0.94 to 3.0), but was significantly greater with IgG at 3 months after end of treatment (72% vs. 44%, RR 1.64, 95% CI 1.07 to 2.5).

In adolescents, IgG infusion was associated with increased likelihood of severe headache versus placebo following first infusion (64% vs. 20% RR 3.14 95% CI 2.09 to 4.73) following first infusion; withdrawals due to adverse events or serious adverse events were not reported.[107] In adults, the study of 1 gm/kg infusions found IgG associated with increased risk of severe infusion-related headache (93% vs. 60%, RR 1.56, 95% CI 1.0 to 2.4).[105] Withdrawals due to adverse events (13% vs. 13%, RR 1.00, 95% CI 0.16 to 6.2) and serious adverse events (13% vs. 20%, RR 0.67, 95% CI 0.13 to 3.4) were not different between groups. The dose-ranging study of IgG in adults found similar incidence of "constitutional symptoms" including headache, fatigue, malaise, and concentration problems across IgG doses and placebo, with 71% versus 88% in the 1 gm/kg vs placebo groups (RR 0.81, 95% CI 0.62 to 1.06).[104] More withdrawals due to adverse events occurred in the IgG groups than placebo, but the estimate was very imprecise and not statistically significant (5.7% vs. 0%, RR 10.1, 95% CI 0.57 to 178.5).

Rituximab versus placebo. Two trials (N=181) evaluated the anti-CD20 monoclonal antibody rituximab, which results in depletion of B-lymphocytes, versus placebo infusions in adults who met the Fukuda case definition[94] or the Canadian consensus criteria.[95] Dosing in the earlier, medium risk of bias, pilot study was two infusions of 500 mg/m^2 or saline given two weeks apart with 12 months of follow-up,[94] while in the second, low risk of bias trial the same initial dosing was used, followed by fixed-dose infusions of 500 mg at 3, 6, 9, and 12 months.[95] Mean baseline fatigue was 8.0 on a self-rated 1 to 10 severity scale in the pilot study (based on scoring of fatigue, post-exertional exhaustion, need for rest, and daily functioning) and 59.6 on the Fatigue Severity Scale (range 9 to 63) in the subsequent study. Mean function at baseline in the second trial was 19 on a 0 to 100% scale, and 34 on the SF-36 Physical Function scale (range 0 to 100).[95] The pilot study reported on chronic fatigue symptoms at baseline, with a mean of 8.1 on a 1 to 10 self-assessed scale.[94]

In the pilot study (N=30), the primary endpoint of CFS symptoms (Fatigue scores) at 3 months was not significantly different between groups when assessed by the patients (mean difference 0.00, 95% CI -0.31 to 0.31) or by a physician (mean difference 0.13, 95% CI -0.35 to 0.61);[94] however, a significant interaction was found for symptom scores based on an analysis of intervention and time (p=0.018). Overall response was defined as a fatigue score of ≥4.5 for at least 6 consecutive weeks, further categorized as major or moderate. A major response required some fatigue symptoms rated as having major improvement by the patient (6 points on a 0 to 6 scale). Rituximab was associated with increased likelihood of experiencing an overall response (67% vs. 13%, RR 5.0, 95% CI 1.31 to 19.1). Most of the patients with response met criteria for a major response (60% vs. 7%, RR 9.00, 95% CI 1.30 to 62.51). The mean response duration was 25 weeks (range 8 to >44) with rituximab and 41 weeks (range 34 to >48) with placebo. Function, as assessed by the SF-36 Physical health summary score was significantly improved in the rituximab group compared with the placebo group (mean maximum change 54% vs. 26%, mean difference 28%, 95% CI 1.8% to 54%). The SF-36 mental health component score was not significantly different between groups (mean maximum change 9% vs. 5%, mean difference 4%, 95% CI -38% to 29%). This study reported no withdrawals due to adverse events, or serious adverse events. Overall, there were similar numbers of patients reporting infusion-related adverse events (33% vs. 27%, RR 1.25, 95% CI 0.42 to 3.77).

The second, larger (N=151) RCT of rituximab, which added additional infusions every 3 months after the initial set of two infusions, did not find rituximab associated with increased likelihood of a response using a similar definition as the pilot study, but requiring 8 rather than 6 consecutive weeks of improvement (35% vs. 26%, mean difference 9.2%, 95% CI -5.5 to 23.3).[95] There was also no difference in mean fatigue scores over 24 months (mean difference 0.02, 95% CI -0.27 to 0.31) and no effect on secondary measures, including the SF-36 Physical Function score (mean difference -0.41, 95% CI -7.73 to 6.92), function level (0 to 100% scale, mean difference -0.21%, 95% CI -4.18% to 3.76%), the Fatigue Severity Scale (mean difference, -0.25, 95% CI -2.44 to 1.95) or the number of steps per 24 hours (mean difference -177, 95% CI -1004 to 749). No patients withdrew due to adverse events, and 26% in the rituximab group versus 19% in the placebo group had serious adverse events (RR 1.37, 95% CI 0.75 to 2.5). Infusion-related adverse events were reported in 10% of rituximab patients and zero placebo patients (RR 26, 95% CI 1.57 to 429).

Anakinra versus placebo. One low risk of bias trial (N=50) compared anakinra, an interleukin-1 receptor antagonist (100 mg subcutaneously for 28 days) versus placebo in adults who met the Fukuda case definition.[106] Baseline fatigue was 51 on the CIS fatigue subscale (range 8 to 56), and mean baseline functional impairment based on the Sickness Impact Profile was 1677 (0 to 5799 scale). There were no differences between anakinra versus placebo in SF-36 physical function, SIP functional impairment, and the CIS fatigue subscale at the end of treatment or at 24 weeks. There were also no differences between groups in psychological symptoms measured using the SCL-90 or pain measured on a visual analog scale.

Anakinra was associated with increased risk of any adverse event versus placebo (95% vs. 56%, RR 1.71 95% CI 1.20 to 2.45), with one patient discontinuing treatment in the anakinra group (4% vs. 0%, RR 3.00, 95% CI 0.13 to 70.30). There were no serious adverse events. Anakinra was also associated with increased likelihood of injection site reactions (68% vs 4%, RR 17.00 95% CI 2.44 to 118.20) and infections (24% vs. 16%, RR 1.50, 95% CI 0.48 to 4.78).

Alfa-2a Interferon versus placebo. One small (N=26), high risk of bias, crossover trial compared alfa-2a interferon (3 million units subcutaneously three times per week for 12 weeks) versus placebo in adults who met the Holmes case definition.[99] Quality of life was assessed using a 10-item clinical well-being scale addressing many symptoms found in ME/CFS patients (fatigue, fevers, sore throat, lymphadenopathy, muscle aches, headaches, joint pains, depression, concentration, and insomnia, range 0 to 60, with lower score representing greater well-being). The mean score at baseline was 35.7. After 12 weeks there was no significant difference between groups (31.4 vs. 28.4, mean difference 3.0, 95% CI -5.6 to 11). This study reported the results for interferon from both the initial and crossover phase (n=26), but only reported results from the initial phase for placebo (n=13). No other clinical outcomes were reported. A subgroup analysis found that patients with NK cell dysfunction (N=10) at baseline experienced improvement in quality of life with interferon (mean difference 23.4, 95% CI -35.3 to -11.5). However, this was a very small subgroup, with only 7 interferon patients and 3 placebo patients.

Although no serious adverse events were reported, 27% of interferon patients withdrew due to adverse events, compared with none in the placebo group (RR 9.00 95% CI 0.53 to 151.95). Adverse events in the interferon group included flu-like syndrome (27% vs. 0%, RR 9.00 95% CI 0.53 to 151.95) and diarrhea (13% vs. 0%, RR 5.00 95% CI 0.26 to 95.02).

Antidepressants

Four moderate risk of bias RCTs (N=285) evaluated antidepressants.[59,75,103,110] Two trials evaluated the SSRI antidepressant 20 mg daily of fluoxetine,[59,103] one evaluated the serotonin-norepinephrine reuptake inhibitor duloxetine at 120 mg daily,[110] and one evaluated the noradrenergic and specific serotonergic mirtazapine at 15 mg to 45 mg daily.[75]

Fluoxetine versus placebo. Two fluoxetine trials enrolled adults meeting the Oxford case definition for ME/CFS, but were too heterogeneous to combine.[59,103] In one trial, 10% of patients had major depressive disorder (MDD) at baseline, with a mean HADS-depression scale score of 8.8 (0 to 21 scale) and mean fatigue severity score of 34 (14-item 0 to 42 Chalder scale).[59] After 6 months of treatment, there was no statistically significant difference between groups in Chalder fatigue scores (-0.30, 95% CI -4.3 to 3.7) or functional work capacity (-1.1, 95% CI -3.7 to 1.5). There was also no difference between groups in depression severity (mean difference 0.40, 95% CI -1.23 to 2.03) on the 0 to 21 HADS-depression scale. In the second trial, randomization to 8 weeks of fluoxetine or placebo was stratified by presence/absence of MDD, and mean baseline fatigue was 9.4 on the Subjective Daily Observed Fatigue Scale (0 to 16 scale).[103] Baseline function was not reported in either trial. Although randomization was stratified by presence of depression, main results were not stratified by depression status. However, graphical presentation of results stratified by depression status showed very similar findings. The difference between groups at 8 weeks was not significant for fatigue (mean difference -0.16, 95% CI -0.64 to 0.31). While the effect on depression severity was statistically significant (mean difference -0.19, 95% CI -0.35 to -0.0 on the 0 to 63 Beck Depression Inventory), the difference was very small (less than 0.25 points).[103] More patients in the fluoxetine groups reported deterioration in CFS symptoms, but the differences were not statistically significant (patients with major depression 38% vs. 26%, RR 1.46, 95% CI 0.6 to 3.5 and patients without major depression 35% vs. 14%, RR 2.43, 95% CI 0.84 to 7.07).

In both trials, fluoxetine was associated with increased risk of withdrawal due to adverse events (13% vs. 4%, RR 3.93, 95% CI 0.87 to 17.64[103] and 15% vs. 3%, RR 4.37, 95% CI 1.02 to 18.78[59]). One trial found fluoxetine associated with increased risk of tremor (67% vs. 40%, RR 1.57 95% CI 0.87 to 2.83), perspiration (40% vs. 26%, RR 1.70 95% CI 1.14 to 2.53), discontinuations due to skin reactions (6% vs. 2%, RR 2.94 95% CI 0.32 to 27.42), and headache (2% vs. 4%, RR 1.96 95% CI 0.18 to 21.01).[103] The other trial did not report specific adverse events.[59] Serious adverse events were not reported in either trial.

Duloxetine versus placebo. One trial (N=57) compared 12 weeks of duloxetine versus placebo in patients who met the Fukuda case definition.[110] Patients with major depression were excluded (mean HADS depression score 9.27 on a 0 to 21 scale). At baseline, the mean CDC Fukuda CFS case definition symptom score was 40 (0 to 152 scale) and mean SF-36 physical function score was 59.5 (0 to 100 scale). At 12 weeks, there was no difference in function based on the SF-36 physical function subscale score (mean difference -2.7, 95% CI -15.5 to 10.1) or other SF-36 subscales. Fatigue was also not significantly different between groups, based on the MFI general fatigue scale or subscales on physical fatigue, reduced activity, or reduced motivation subscales. The mental fatigue subscale showed more change in the duloxetine group, but the difference was very small (-0.1, 95% CI, -0.3 to 0.0 on a 4 to 20 scale). There was no significant difference in improvement in the CDC Symptom Inventory overall or for CFS symptoms (mean difference -1.5, 95% CI -9.9 to 6.9). There were also no significant differences between groups in depression or anxiety (mean difference -0.9, 95% CI -2.4 to 0.6 on the HADS anxiety scale and 0.94, 95% CI 0.72 to 1.23 on the HADS depression scale). Patient assessments

of improvement in disease severity were greater with duloxetine (-1.1 vs. -0.4, p=0.06. 1 to 7 scale). Duloxetine was associated with decreased pain severity (mean difference -0.73, 95% CI -1.00 to -0.54 on the 0 to 10 BPI pain intensity scale) and pain interference (mean difference -0.70, 95% CI -0.96 to -0.51 on the 0 to 10 BPI pain interference scale).

Duloxetine was also associated with increased risk of withdrawal due to adverse events versus placebo (10% vs. 0%, RR 7.00 95% CI 0.38 to 129.93).[110] One patient assigned to duloxetine had suicidal ideation; no other serious adverse events were reported. Duloxetine was associated with increased likelihood of dry mouth (21% vs. 3.3%, RR 6.21, 95% CI 0.80 to 48.4).

Mirtazapine versus placebo. One trial (N=49) compared mirtazapine versus placebo in patients who met the Oxford or Fukuda case definitions (results for CBT arm reported in the CBT section). It did not report the proportion of patients with major depression; the mean HADS depression score at baseline was 14.51 (0 to 21 scale).[75]At baseline, the mean fatigue score was 24.97 on a 0 to 100 scale, and a mean score of 28.94 on the SF-36 Physical Function scale (0 to 100). After 12 weeks, there were no differences between mirtazapine and placebo in fatigue severity (mean difference 1.00, 95% CI -2.10 to 4.1) or depression severity (mean difference 1.2, 95% CI -2.7 to 5.1). Effects on function were not reported.

Mirtazapine was associated with increased risk of any adverse event versus placebo (100% vs. 45%, RR 2.18, 95% CI 1.41 to 3.37). Sedation was the most common adverse event in patients randomized to mirtazapine (proportion not reported in placebo group). Withdrawal due to adverse events and serious adverse events were not reported.

Corticosteroids versus placebo. Two moderate risk of bias trials evaluated corticosteroids versus placebo in adults meeting the Fukuda case definition.[96,111] One parallel group trial (N=70) compared oral hydrocortisone (20-30 mg am and 5 mg pm for 12 weeks) versus placebo[96] and one crossover trial (N=100) compared hydrocortisone (5 mg daily) plus 9-alpha fludrocortisone (50 µg daily) for 12 weeks versus placebo.[111] Neither trial reported statistically significant differences between corticosteroids versus placebo in fatigue or function. In the hydrocortisone (only) trial, the mean difference between groups in change in score on a 10-point activity scale was -0.4 (p=0.32), and -1.9 on the POMS Fatigue subscale (range 0 to 28, p=0.21).[96] The trial of hydrocortisone/fludrocortisone also did not find a significant difference in fatigue (mean difference on a 0 to 100 visual analogue scale [VAS] 0.1, 95% CI -0.3 to 0.6 and Abbreviated Fatigue Questionnaire -1, 95% CI -2 to 1, 7-point scale) or on the SF-36 Physical Component Summary scale (mean difference -1.3, 95% CI -4.7 to 2.1, 0 to 100 scale).[111] There was no correlation between cortisol levels at baseline or during treatment or follow-up and the primary outcome of Global Wellness.[96]

In the hydrocortisone/fludrocortisone trial, one patient withdrew from the steroid arm due to acne and weight gain.[111] Stimulated cortisol was significantly suppressed with treatment compared with placebo (mean difference 127 nmol/L, 95% CI 81 to 171). Adverse events were not otherwise reported. In the hydrocortisone trial, the steroid group had increased incidence of suppression of adrenal glucocorticoid responsiveness (34% vs. 0%, RR 1.00, 95% CI 1.54 to 406.5); increased appetite (49% vs. 23%, RR 2.12, 95% CI 1.06 to 4.27), weight gain (54% vs. 23%, RR 2.38, 95% CI 1.21 to 4.69); and difficulty sleeping (49% vs. 23%; RR 2.12, 95% CI 1.06 to 4.27).[96]

Other Drugs

Valganciclovir versus placebo. One moderate risk of bias trial (N=30) evaluated the antiviral medication valganciclovir in patients with suspected viral onset of ME/CFS, based on history and presence of elevated HHV-6 or Epstein-Barr Virus antibody titers.[98] Patients met the Fukuda case definition for ME/CFS. Patients were randomized to oral valganciclovir (900 mg twice daily for 21 days, then 900 mg once daily for a total of 6 months) versus placebo, with final outcomes measured at 9 months. Baseline fatigue was 78.6 on the Multidimensional Fatigue Inventory 20-item (MFI-20, 20 to 100 scale). Function was not reported. Valganciclovir was associated with decreased fatigue severity versus placebo based on the Fatigue Severity Scale (mean difference -0.06 vs. 0.02, p=0.006 for interaction of time and study arm); however, the difference was small (<0.1 point on a 9 to 63 scale). No statistically significant differences were found between valganciclovir and placebo on the MFI-20 total score (-0.88 vs. 0.29, p=0.11), MFI-20 mental fatigue subscale (-0.27 vs. -0.05, p=0.05), the CDC Symptom Inventory total scores (-2.63 vs. -2.69 on a 0 to 304 scale, range 0 to 304, p=0.96), the sleep assessment questionnaire (-0.17 vs. -0.14 on a 0 to 68 scale, p=0.86), and HADS depression (typical: 0.01 versus -0.14, p=0.66 and atypical: 0.07 vs. 0.04, p=0.54). Valganciclovir was associated with greater improvement in self-reported cognitive functioning (1.72 vs. 0.59, p=0.02), but not self-reported physical functioning (1.02 vs. 0.46, range 1% to 100%, p=0.22). The study evaluated the effects of treatment on monocyte counts but did not evaluate whether there was a subgroup effect according to baseline levels.

Valganciclovir was not discontinued due to hematologic or hepatic adverse events. There were two serious adverse events (cancer diagnosis) in the valganciclovir group that were deemed unrelated to the medication.

Galantamine versus placebo. One moderate risk of bias trial (N=423) evaluated galantamine (an acetyl-cholinesterase inhibitor) at various doses (7.5, 15, 22.5, or 30 mg daily for 16 weeks) versus placebo in patients who met the Fukuda case definition.[108] It was the largest of the medication trials. Baseline function was 13.5 on the Fibromyalgia Impact Questionnaire-Physical scale (0 to 10 scale); baseline fatigue was not reported. The primary outcome was response, defined as a score of zero or one (very much or much improved) on the Clinical Global Impression scale, with a clinically important difference defined as at least a 25% improvement. There were no differences between galantamine versus placebo in likelihood of a response (35% to 45% for galantamine at various doses, vs. 30% for placebo, p>0.05 for each galantamine dose vs. placebo). The study also reported no statistically significant differences between groups on other outcomes. The change from baseline on the Chalder fatigue 14-item scale physical subscale ranged from 8.77 to 11.02 with galantamine and was 9.86 with placebo; for the mental subscale the change from baseline ranged from 5.80 to 7.74 with galantamine and was 6.80 for placebo. There were also no differences in quality of life measured by the Nottingham Health profile or sleep quality based on the Pittsburgh Sleep Quality Index.

The likelihood of withdrawal due to adverse events was similar between the lowest dose galantamine (7.5 mg/d) and placebo (14% vs. 15%), but higher in the other doses of galantamine (23%, 24%, 26%). Serious adverse events were reported in 8 of 352 (2.3%) patients assigned to galantamine (any dose), with none in the placebo group. The most common adverse events were nausea, headache, and symptoms of depression in both galantamine and placebo groups. One patient committed suicide (galantamine 10 mg daily), and three others had suicidal ideation (1 each in the galantamine 7.5 mg, 22.5 mg and placebo groups).

Clonidine versus placebo. One low risk of bias trial (N=120) evaluated clonidine (25 or 50 mcg based on body weight) versus placebo in adolescents (mean age 15 years) who met the Fukuda case definition.[102] Baseline fatigue was 19.2 on the 11-item 0 to 33 Chalder scale, and baseline function was 24 on the Functional Disability Inventory scale (0 to 60 scale). There were no significant differences between clonidine versus placebo in fatigue, function, CFS hypersensitivity symptoms, insomnia or pain after 8 weeks of treatment or at 30 weeks of follow-up. The mean differences at these time points on the Chalder fatigue 11-item scale (0 to 44) were 1.7, 95% CI −2.3 to 5.6 and 0.5, 95% CI −14.7 to 15.7. The mean differences on the Functional Disability Inventory (0 to 60) scale were 0.2, 95% CI −10.3 to 10.8 and 0.2, 95% CI −13.3 to 13.6. The mean differences on the CFS symptom inventory hypersensitivity subscale score (0 to 10) were 0.1, 95% CI -0.2 to 0.5 and −0.03, 95% CI −0.4 to 0.3. The mean differences on the Karolinska Sleep Questionnaire insomnia subscale (range 1 to 6) were 0.1, 95% CI −0.4 to 0.2 and 0.1, 95% CI −0.3 to 0.4. The mean differences on the BPI (range 0 to 10) were 0.5, 95% CI −0.16 to 1.16 and 0.4, 95% CI −0.4 to 1.1. There was also no significant difference in the primary outcome of change in number of steps daily (51 vs. -560, mean difference at 8 weeks (−637, 95% CI −1328 to 53) and (119, 95% CI −796 to 1035) at 30 weeks of follow-up.

The trial did not report withdrawal due to adverse event or serious adverse events. There was no statistically significant difference in risk of any adverse event (75% vs. 65%, RR 1.30, 95% CI 0.99 to 1.72). Clonidine was associated with increased likelihood of dizziness (28% vs. 10%, RR 3.2, 95% CI 1.25 to 8.18).

Methylphenidate versus placebo. One moderate risk of bias trial (N=128) evaluated the stimulant methylphenidate (20 mg daily) given with a nutritional supplement designed to improve mitochondrial function (consisting of amino acids, vitamins, and other supplements) versus placebo.[109] Baseline fatigue severity was 112 on the 20 to 140 CIS scale; baseline function was not reported. At 12 weeks, there was no significant difference between methylphenidate versus placebo in fatigue severity. Pre-planned subgroup analyses evaluated patients with more severe ME/CFS symptoms and those taking analgesics at baseline. While both subgroups showed larger effects than the overall group, there were no statistically significant subgroup effects. Methylphenidate plus nutritional supplement was associated with increased likelihood of withdrawal due to adverse events (13% vs. 5%, RR 2.83, 95% CI 0.79 to 10.21) and dizziness (7% vs. <2%, RR 4.18, 95% CI 0.48 to 36.47).

Subgroup Effects

The ability to evaluate how effects of medications vary in subgroups was limited. The trials did not report subgroup analyses based on factors such as age, sex, race/ethnicity, and ME/CFS severity or type of onset. Most trials used the Fukuda or Holmes CDC case definitions for ME/CFS and there were too few trials of each medication to perform reliable cross-trial comparisons. A trial of alfa-interferon (rated high risk of bias) evaluated the subgroup effects related to baseline NK cell dysfunction, but the trial and subgroups had very small samples and no statistical test for subgroup effects was performed.[99] A trial of rintatolimod found no subgroup effect based on presence of markers of HHV-6 reactivation at baseline.[100] Among trials of IgG, [104,105,107]one trial of adolescents found an effect on function, [107] but two trials in adults did not. The small number of trials and heterogeneity (e.g., dosing and duration, type of IgG used eligibility criteria and outcome measures) precludes meaningful conclusions about differences in effects based on age. The trial in adolescents performed a stratified analysis based on the duration of ME/CFS symptoms and found no subgroup effect.[107] One trial of corticosteroids

found no correlation between cortisol levels at baseline and the primary outcome of Global Wellness.[96]

Complementary and Alternative Therapies

Dietary Interventions, Herbal Supplements, or Homeopathy

Nine trials evaluated dietary interventions, herbal supplements or homeopathy in adult patients with ME/CFS (**Tables 18 and 19, Evidence Table Appendix E2**). Sample sizes ranged from 14 to 268 (total N=739). Five trials evaluated various dietary changes or supplements. Three[112-114] trials compared dietary supplements versus placebo, and two trials compared one dietary supplement versus another (one of the trials also evaluated the combination of supplements). Four trials evaluated homeopathy or herbal supplements. One trial[115] compared melatonin supplements versus phototherapy, one trial[116] compared pollen versus placebo, one compared dengzhanshengmai herbal supplement versus placebo when used with an SSRI, and one compared[117] homeopathy versus placebo. Six trials were included in the prior report.[114-119] No trials were conducted in the United States, eight trials in Europe, and one trial in China. The mean age of participants ranged from 35 to 50 years in all but one trial of older patients (mean age 76.2 years) and the proportion female ranged from 49 percent to 100 percent. The case definition for ME/CFS was the Oxford criteria in two trials and the Fukuda criteria in 6 trials, and the Fukuda criteria or Holmes[10] criteria in one trial. The duration of ME/CFS ranged from 14.5 months to 6 years in four trials that reported this information. Baseline fatigue was measured using a variety of scales (**Table 18**). Details regarding the presence of post-exertional fatigue and activity patterns were lacking.

No trial was rated low risk of bias, four trials were rated medium risk of bias[112,114,117,119] and five trials[113,115,116,118,120] were rated high risk of bias (**Risk of Bias Table Appendix F**). Methodological limitations included failure to report attrition, inadequate description of randomization or allocation concealment methods, and failure to blind or unclear blinding status of outcomes assessors and data analysts.

Table 18. RCTs of dietary interventions, herbal supplements, or homeopathy: study characteristics

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Hobday, 2008[118] United Kingdom High	n: 39 Age: 44 vs. 42 % Female: 88 vs. 78	Criteria: Fukuda Duration: Not reported	Fatigue scale: Chalder Fatigue Scale 14-item (0 to 56) Baseline fatigue: 23.0 vs. 22.0	Baseline depression: HADS depression (0 to 21): 8.1 vs. 7.0 Baseline function: SF-36 physical function (0 to100): 34.6 vs. 38.7	A: Low sugar/low yeast B: Healthy eating Frequency: Daily Duration of treatment: 24 weeks Duration of follow-up: End of treatment

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Li, 2015[120] China High	n: 268 (unclear if 45 dropouts included in analyses) Age: 35.1 vs. 36.8 % Female: 59	Criteria: Fukuda Duration, months: 15.7 vs. 14.5	Fatigue scale: MFI-20 general fatigue subscale (4 to 20) Baseline fatigue: 10.7 vs. 10.2	Baseline depression: not reported Baseline function: not reported	A: Dengzhanshengmai herbal supplement 1.08g + SSRI, Seroxat 10 to 30 mg or Zoloft 25 to 100mg B: SSRI, Seroxat 10 to 30 mg or Zoloft 25 to 100mg Frequency: Daily Duration of treatment: 12 weeks Duration of follow-up: End of treatment
Malaguarnera, 2008[112] Italy Medium	n: 96 Age: 76.2 vs. 78.4 % Female: 49	Criteria: Fukuda or Holmes (CDC, 1998) Duration: Not reported	Fatigue scale: Fatigue Severity Scale 9 to 63) and Wessely and Powell Scales (8-item physical, 5-item mental, maximum score 26) Baseline fatigue: Fatigue Severity Scale (9 to 63), mean: 50.4 vs. 50.1 Wessely and Powell Scales: 13.4 vs. 13.1	Baseline depression: not reported Baseline function: PF-10 (0 to 100): 69.8 vs. 70.2	A: Acetyl L-carnitine 2 g B: Placebo Frequency: twice daily Duration of treatment: 180 days Duration of follow-up: End of treatment
Ockerman, 2000[116] Sweden High	n: 43 Age: 50 % Female: 86	Criteria: Fukuda Duration: Not reported	Fatigue scale: Unclear which scale was used, higher score indicates worse outcome Baseline fatigue: 7.95 vs. 7.32 (this includes participants at the start of each 3-month phase)	Baseline depression: unclear scale: 5.9 vs. 6.7 Baseline function: not reported	A: Pollen extract (Polbax), 7 tablets B: Placebo Frequency: Daily for 3 months, then crossover to other arm Crossover design: Pollen/Placebo, n=5 Placebo/Pollen, n=5 Pollen/Pollen, n=6 Placebo/Placebo, n=6 Duration of treatment: 3 months Duration of follow-up: End of treatment
Ostojic, 2016[113] Serbia High	n: 14 Age: 39.3 % Female: 100	Criteria: Fukuda Duration: not reported	Fatigue scale: MFI-20 general fatigue subscale (4 to 20) Baseline mean: 12.1	Baseline depression: not reported Baseline function: not reported	A: Guanidinoacetic acid supplement, 2.4 grams B: Placebo, cellulose Frequency: Daily Duration of treatment: 3 months, Duration of follow-up: End of first treatment period; 3 months after randomization

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
The, 2007[114] The Netherlands Medium	n: 57 Age: 40.9 vs. 43.4 % Female: 77 vs. 59	Criteria: Fukuda Duration: not reported	Fatigue scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) Baseline fatigue: 46.5 (7.4) vs. 46.2 (7.9)	Baseline depression: not reported Baseline function: Sickness Impact Profile (SIP-8) (0 to5,799): 1484 vs. 1317	A: Acclydine supplement, declining dose from 1000 mg daily to 250 mg/2 days + amino acid supplement B: Placebo acclydine + placebo amino acid supplement Frequency: Daily decreasing to every other day Duration of treatment: 14 weeks Duration of follow-up: End of treatment
Vermeulen, 2004[119] The Netherlands Medium	n: 89 Age: 42 vs. 37 vs. 38 % Female: 76	Criteria: Fukuda Duration, years: 6.0 vs. 5.5 vs. 3.0	Fatigue Scale: MFI-20 general fatigue subscale (4 to 20) Baseline: 19.0 vs. 17.6 vs. 18.0	Baseline depression: not reported Baseline function: not reported	A: Acetyl-L-carnitine 2g + Propionyl-L-carnitine 2g B: Acetyl-L-carnitine 2g C: Propionyl-L-carnitine 2g Frequency: Daily Duration of treatment: 24 weeks Duration of follow-up: 2 weeks after end of treatment
Weatherley-Jones, 2004[117] United Kingdom Medium	n: 103 Age: 38.9 vs. 38.8 % Female: 57 vs. 62	Criteria: Oxford Duration, years: 4.8 vs. 3.7	Fatigue scale: MFI-20 (20 to 100) Baseline fatigue: MFI-20 general fatigue subscale (4 to 20): 18.4 vs. 18.1	Baseline depression: not reported Baseline function: Functional Limitations Profile physical dimension: 20.4 vs. 22.1 Functional Limitations Profile psychosocial dimension: 35.1 vs. 36.3	A: Homeopathy B: Placebo Frequency: Monthly visits to homeopath, treatments varied Duration of treatment: 6 months Duration of follow-up: 1 month after end of treatment; 7 months after randomization
Williams, 2002[115] United Kingdom Medium	n: 30 Age: 44.5 % Female: 57	Criteria: Oxford Duration, years: 3.6	VAS (0 to 10): 7.1 vs. 6.6 Mental Fatigue Inventory (0 to 36): 25 vs. 24	Baseline depression: HADS depression (0 to 21): 11 vs. 9 Baseline function: SF-36 physical function (0 to100): 25 vs. 42.2	A: Melatonin 5mg B: Phototherapy with 2500 Lux Lightbox for 30 minutes in the morning Frequency: Daily Duration of treatment: 60 weeks: 12 weeks placebo, 12 weeks treatment, 12-week washout or placebo, then 12-week crossover and 12-week washout or placebo Duration of follow-up: End of treatment

Abbreviations: CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; HADS-D = Hospital Anxiety and Depression Scale-depression; ME = myalgic encephalomyelitis; MFI-20 = Multidimensional Fatigue Inventory 20-item; RCT = randomized controlled trial; SF-36 = 36-item Short Form Health Survey; SIP-8 = Sickness Impact Profile; SSRI = selective serotonin reuptake inhibitor; VAS = visual analogue scale

Table 19. RCTs of dietary interventions, herbal supplements, or homeopathy: study results

Author, Year ME/CFS criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Hobday, 2008[118] Fukuda	A: Low sugar/low yeast (19) B: Healthy eating (20) Duration of treatment: 24 weeks Duration of follow-up: End of treatment	Chalder Fatigue Scale 14-item (0 to 56)mean (SD): 16.0 (8.2) vs. 17.7 (10.0); p=0.6	HADS depression (0 to 21), mean (SD): 6.5 (3.6) vs. 5.4 (3.7); p=0.33	SF-36 physical function (0 to 100), mean: 42.3 (29.2) vs. 52.2 (24.1); p=0.25 SF-36 social functioning subscale, mean: 42.0 (29.3) vs. 50.6 (29.4), p=0.35
Li, 2015[120] Fukuda	A: Dengzhanshengmai herbal supplement + SSRI (134) B: SSRI alone (134) Duration of treatment: 12 weeks Duration of follow-up: End of treatment	MFI-20 general fatigue subscale (4 to 20), mean improvement: Improvement from week 2 to end of treatment General Fatigue: 1.3 (0.7) vs. 0.8 (0.6), p<0.01 Physical Fatigue: 1.0 (0.4) vs. 0.6 (0.3), p<0.01 Reduced Activity: 1.3 (0.6) vs. 1.0 (0.5), p<0.01 Improvement from week 8 to end of treatment Reduced Motivation: 2.4 (1.0) vs. 2.1 (0.8), p<0.01 No improvement Mental Fatigue: data not shown, p>0.05	No significant differences, data not shown	Not reported
Malaguarnera, 2008[112] Fukuda or Holmes	A: Acetyl L-carnitine 2 g (48) B: Placebo (48) Duration of treatment: 180 days Duration of follow-up: End of treatment	Fatigue Severity Scale 9-item (1 to 7), mean (SD): 27.9 (9.7) vs. 48.9 (6.9), p=0.000 Physical Fatigue: Wessely and Powell Scales(8-item physical, 5-item mental, maximum score 26): 6.4 (2.2) vs. 12.6 (2.4), p=0.000	Not reported	Physical function: PF-10 (0 to 100), mean (SD): 86.9 (17.40 vs. 70.8 (19.1), p=0.000
Ockerman, 2000[116] Fukuda	A: Pollen extract (Polbax) (21) B: Placebo (22) Duration of treatment: 3 months Duration of follow-up: End of treatment	Fatigue, Mean score (Likert scale 0=no problem to 10=extremely serious symptom) 7.52 vs. 7.14; p=NR	Mean depression score (Likert scale 0=no problem to 10=extremely serious symptom) 5.16 vs. 6.60; p=NR	Not reported

Author, Year ME/CFS criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Ostojic, 2016[113] Fukuda	A: Guanidinoacetic acid supplement (not reported) B: Placebo, cellulose (not reported) Duration of treatment: 3 months, Duration of follow-up: End of first treatment period; 3 months after randomization	Fatigue scale: MFI-20 (20 to 100) General fatigue subscale (4 to 20): 11.6 (1.3) vs. 11.8 (1.5), p=0.44 Physical fatigue subscale (4 to 20): 11.7 (1.2) vs. 11.6 (1.4), p=0.99 Reduced activity subscale (4 to 20): 13.9 (1.2) vs. 11.7 (1.8), p=0.00 Reduced motivation subscale (4 to 20): 13.1 (1.9) vs. 15.0 (1.8), p=0.03 Mental fatigue subscale (4 to 20): 12.2 (1.7) vs. 14.0 (0.9), p=0.01	Not reported	Not reported
The, 2007[114] Fukuda	A: Acclydine + amino acid supplements (30) B: Placebo (27) Duration of treatment: 14 weeks Duration of follow-up: End of treatment	Fatigue scale: Checklist Individual Strength, fatigue severity subscale (8 to 56) 14 weeks: 42.4 (11.6) vs. 43.0 (12.6); p=0.70	Not reported	14 weeks: 1,228.1 (619.7) vs. 1,120.2 (543.0); p=0.65
Vermeulen, 2004[119] Fukuda	A: Acetyl-L-carnitine + Propionyl-L-carnitine (30) B: Acetyl-L-carnitine (29) C: Propionyl-L-carnitine (30) Duration of treatment: 24 weeks Duration of follow-up: 2 weeks after end of treatment	General fatigue at 24 weeks; MFI-20 general fatigue subscale (4 to 20): 17.3 (3.3) vs. 15.9 (4.2) vs. 16.5 (3.1); p=0.004 for propionyl-L-carnitine change from baseline; p=0.000 for combination change from baseline Physical fatigue subscale (4 to 20) at 24 weeks: 16.5 (3.4) vs. 15.7 (4.4) vs. 16.4 (3.2), not significant Mental fatigue subscale (4 to 20) at 24 weeks: 14.6 (4.0) vs. 15.1 (3.6) vs. 13.9 (3.5); p=0.015 for acetyl-L-carnitine change from baseline	Not reported	Not reported

Author, Year ME/CFS criterion	Intervention A: Intervention (n) B: Control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Weatherley-Jones, 2004[117] Oxford	A: Homeopathy (50) B: Placebo (53) Duration of treatment: 6 months Duration of follow-up: 1 month after end of treatment; 7 months after randomization	MFI-20 (20 to 100) : General fatigue subscale (4 to 20), mean change (SD): 2.70 (3.93) vs. 1.35 (2.66), p=0.04 Physical fatigue subscale (4 to 20), mean change (SD): 2.13 (4.00) vs. 1.28 (2.74), p=0.21 Mental fatigue subscale (4 to 20), mean change (SD): 2.70 (4.01) vs. 2.05 (86), p= Reduced activity subscale (4 to 20), mean change (SD): 2.72 (4.47) vs. 1.81 (2.82), p=0.16 Reduced motivation subscale (4 to 20), mean change (SD): 1.35 (4.15) vs. 1.65 (3.02), p=0.82 Fatigue Impact Scale, mean change (SD): Cognitive dimension: 4.88 (9.3) vs. 4.21 (7.18); p=0.61 Physical dimension: 4.98 (8.5) vs. 5.30 (6.69); p=0.98 Social dimension: 7.92 (18.02) vs. 8.20 (14.06); p=0.79	Not reported	Functional Limitations Profile, mean change (SD): Physical dimension: 5.11 (8.82) vs. 2.72 (8.40), p=0.04 Psychosocial dimension: 9.81 (14.19) vs. 6.76 (10.67); p=0.14
Williams, 2002[115] Oxford	A: Melatonin 5mg (42) B: Phototherapy with 2500 Lux Lightbox for 30 minutes in the morning (42) All 30 patients received both treatments, in two possible orders. Duration of treatment: 60 weeks: 12 weeks placebo, 12 weeks treatment, 12-week washout or placebo, then 12-week crossover and 12-week washout or placebo Duration of follow-up: End of treatment	Median (IQR) VAS score for How fatigued are you? (1 to 10 scale, lower score indicates better health) After treatment: 6.1 (4.8 to 8.0) vs. 7.2 (5.5 to 8.3); p=NS Median (IQR) Mental Fatigue Inventory (0 to 36) After treatment: 23 (15.0 to 27.0) vs. 24 (21.0 to 29.0); p=NS	HADS depression (0 to 21): 10 (7.7 to 11.2) vs. 10 (6.0 to 14.0), p= NS	Overall Function: Median (IQR) SF-36 physical function (0 to100) After treatment: 42.5 (16.3 to 53.8) vs. 45 (22.5 to 60.0); p=NS

Abbreviations: CFS = chronic fatigue syndrome; HADS = Hospital Anxiety and Depression Scale; HADS-D = Hospital Anxiety and Depression Scale-depression; IQR = interquartile range; ME = myalgic encephalomyelitis; MFI-20 = Multidimensional Fatigue Inventory 20-item; NR = not reported; NS = not significant; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey; SSRI = selective serotonin reuptake inhibitor

Dietary supplements versus placebo

One medium risk of bias trial[112] (N=96) compared acetyl L-carnitine (2 grams twice daily) versus placebo in older (age >70 years) patients with CFS. Carnitines are amino acid compounds that have a role in metabolism and have been promoted for cognitive benefits. All patients met either the Fukuda or Holmes[10] criteria; in addition, all patients were positive on the Fukuda

minor criterion of prolonged post-exercise fatigue. Patients were evaluated at the end of 6 months of therapy. Acetyl L-carnitine was associated with decreased fatigue severity (mean difference -21.00, 95% CI -24.41 to 17.59 on the 9 to 63 Fatigue Severity Scale), decreased functional limitations (mean difference 16.10, 95% CI 8.70 to 23.50 on the 0 to 100 Physical Function functional limitations scale), and improved cognitive status (mean difference 2.70, 95% CI 1.48 to 3.92 on the 0 to 30 Mini-Mental Status Examination). There was no difference between acetyl L-carnitine versus placebo in severity of disability (mean difference -0.60, 95% CI -8.36 to 7.16 on the 0 to 100 Physical Function disability scale). On individual CFS case definition criteria, acetyl L-carnitine was associated with decreased likelihood of prolonged post-exercise fatigue (48% vs. 96%, RR 0.50, 95% CI 0.37 to 0.68), activity reduction >50% (56% vs. 75%, RR 0.75, 95% CI 0.56 to 1.01), muscle pain (67% vs. 90%, RR 0.74, 95% CI 0.60 to 0.93), and sleep disorder (62% vs. 84%, RR 0.75, 95% CI 0.58 to 0.97), with no statistically significant differences in likelihood of painful throat, painful lymph nodes, neuropsychiatric complaints, spreading arthralgias, or headaches. The trial reported no adverse events or laboratory abnormalities in either group.

One small (N=14)[113] high risk of bias crossover trial evaluated guanidinoacetic acid (GAA, 2.4 g daily) versus placebo in women meeting the Fukuda case definition for CFS. GAA naturally occurs in the body as an immediate precursor of creatine. At the end of 3 months of therapy, GAA was associated with improved scores on the reduced activity (mean difference -2.2, p<0.005), reduced motivation (mean difference -1.9, p=0.03), and mental fatigue (mean difference -1.8, p=0.01) subscales of the MFI-20 (each on a 4 to 20 scale), but no difference on the general fatigue (mean difference -0.2, p=0.44) or physical fatigue (mean difference 0.1, p=0.99) subscales. There were no differences in musculoskeletal soreness at rest (mean difference -0.2 on a 0 to 10 VAS, p=0.31) or during activity (mean difference -0.6 on a 0 to 10 VAS, p=0.18). GAA was also associated with better scores on the SF-36 physical (mean difference 2.4, p=0.04) and mental (mean difference 5.3, p<0.005) component summary scores (both on a 0 to 100 scale, higher scores indicating higher quality of life). The trial reported no harms in either group.

One medium risk of bias trial[114] (N=57) comparing the dietary supplement acclydine plus amino acids versus placebo in patients meeting the Fukuda case definition. Acclydine has been claimed to increase insulin-like growth factor 1 concentrations by stimulating growth hormone releasing hormone. At the end of 14 weeks of treatment, there were no differences between acclydine plus amino acids versus placebo in fatigue severity (mean difference in change from baseline 1.1, 95% CI -4.4 to 6.5 on the 8 to 56 Checklist Individual Strength fatigue subscale, p=0.70), functional impairment (mean difference in change from baseline 59.1, 95% CI -201.7 to 319.8 on the 0 to 5,799 Sickness Impact Profile-8 scale) or activity level (mean difference in change from baseline 4.1, 95% CI -5.9 to 14.0 measured with an actometer). No “important” (not defined) side effects were reported in either group.

Head-to-head comparisons of dietary interventions

One medium risk of bias, open-label trial[119] (N=89) compared dietary supplementation with acetyl L-carnitine (2 grams daily) versus propionyl L-carnitine (2 grams daily) versus both in patients meeting the Fukuda case definition. Acetyl L-carnitine and propionyl L-carnitine are supplements believed to promote mitochondrial energy and decrease oxidative stress. At the end of 24 weeks of treatment, the acetyl L-carnitine and propionyl L-carnitine groups were both associated with higher likelihood of improvement (based on Global Impression of Change score

≥2 on a -3 to +3 scale) than the combination (59% vs. 63% vs. 37%). However, at 2-week post-intervention follow-up, no patient met criteria for improvement in any group. There were no differences between groups in severity general, physical, or mental fatigue assessed with the MFI-20. Harms were not reported.

One small (N=39), high risk of bias, trial[118] randomized patients meeting the Fukuda case definition to either a low sugar low yeast diet or a healthy eating diet for 24 weeks. Full compliance occurred in only 24% of patients assigned to the low sugar low yeast diet and 67% of those assigned to healthy diet. At the end of the intervention, there were no differences between dietary interventions in fatigue severity (mean difference -1.7, 95% CI -7.5 to 4.1 on the 14-item, 0 to 42 Chalder scale), depression severity (mean difference 1.1, 95% CI -1.2 to 3.5 on the 0 to 21 HADS depression scale), or anxiety severity (mean difference 1.2, 95% CI -1.8 to 4.2 on the 0 to 21 HADS anxiety scale). Differences in the 8 SF-36 subscales ranged from -15.1 (body pain) to 2.9 (mental health); however, none of the differences were statistically significant. The trial did not report harms.

Homeopathy or herbal supplements versus placebo

One medium risk of bias trial[117] (n=103) compared homeopathy for 6 months versus placebo in patients meeting the Oxford criteria for CFS. Homeopathic remedies were individualized by the pharmacy for each patient. Fatigue was measured using the MFI-20 (score on each subscale 4 to 20). Homeopathy was associated with greater improvement from baseline versus placebo on the general fatigue subscale (-2.70 vs. -1.35, p=0.04), but differences in the physical fatigue, mental fatigue, activity, and motivation subscales were not statistically significant (differences ranged from -0.91 [reduced activity subscale] to 0.30 [reduced motivation subscale]). There were also no differences between homeopathy versus placebo in the likelihood of ≥3-point improvement in any MFI-20 subscale or the likelihood of ≥3-point improvement in all MFI-20 subscales. The trial did not report harms.

A small (N=43), high risk of bias[116] crossover trial compared of pollen extract (Polbax) versus placebo in adults meeting the Fukuda case definition. The pollen extract is believed to have an antioxidative effect. At the end of 3 months of treatment, there were no differences between pollen extract versus placebo in total well-being, fatigue, fatigability, sleep problems, depression, or intestinal problems. The trial reported no clear side effects except for slight gastrointestinal complaints in 1 or 2 patients.

Other comparisons involving herbal supplements

One high risk of bias, open label trial[120] (N=268) compared the Chinese herbal supplement Dengzhanshengmai (1.08 grams daily) plus an SSRI (paroxetine 10 to 30 mg daily or sertraline 25 to 100 mg daily) versus an SSRI alone for 12 weeks in patients meeting the Fukuda case definition. Results were reported poorly in this trial. At 12 weeks, Dengzhanshengmai plus SSRI was associated with greater change from baseline in the MFI-20 general fatigue, reduced activity, physical fatigue, and reduced motivation subscales versus SSRI alone ($p<0.05$ for general fatigue and $p<0.01$ for the other subscales), but the effects were small (~0.5 on each subscale, each on a 4 to 20 scale). There was no difference in the MRI mental fatigue subscale. There were no differences between Dengzhanshengmai plus SSRI versus an SSRI alone in severity of depression or anxiety (measured using the HADS scales). Rates of any adverse event (41.0% vs. 41.8%) and specific adverse events were similar between groups, with the exception of an increased likelihood of hypertension with the combination (5.8% vs. 1.5%, p=0.05).

A small (N=30),[115] high risk of bias crossover trial evaluated melatonin (5 mg daily) versus phototherapy (2500 Lux lightbox for 1 hour each morning) for 12 weeks in patients who met the Oxford case definition. The purpose of the therapies was to alleviate circadian rhythm disturbances which are often present in CFS. Neither melatonin nor phototherapy were associated with significant improvements (based on assessments using a 0 to 10 VAS) in fatigue, depression, anxiety, waking refreshed, low energy, poor concentration, or muscle pain during treatment with melatonin or phototherapy. Although phototherapy, but not melatonin, was associated with a statistically significant improvement from baseline in severity of sleep disturbance (mean change 1.5 points, p=0.03), the scores at the end of treatment were similar in the melatonin and phototherapy groups (5.5 vs. 5.1). Neither melatonin nor phototherapy were associated with improvements from baseline in any SF-36 subscale or HADS depression or anxiety. Harms were not reported, but the study stated that both treatments were well tolerated and not considered responsible for study withdrawals.

Qigong, Yoga, or Abdominal Tuina

Four trials evaluated qigong (two studies),[121-123] yoga (one study),[124] or abdominal tuina (one study)[125] in adult patients with ME/CFS (**Tables 20 and 21, Evidence Tables Appendix E2**). One trial was included in the prior report.[121,123] Sample sizes ranged from 28 to 137 (total N=272). One trial was conducted in Europe, one trial in Hong Kong, one trial in China, and one trial in Japan. The mean age of participants ranged from 38 to 44 years and the proportion female ranged from 77 percent to 88 percent. The case definition for ME/CFS was the Fukuda criteria in all four trials. The duration of ME/CFS ranged from 10.4 months to 11.9 years in seven trials that reported this information. Baseline fatigue was measured using a variety of scales (**Table 20**). Details regarding the presence of post-exertional fatigue and activity patterns were lacking.

All four trials were rated medium risk of bias (**Risk of Bias Table Appendix F**). Methodological limitations included inadequate description of randomization or allocation concealment methods, and failure to blind or unclear blinding status of outcomes assessors and data analysts.

Qigong versus wait list or no treatment

Two trials (N=137 and 28)[121-123] investigated Qigong versus wait list or no treatment. The larger trial (N=137)[121] compared Qigong versus wait list. The Qigong intervention consisted of training twice weekly in 2-hour sessions for 5 weeks, followed by 12 weeks of home exercise (participants asked to practice at least 30 minutes daily). After 4 months of treatment, Qigong was associated with decreased fatigue severity versus wait list (mean change from baseline -13.1 vs. -6.6 on the 14-item 0 to 42 Chalder scale total fatigue score, p<0.0005). Qigong was also associated with decreased depression severity (mean change from baseline -1.3 vs. 0.4 on the 0 to 21 HADS depression scale, p=0.002), though there was no difference in anxiety severity (mean change from baseline -1.1 vs. -0.6 on the 0 to 21 HADS anxiety scale, p=0.58). The trial did not report harms.

A second (N=28) trial[122] compared Qigong versus no Qigong. Qigong training consisted of 15 weekly, 2-hour sessions of gradually more complex exercises. At completion of therapy, Qigong was associated with greater improvement in fatigue severity (mean difference in change from baseline -0.5, 95% CI -0.9 to -0.02 on the 1 to 7 Fatigue Severity Scale). However, the Fatigue Severity Scale was ≥5 in all patients at the end of the trial, indicating significant fatigue remained present. There were no differences between Qigong versus no Qigong in any SF-36

subscale, including physical function and bodily pain. However, Qigong was associated with greater reduction in pain intensity that was borderline statistical significance (mean change from baseline 1.4 vs. "similar" [data not provided], p=0.05). There were no effects on depression severity (HADS). The trial did not report harms.

Abdominal tuina versus acupuncture

One medium risk of bias[125] trial (N=77) compared abdominal tuina versus acupuncture 5 days per week for 4 weeks in patients who met the Fukuda case definition. Abdominal tuina is a massage technique that uses traditional Chinese medicine principles. At 3 months following the completion of therapy, tuina was associated with decreased fatigue severity versus acupuncture (mean 6.6 vs. 7.6, mean difference 1.0, 95% CI 0.11 to 1.88 on the 0 to 14 Fatigue-Scale 14, p=0.015). Tuina was also associated with decreased depression severity (6.3 vs. 7.0, mean difference 0.70, 95% CI 0.13 to 1.27 on the 0 to 52 Hamilton Depression Rating scale, p=0.044), and borderline associated with anxiety severity (47.0 vs. 49.0, mean difference 2.0, 95% CI -0.05 to 4.05 on the 20 to 80 Zung Self-Rating Anxiety Scale). No serious harms were reported, and there was no significant difference in the likelihood of adverse events.

Yoga versus plus conventional pharmacotherapy versus pharmacotherapy alone

One small (N=30),[124] medium risk of bias compared yoga plus conventional pharmacotherapy (medications not described) versus pharmacotherapy alone in patients who met the Fukuda case definition. Yoga sessions occurred every 2 to 3 weeks for 2 months (mean number of visits: 5.6), with follow-up 2 months after completion of therapy. Yoga was associated with decreased fatigue severity versus pharmacotherapy alone (mean 19.2 vs. 25.8, difference 6.6, 95% CI 1.55 to 11.65 on the 14-item 0 to 42 Chalder fatigue scale total score). Reported harms were minor and primarily involved occasional dizziness when practicing yoga. No patient reported post-exertional malaise after yoga sessions.

Distant Healing Versus no Treatment

A low risk of bias trial[126] (N=409) compared distant healing (prayer, or imagining transmission of healing energy, light or healing power) versus wait list in patients that met the Oxford case definition. This study was included in the prior report. Patients had a mean age of 50 years, were 74% female, and had a mean duration of ME/CFS of 11.9 years. Patients were randomized to blinded or unblinded distant healing and wait list (4 arms total). After 6 months of treatment, there were no differences between distant healing versus wait list in SF-36 mental (p=0.18) or physical (p=0.32) component summary scores. There was no interaction between blinding status and effects of distant healing. The trial did not report effects on fatigue. Harms were not reported.

Table 20. RCTs of Qigong, yoga, abdominal tuina, or distant healing: study characteristics

Author, Year Country Risk of Bias	Study n (analyzed) Age, Mean Years % Female	ME/CFS Criterion ME/CFS Duration	Fatigue Scale Baseline Fatigue	Baseline Depression Baseline Function	Intervention Frequency, Duration, and Intensity Duration of Treatment Duration of Follow-up
Chan, 2013[121] Hong Kong Medium	n: 137 Age: 42.4 % Female: 77	Criteria: Fukuda Duration: not reported	Fatigue Scale: Chalder Baseline (14-item, 0 to 56): 39.8 (SD 6.5) Post-exertional fatigue or malaise: not reported	Major depression: Excluded Baseline depression: HADS depression (0 to 21): 9.2 (SD 2.1) Baseline function: not reported	A: Qigong B: Wait list Frequency: Twice weekly Session length: 2 hours Exercise intensity: Qigong ≥30 minutes daily Duration of treatment: 4 months (5 weeks Qigong training and 12 additional weeks home exercise) Duration of follow-up: 4 months
Dybwad, 2007[122] Norway Medium	n: 28 Age: 36 % Female: 84	Criteria: Fukuda Duration: 8.1 years	Fatigue scale: Fatigue Severity Scale 9-item (1 to 7) Baseline fatigue, mean: 6.5	Baseline depression: HADS depression (0 to 21): 4.9 Baseline function: SF-36 physical function (0-100 scale, lower score indicates better health): 48	A: Qigong B: No Qigong Frequency: Once weekly Session length: 2 hours Exercise intensity: not reported, but gradually progressed in complexity Duration of treatment: 15 weeks Duration of follow-up: end of treatment
Huanan, 2017[125] China Medium	n: 77 Age: 41.8 vs 42.63 % Female: 44 vs. 37	Criteria: Fukuda Duration, months: 10.4 vs. 10.6	Fatigue scale: Fatigue Scale- 14 (0 to 14, higher score indicates greater severity of fatigue) Baseline fatigue, mean: 8.9 vs. 9.3	Baseline depression: Hamilton Rating Scale for Depression (0 to 52): 11.0 vs. 10.9 Baseline function: Not reported	A: Abdominal tuina B: Acupuncture Frequency: 5 days per week with 2 days off between weeks Duration of treatment: 4 weeks Duration of follow-up: 3 months after treatment
Oka, 2014[124] Japan Medium	n: 30 Age: 38 % Female: 80	Criteria: Fukuda Duration: Not reported	Fatigue scale: Chalder Fatigue Scale 14-item (0 to 56) Baseline fatigue, mean: 30.8	Baseline depression: not reported Baseline function: not reported	A: Yoga + pharmacotherapy B: Pharmacotherapy alone Frequency: Every 2-3 weeks Duration of treatment: 2 months, mean 5.6 visits Duration of follow-up: 2 months after yoga ended
Walach, 2008[126] Germany and Austria Low	n: 409 Age: 47.5, 48.1, 46.2, 50.4 % Female: 74.3, 76.5, 76.6, 75	Criteria: Fukuda or Oxford severe idiopathic CFS Duration, years: 11.3, 9.6, 9.6, 11.9	Fatigue scale: Fatigue Severity Scale 9-item (1 to 7) Baseline fatigue: 6.2, 6.1, 6.1, 6.0	Baseline depression: not reported Baseline function: SF-36 mental health (0 to 100): 36.67, 34.88, 37.28, 35.16 SF-36 physical function (0 to100): 31.02, 31.75, 31.78, 32.71	A: Distant Healing (blinded) B: Distant Healing (unblinded) C: Deferred treatment (blinded) D: Deferred treatment (unblinded) Frequency: Duration of treatment: 6 months Duration of follow-up: 6 months

Abbreviations: CFS = chronic fatigue syndrome; HADS = Hospital Anxiety and Depression Scale; HADS-D = Hospital Anxiety and Depression Scale-depression; HAM-D = Hamilton Depression Rating Scale; ME = myalgic encephalomyelitis; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey

Table 21. RCTs of Qigong, yoga, abdominal tuina, or distant healing: study results

Author, Year ME/CFS Criterion	Intervention A: intervention (n) B: control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Chan, 2013[121] Fukuda	A: Qigong (72) B: Wait list (65) Duration of treatment: 4 months Duration of follow-up: 4 months	Fatigue score change at 12 weeks, mean (SD): Chalder Fatigue Scale 14-item (0 to 56) total fatigue scores: -13.1 (11.7) vs. -6.6 (8.3), p=0.000 Chalder Fatigue Scale physical fatigue scores: -8.8 (7.3) vs. -3.8 (5.0), p=000 Chalder Fatigue Scale mental fatigue scores: -4.3 (5.3) vs. -2.7 (3.9), p=0.50	Depression score HADS depression (0 to 21) change, mean (SD): -1.3 (2.7) vs. 0.4 (3.7), p=0.002	Not reported
Dybwad, 2007[122] Fukuda	A: Qigong (14) B: No Qigong (14) Duration of treatment: 15 weeks Duration of follow-up: end of treatment	Fatigue Severity Scale 9-item (1 to 7), mean change (SD): -0.44 (0.60) vs. 0 (0.6), p=0.04, adjusted for baseline values All participants in both groups still clinically fatigued	No significant changes observed after intervention within or between groups, data not shown	SF-36, physical functioning, mean change (SD): 4.7 (13) vs. 1.3 (16), p=0.34 adjusted for baseline value.
Huanan, 2017[125] Fukuda	A: Abdominal tuina (39) B: Acupuncture therapy (38) Duration of treatment: 4 weeks Duration of follow-up: 3 months after treatment	Fatigue Scale-14 (0 to 14, higher score indicates greater severity of fatigue): 7.1 (1.7) vs. 8.2 (2.0), p=0.015	Hamilton Rating Scale for Depression (0 to 52): 6.3 (1.2) vs. 7.0 (1.5), p=0.044	Not reported
Oka, 2014[124] Fukuda	A: Yoga + pharmacotherapy (15) B: Pharmacotherapy alone (15) Duration of treatment: 2 months, mean 5.6 visits Duration of follow-up: 2 months after yoga ended	Fatigue, mean (SD), time x group interaction Chalder Fatigue Scale physical fatigue (14-item, 0 to 56): 12.3 (3.8) vs. 16.1 (3.6), p=0.009 Chalder Fatigue Scale mental fatigue: 6.9 (4.4) vs. 9.7 (3.1), p=0.007 Chalder Fatigue Scale total fatigue: 19.2 (7.5) vs. 25.8 (5.9), p=0.003	Not reported	Not reported

Author, Year ME/CFS Criterion	Intervention A: intervention (n) B: control (n) Duration of Treatment Duration of Follow-up	Fatigue Outcomes (fatigue and post-exertional fatigue)	Depression Outcomes	Function Outcomes
Walach, 2008[126] Fukuda or Oxford	A: Distant Healing, blinded (105) B: Distant Healing, unblinded (102) C: Deferred treatment, blinded (94) D: Deferred treatment, unblinded (108) Duration of treatment: 6 months Duration of follow-up: 6 months	Not reported	Not reported	Change from baseline: SF-36 Physical Function (0-100 scale, lower score indicates better health): 3.66 (6.83) vs. 3.04 (7.38) vs. 3.29 (7.28) vs. 0.75 (7.85); p=not significant, data not shown SF-36 mental health (0 to100): -0.29 (9.54) vs. 1.74 (10.25) vs. 1.16 (11.07) vs. 0.81 (10.45); p=not significant, data not shown

Abbreviations: CFS = chronic fatigue syndrome; HAM-D = Hamilton Depression Rating Scale; ME = myalgic encephalomyelitis; RCT = randomized controlled trial; SD = standard deviation; SF-36 = 36-item Short Form Health Survey

Discussion

This review synthesizes evidence on the evaluation and management of ME/CFS. Specifically, it addresses the prevalence of non-ME/CFS conditions in persons presenting with fatigue and the benefits and harms of treatments for ME/CFS. Although we also sought to synthesize evidence on the benefits and harms of receiving an ME/CFS diagnosis versus no diagnosis, no study met inclusion criteria. The key findings of this review are summarized in the summary of evidence table (**Table 22**).

The bulk of the evidence in this report addressed the effectiveness of treatments for ME/CFS. The prior AHRQ report found GET and CBT were associated with improved fatigue and function versus inactive controls, but the applicability of findings to more disabled populations with ME/CFS and those diagnosed using more current ME/CFS criteria (including criteria that require presence of post-exertional malaise) was uncertain.[1] The prior report also found that rintatolimod was associated with improved exercise performance in some patients, but the strength of evidence was low. There was insufficient evidence to determine effects of other treatments and harms of therapy. The findings of this report regarding treatments for ME/CFS expands upon the prior AHRQ report by including new studies, adding children as a population of interest, and evaluating outcomes in addition to fatigue and function (e.g., depression, anxiety, sleep, and pain). We also sought to determine how effects of treatment differed in subgroups based on ME/CFS disease severity, duration of symptoms, ME/CFS case definition used, demographic factors, comorbidities, and intervention characteristics (e.g., type of exercise, method of delivering CBT).

As in the prior AHRQ report, exercise and CBT remained the most frequently studied interventions in adults with ME/CFS, with 33 new trials of these interventions added for this update. Like the prior report, we found some evidence that graded exercise and CBT are associated with improved fatigue and function versus inactive controls (wait list, usual care, usual specialist care, attention control, or placebo) at the end of the intervention, with more limited data indicating sustained benefits at post-intervention follow-up 2 to 12 months following completion of therapy. However, these findings must be interpreted with caution. In a number of pooled analyses, statistical heterogeneity was large and not fully explained in subgroup analyses based on the type of control, the scale used to measure fatigue or function, or the method for delivering CBT. Although an outlier trial[56] of exercise that reported unusually strong results was identified, it was unclear why results of this trial differed from the others and statistical heterogeneity remained present when it was excluded. In addition, the magnitude of effects was relatively modest. For fatigue and function, pooled SMD's ranged from -0.55 to -0.76, or within the "moderate" range (0.50 to 0.80).[35] However, mean differences based on the original scales were 2 to 3 points in trials that used the 11-item 0 to 33 Chalder fatigue scale, 3 to 4 points in trials that used the 11-item 0 to 11 Chalder fatigue scale, and 5 to 6 points on the 0 to 100 SF-36 physical function subscale. Therefore, observed effects ranged from below to slightly above proposed thresholds for minimum clinically important differences (2.3 points on the 11-item 0 to 11 Chalder scale and 10 points on the SF-36 physical function subscale). In addition, minimum clinically important differences are not well-established in patients with ME/CFS. For exercise, effects on fatigue and function were attenuated and below the "moderate' threshold when the outlier trial described above was excluded (SMD's 0.32 to 0.35). The applicability of results to patients with more severe ME/CFS symptoms also remains unclear. At baseline, mean scores on the SF-36 physical function subscale ranged from ~30 to ~60, with most trials reporting scores in the 40 to 50 range. Whether similar effects would occur in patients with more severe functional

limitations and symptoms due to ME/CFS is unknown. The applicability of findings to patients with severe post-exertional malaise is also uncertain. Only one exercise trial required patients to have post-exertional fatigue or malaise at baseline, but it did not report the proportion of patients with post-exertional malaise;[52] the other trials did not report the baseline prevalence or severity of post-exertional fatigue or malaise. Similarly, the applicability of findings to patients diagnosed using more recent, specific ME/CFS case definitions (including those requiring presence of post-exertional malaise) is uncertain. Although we found similar results for exercise and CBT regardless of the ME/CFS case definition used, subgroup analyses were limited by small numbers of trials. In addition, the most commonly used case definition in the exercise trials was the Oxford case definition, in which has been shown to result in a substantially higher prevalence of ME/CFS than case definitions that require presence of post-exertional malaise. This could have resulted in potential misclassification (overdiagnosis) of patients with non-ME/CFS fatiguing conditions. There were also methodological limitations in the trials. It was not possible to blind patients and care providers to the exercise and CBT interventions, potentially resulting in bias related to delivery of interventions, expectations of treatment, provider enthusiasm for different treatments, and assessment of outcomes.[127] The inability to blind is of particular concern for subjective outcomes such as fatigue and function. Most trials also had other methodological limitations, including unclear randomization or allocation concealment methods and high attrition. There were also challenges related to interpretation of outcomes, including lack of validation of outcome measures in patients with ME/CFS (including ceiling and floor effects) and poor standardization in selection and reporting of outcomes. For example, post-exertional malaise was not reported in most trials; when reported it was not defined well and may not have adequately distinguished post-exertional malaise from post-exertional fatigue.[128] Because of these issues, the strength of evidence for exercise therapy and CBT versus inactive therapies was rated low, even though the literature for these treatments represented the most robust bodies of evidence on treatments for ME/CFS.

There were other challenges in interpreting our findings. Exercise and CBT were also associated with increased likelihood of improvement in fatigue, function, and recovery, based on differences in the proportion of patients meeting a defined threshold. Such dichotomous outcomes can be more informative than continuous outcomes based on average improvements, because they indicate the proportion of patients who experience clinically meaningful benefits. However, few trials reported dichotomous outcomes, the thresholds used to define dichotomous outcomes varied, and findings were largely driven by the largest trial, PACE. The PACE trial has been criticized because of protocol modifications involving the definitions used for dichotomous outcomes, including the primary outcome of overall efficacy (a composite of fatigue and function), recovery, and improvement in fatigue and function.[66] These changes resulted in less stringent thresholds that were met by higher proportions of patients. The modified definition for recovery was of particular concern because it encompassed patients with significant symptoms; furthermore, some patients met the criteria for recovery at the time of study entry, contradicting the concept of recovery. However, findings were similar in sensitivity analyses that utilized data based on the original PACE protocol definitions for these outcomes.[40] The PACE trial has also been critiqued for other issues, including perceived conflicts of interest, failure to comply with requests with data sharing requests, and methodological limitations. However, a review of PACE by the UK Health Research Authority found that protocol modifications were reported appropriately, that there was no evidence that outcomes were changed to favor certain outcomes, that conflicts of interests were reported and handled appropriately, that requests for data were

handled responsibly, and that PACE had appropriate oversight without indication that it was a poorly conducted trial.[129]

Exercise and CBT were also associated with small beneficial effects on depression and anxiety versus inactive therapies (differences 1.2 to 2.4 points on the 0 to 21 HADS depression or anxiety scales; proposed minimum clinically important difference 1.7) and sleep. Data on effects on the 6-minute walk test were limited but indicated small but statistically significant for exercise versus inactive controls (mean difference 31 minutes), but not CBT. Data on harms were very limited, mostly reported from the PACE and GETSET trials. There was no increase in risk of serious adverse events, withdrawal due to worsening, or physical function worsening, but some estimates were imprecise and harms were not well-reported. In the PACE trial, exercise and CBT were associated with decreased likelihood of post-exertional malaise versus usual specialist care.[38]

Evidence on the comparative effectiveness of exercise or CBT versus other active therapies for ME/CFS was limited. There were no differences between exercise versus CBT in fatigue, function, or other outcomes, though findings were based on one or two trials (including PACE), with imprecise estimates. Data were also too limited for reliable comparisons of exercise or CBT versus other active therapies (relaxation, cognitive therapy, adaptive therapy, medications), though the PACE trial[38] generally found that adaptive pacing generally performed worse than either exercise or CBT.

For adolescents with ME/CFS, evidence on the effectiveness of treatments mostly focused on CBT. Evidence found CBT with a family focus or parental involvements associated with decreased fatigue severity versus inactive controls, though there was heterogeneity in the magnitude of effects, which ranged from small to large. Although estimates also favored CBT for severity of functional impairment and school attendance, the differences were not statistically significant. All estimates were based on a small number of trials (three) and statistical heterogeneity was very high. No trial evaluated the effectiveness of exercise in adolescents. One new trial found an osteopathy, life coaching, and neurolinguistic programming intervention ("Lightning Process") associated with improved function versus usual specialist care, with no statistically significant effects on fatigue, anxiety, depression, or quality of life.[89] This trial involves an intensive but brief intervention with limited follow-up sessions; additional research is needed to reproduce and verify these findings.

Regarding medications for ME/CFS, a new placebo-controlled trial of rituximab, a monoclonal antibody against the CD20 protein with immunomodulatory effects, failed to confirm positive results of an earlier pilot trial, showing no statistically significant effects on fatigue, function, or other clinical outcomes.[94,95] Small, more recent trials of anakinra,[106] duloxetine,[110] mirtazapine,[75] clonidine, and methylphenidate[109] in patients with ME/CFS also found no beneficial effects on fatigue, function or other clinical outcomes. The only trial of a medication in adolescents with ME/CFS found IgG associated with improvement in function after 3 months versus placebo, but not at other measurement time points.[107] Two trials of adults found no differences between IgG versus placebo in fatigue, function, or other clinical outcomes.[104,105] There was no new evidence on rintatolimod, which is not approved by the FDA for any indication. Rintatolimod was reviewed by the FDA in 2012 and failed to receive approval. Two trials included in the prior report found some evidence of improved exercise tolerance with rintatolimod versus placebo, but improvements in overall function seen in an initial, smaller trial were not replicated in a subsequent, larger trial.[100,101] Rintatolimod was associated with infusion-related headaches and flu-like symptoms in one trial. No new trial of

rintatolimod has been published since 2012. A small trial of patients with suspected viral onset of ME/CFS found no statistically significant differences between valganciclovir versus placebo in the MFI-20 total fatigue, CDC symptom inventory scores, sleep quality, or depression, though valganciclovir was associated with improved self-reported cognitive functioning and severity of mental fatigue.[98] Antidepressants did not significantly improve depressive symptoms in those without MDD at baseline, and improvements did not meet clinically important differences in one study that enrolled patients with MDD. The evidence on other medications for ME/CFS was sparse and did not indicate benefits.

Evidence on the effectiveness of other therapies was limited. Limited evidence indicated that Qigong was associated with beneficial effects on fatigue and function; evidence on other complementary and alternative approaches, dietary supplements, and dietary interventions was limited but did not indicate beneficial effects. One trial of home orthostatic training found orthostatic training associated with a decrease in blood pressure drop with standing versus sham training, but did not measure effects on orthostatic symptoms.[130]

Our findings are generally consistent with a recent systematic review on exercise therapy for ME/CFS that concluded that exercise probably has a positive effect on fatigue in adults compared to usual care or passive therapies, but noted uncertain applicability to patients diagnosed with case definitions other than the Oxford and Fukuda criteria.[131] This review also found that there was limited evidence on exercise versus other active therapies, limiting its ability to determine comparative effectiveness. With regard to the use of medications for ME/CFS, our findings are consistent with a recent systematic review that found limited evidence on medications for ME/CFS, which was insufficient to draw conclusions.[132] An older systematic review found CBT associated with small improvement versus usual care, but was conducted in 2008, prior to the publication of PACE and other trials of CBT.[133]

Evidence to determine the degree to which patient characteristics predict treatment effects was limited. The PACE trial, which enrolled patients using the Oxford case definition, found no interaction between whether patients met alternative case definitions (2003 Reeves, 1994 London)[92] or presented with a primary depressive or anxiety disorder and estimates of effectiveness. Evidence on the interaction between severity of baseline functional impairment and effects of exercise or CBT was limited and inconsistent, with some trials finding worse baseline function associated with greater response to therapy and others finding worse baseline function associated with worse response. Three trials found no interaction between baseline depression and effects of exercise or CBT.

The prevalence of non-ME/CFS conditions in patients presenting with fatigue is high, though estimates appeared to vary depending on the setting. In addition, the reliability of methods used to diagnose or exclude ME/CFS was uncertain. In a systematic review of patients who sought care for fatigue or tiredness in primary care settings, the most common non-ME/CFS conditions were depression (18.5%), anemia (2.8%), malignancy (0.6%), and serious somatic diseases (including diabetes, anemia, hypothyroidism, and malignancy, 4.3%).[45] In specialty settings of patients referred for evaluation of possible ME/CFS, the most common non-ME/CFS conditions were psychological (15% to 51%) and sleep disorders (6% to 30%). A variety of other non-ME/CFS conditions were reported in both primary care and specialty settings, highlighting the importance of the clinical and diagnostic evaluation of patients with fatigue.

Table 22. Summary of evidence

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
Exercise vs. inactive therapies (adults)	Fatigue, end of intervention	6 (1034)	Direct	Precise	Moderate	Inconsistent	SMD –0.55 (-0.91 to -0.19); SMD -0.32 (-0.52 to -0.12) without outlier trial	Low
	Fatigue, post-intervention	3 (625)	Direct	Precise	Moderate	Inconsistent	SMD -0.76 (-1.48 to -0.05); SMD -0.35 (-0.58 to -0.12) without outlier trial	Low
	Fatigue improvement	1 (305)	Direct	Precise	Moderate	Unable to assess	RR 1.23 (1.07 to 1.42)	Low
	SF-36 physical function (0 to 100), end of intervention	5 (965)	Direct	Precise	Moderate	Inconsistent	MD 11.73 (2.33 to 21.14); MD 5.89 (2.52 to 9.25) without outlier trial	Low
	SF-36 physical function (0 to 100), post-intervention	3 (711)	Direct	Precise	Moderate	Inconsistent	MD 17.07 (-2.02 to 36.16); MD 6.37 (1.89 to 10.85) without outlier trial	Low
	Functional improvement	3 (618)	Direct	Imprecise	Moderate	Inconsistent	RR 2.48 (0.77 to 7.97); RR 1.41 (1.15 to 1.74) without outlier trial	Low
	HADS depression (0 to 21), end of intervention	4 (688)	Direct	Precise	Moderate	Inconsistent	MD -1.83 (-3.65 to -0.01); MD -0.97 (-1.71 to -0.23) without outlier trial	Low
	HADS depression (0 to 21), post-intervention	3 (699)	Direct	Imprecise	Moderate	Inconsistent	MD -2.36 (-4.98 to 0.27)	Low
	HADS anxiety (0 to 21), end of intervention	3 (620)	Direct	Precise	Moderate	Consistent	MD -1.59 (-2.41 to -0.77); MD -1.31 (-2.12 to -0.51) without outlier trial	Low
	HADS anxiety (0 to 21), post-intervention	3 (697)	Direct	Precise	Moderate	Inconsistent	MD -1.07 (-2.64 to 0.49); MD -0.38 (-1.52 to 0.76) without outlier trial	Low
	Sleep, end of intervention	2 (420)	Direct	Precise	Moderate	Consistent	SMD -0.35 (-0.56 to -0.13); SMD -0.31 (-0.57 to -0.05) without outlier trial	Low
	Sleep, post-intervention	3 (700)	Direct	Precise	Moderate	Inconsistent	SMD -0.39 (-0.71 to -0.07); SMD -0.26 (-0.56 to 0.03) without outlier trial	Low
	Recovery	3 (536)	Direct	Precise	Moderate	Consistent	RR 2.73 (1.65 to 4.52)	Low
	Serious adverse events	2 (518)	Direct	Imprecise	Moderate	Consistent	RR 1.59 (0.69 to 3.66)	Low

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
	Withdrawal due to worsening	1 (320)	Direct	Imprecise	Moderate	Unable to assess	RR 2.00 (0.18 to 21.84)	Low
	Physical function worsening	2 (518)	Direct	Imprecise	Moderate	Consistent	RR 0.83 (0.52 to 1.34)	Low
	Post-exertional malaise	1 (320)	Direct	Precise	Moderate	Unable to assess	RR 0.70 (0.57 to 0.87)	Low
CBT vs. inactive therapies (adults)	Fatigue, end of intervention	7 (1129)	Direct	Precise	Moderate	Inconsistent	SMD -0.61 (-0.83 to -0.40)	Low
	Fatigue, post-intervention	3 (489)	Direct	Precise	Moderate	Inconsistent	SMD -0.57 (-0.89 to -0.25)	Low
	Fatigue improvement	4 (784)	Direct	Precise	Moderate	Inconsistent	RR 3.00 (0.95 to 9.49)	Low
	SF-36 physical function (0 to 100), end of intervention	5 (1024)	Direct	Precise	Moderate	Consistent	MD 6.58 (3.76 to 9.39)	Low
	Function, post-intervention	3 (489)	Direct	Precise	Moderate	Inconsistent	SMD 0.37 (0.08 to 0.66)	Low
	Functional improvement	3 (488)	Direct	Imprecise	Moderate	Inconsistent	RR 1.06 (0.83 to 1.35)	Low
	Depression, end of intervention	5 (660)	Direct	Precise	Moderate	Inconsistent	SMD -0.26 (-0.49 to -0.03)	Low
	HADS depression (0 to 21), post-intervention	3 (483)	Direct	Precise	Moderate	Consistent	MD -1.24 (-2.01 to -0.47)	Low
	HADS anxiety (0 to 21), post-intervention	3 (481)	Direct	Precise	Moderate	Consistent	MD -1.22 (-1.94 to -0.49)	Low
	Jenkins Sleep Questionnaire (0 to 20), post-intervention	1 (292)	Direct	Precise	Moderate	Unable to assess	MD -1.20 (-2.19 to -0.21)	Low
	Recovery	3 (564)	Direct	Precise	Moderate	Consistent	RR 2.54 (1.53 to 4.22)	Low
	Serious adverse events	1 (321)	Direct	Imprecise	Moderate	Unable to assess	RR 0.99 (0.36 to 2.77)	Low
	Withdrawal due to worsening	1 (321)	Direct	Imprecise	Moderate	Unable to assess	RR 0.33 (0.01 to 8.07)	Low
	Physical function worsening	1 (321)	Direct	Imprecise	Moderate	Unable to assess	RR 0.83 (0.26 to 2.66)	Low
	Post-exertional malaise	1 (321)	Direct	Precise	Moderate	Unable to assess	RR 0.78 (0.64 to 0.95)	Low

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
CBT vs. inactive therapies (adolescents)	Fatigue, end of intervention	3 (263)	Direct	Precise	Moderate	Inconsistent	SMD -0.84 (-1.52 to -0.15)	Low
	11-item Chalder fatigue scale (0 to 33), post-intervention	1 (63)	Direct	Imprecise	Moderate	Unable to assess	MD -1.9 (-5.3 to 1.5)	Low
	Fatigue improvement (dichotomous)	2 (200)	Direct	Imprecise	Moderate	Consistent	RR 3.13 (2.18 to 4.49)	Low
	Function, end of intervention	3 (263)	Direct	Imprecise	Moderate	Inconsistent	SMD 0.49 (-0.34 to 1.32)	Low
	SF-36 physical function (0 to 100), post-intervention	1 (63)	Direct	Imprecise	Moderate	Unable to assess	MD 6.1 (-9.2 to 21.4)	Low
	Functional improvement	2 (200)	Direct	Imprecise	Moderate	Consistent	RR 3.35 (2.25 to 4.99)	Low
	School attendance	3 (251)	Direct	Imprecise	Moderate	Inconsistent	RR 1.96 (0.57 to 6.79)	Low
	Overall improvement	3 (256)	Direct	Imprecise	Moderate	Inconsistent	RR 1.66 (0.67 to 4.10)	Low
	Recovery	1 (131)	Direct	Precise	Moderate	Unable to assess	RR 3.82 (2.31 to 6.31)	Low
Exercise vs. CBT (adult)	Fatigue, function, depression, anxiety, sleep, pain, recovery, fatigue improvement, functional improvement, serious adverse events, withdrawal due to worsening, physical function worsening, post-exertional malaise	1 to 2 (58 to 360)	Direct	Imprecise for most outcomes	Moderate	Varies	Overall no differences between exercise vs. CBT in various outcomes	Low

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
Exercise vs. other active therapies (relaxation, adaptive pacing, biofeedback, fluoxetine) (adults)	Fatigue, function, depression, anxiety, sleep, pain, recovery, fatigue improvement, functional improvement, serious adverse events, withdrawal due to worsening, physical function worsening, post-exertional malaise	1 to 2 (24 to 305)	Direct	Imprecise for most outcomes	Moderate	Varies	Overall no differences between exercise vs. active therapies; 1 trial found exercise associated with better outcomes versus adaptive pacing; 1 trial found exercise associated with lower likelihood of function worsening versus adaptive pacing	Low (adaptive pacing, relaxation) to insufficient (biofeedback, fluoxetine)
CBT vs. other active therapies (relaxation, cognitive therapy, adaptive pacing, mirtazapine)	Fatigue, function, depression, anxiety, sleep, pain, recovery, fatigue improvement, functional improvement, serious adverse events, withdrawal due to worsening, physical function worsening, post-exertional malaise	1 to 2 (57 to 320)	Direct	Imprecise for most outcomes	Moderate	Varies	Overall no differences between CBT vs. other active therapies; 1 trial found CBT associated with better outcomes versus adaptive pacing	Low (adaptive pacing, relaxation) to insufficient (cognitive therapy, mirtazapine)
Illness management and peer counseling vs. wait list (adults)	Fatigue, quality of life, function	1 (47)	Direct	Imprecise	Moderate	Unable to assess	Overall no differences in outcomes	Low
Mindfulness-based cognitive therapy vs. usual care or wait list (adults)	Fatigue, function, depression, anxiety	2 (53)	Direct	Imprecise	High	Consistent	Overall no statistically significant differences	Insufficient
Self-management versus usual care (adults)	Fatigue, function, depression, anxiety	2 (249)	Direct	Imprecise for most outcomes	Moderate	Inconsistent	Inconsistent effects in two trials	Insufficient

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
Osteopathy, life coaching, and neurolinguistic programming versus usual specialist care	Fatigue, function, pain, anxiety, depression, quality of life, school attendance	1 (81)	Direct	Imprecise for most outcomes	Moderate	Unable to assess	Intervention associated with improved function, but no differences in fatigue, pain, anxiety, depression, or quality of life. School attendance improved at 12 months but not at 6 months.	Low
Home orthostatic training vs. sham training	Fatigue, blood pressure drop on standing	1 (27)	Direct	Imprecise for fatigue	Moderate	Unable to assess	No difference in fatigue; orthostatic training associated with reduction in blood pressure drop with standing	Low
Acclydine vs. placebo	Function, fatigue, physical activity (actometer)	1 (57)	Direct	Imprecise	High	Unable to assess	No effect	Insufficient
Anakinra vs. placebo	Fatigue, function	1 (50)	Direct	Imprecise	Low	Unable to assess	No difference in fatigue or function	Insufficient
Alfa-2a interferon vs. placebo	Quality of life	1 (26)	Direct	Imprecise	High	Unable to assess	No difference in quality of life	Insufficient
Clonidine vs. placebo	Fatigue, function	1 (120)	Direct	Imprecise	Moderate	Unable to assess	No difference in fatigue or function	Low
Duloxetine vs. placebo	Fatigue, function	1 (57)	Direct	Imprecise	Moderate	Unable to assess	No difference in fatigue or function	Insufficient
Fluoxetine vs. placebo	Fatigue, function	2 (166)	Direct	Imprecise	Medium	Consistent	No difference in fatigue or function	Low
Galantamine vs. placebo	Fatigue	1 (423)	Direct	Imprecise	Medium	Unable to assess	No difference in fatigue or quality of life	Insufficient
Hydrocortisone vs. placebo	Fatigue, function	1 (80)	Direct	Imprecise	Medium	Unable to assess	No difference in fatigue or function	Insufficient
Immunoglobulin G vs. placebo in Adults	Function, quality of life	2 (127)	Direct	Imprecise	Medium	Consistent	No difference in function or quality of life	Low
Immunoglobulin G vs. placebo in adolescents	Function	1 (70)	Direct	Imprecise	Moderate	Unable to assess	No difference in function at end of treatment, but IgG associated with improved function at 3-month post-intervention follow-up	Insufficient
Methylphenidate vs. placebo	Function	1 (128)	Direct	Imprecise	Moderate	Unable to assess	No difference in function	Low

Intervention	Outcomes	Number of RCTs (number of subjects)	Directness	Precision	Study limitations	Consistency	Findings (95% CI)	Strength of Evidence
Mirtazapine vs. placebo	Fatigue	1 (49)	Direct	Imprecise	Medium	Unable to assess	No difference in fatigue	Insufficient
Rintatolimod vs. placebo	Function, exercise work capacity	2 (316)	Direct	Imprecise	Medium	Consistent	Rintatolimod associated with improved function and measures of exercise capacity	Low
Rituximab vs. placebo	Fatigue	2 (181)	Direct	Imprecise	Medium	Consistent (fatigue); inconsistent (function)	No difference in fatigue; inconsistent effects on function, with no effect in larger trial	Low
Valganciclovir vs. placebo	Fatigue, function	1 (30)	Direct	Imprecise	Medium	Unable to assess	No difference in function, fatigue improved	Insufficient
Dietary interventions, herbal supplements, homeopathy vs. placebo, usual diet, or another dietary/herbal intervention	Fatigue, function, depression, anxiety, sleep, pain, recovery, fatigue improvement, functional improvement	9 (each evaluated a different intervention, n ranged from 14 to 268)	Direct	Imprecise	High	Inconsistent	Evidence insufficient due to imprecision and study limitations, with results based on a single study for specific interventions	Insufficient
Qigong vs. wait list or usual care	Fatigue, depression, function	2 (165)	Direct	Precise	Medium	Consistent (fatigue); inconsistent (depression); unable to assess (function)	Qigong associated with decreased fatigue severity (2 studies); inconsistent effects on depression (1 study); no difference in function (1 study)	Low for fatigue; insufficient for depression and function
Abdominal tuina vs. acupuncture, yoga + pharmacotherapy vs. pharmacotherapy, distant healing vs. no treatment	Fatigue, function, depression, anxiety, sleep, pain, recovery, fatigue improvement, functional improvement	3 RCTs (each evaluated a different intervention, n ranged from 28 to 409)	Direct	Imprecise	High	Inconsistent	Evidence insufficient due to imprecision and study limitations, with results based on a single study for each intervention	Insufficient

Abbreviations: CBT = cognitive behavioral therapy; CI = confidence interval; HADS = Hospital Anxiety and Depression Scale; MD = mean difference; RCT = randomized controlled trial; RR = relative risk; SF-36 = 36-item Short Form Health Survey; SMD = standardized mean difference; SOE = strength of evidence

Limitations

We note that this systematic review has important limitations. These are related to the quality of the clinical trials as well as in methods used to conduct the review.

Study and Clinical Trial Limitations

Lack of blinded outcome assessment. It is not possible to blind patients or healthcare providers to nonpharmacological interventions such as exercise or CBT, potentially resulting in performance bias or differences in effects based on patient expectations of benefits. Although outcome assessors can generally be blinded even when patients and care providers cannot, most trials did not report blinded outcomes assessment, which could have resulted in bias in measurement or analysis of outcomes.

Failure to describe randomization, attrition, and lack of power. Trials often did not describe randomization or allocation concealment methods and experienced high attrition. Most studies were small and many were underpowered to detect significant differences. The largest trial, PACE, incorporated a number of protocol modifications in measurement and definitions of outcomes.

Lack of standardized outcome measures. The variety of outcomes and methods used to measure them limited our ability to compare results across studies except in meta-analyses, where SMD's (a unitless measures) were calculated. However, interpretation of SMD's can be a challenge because the numerator is based on the difference between groups on an outcome measure and the denominator is based on the precision of the estimates, using the pooled standard deviation. This means that trials that report same average difference in fatigue or function but are more precise (e.g., due to larger sample sizes) will have larger SMD's. Therefore, we also calculated pooled estimates based on the original scales used to calculate pooled SMD's. In some cases, the magnitude of the differences on the original scales were below proposed minimum clinically important difference thresholds when the SMD was within the "moderate" range.

Lack of uniform ME/CFS case definition. Many trials used the earlier Oxford case definition, which includes patients with 6 months of unexplained fatigue with physical and mental impairment, but does not require other specific features commonly present in ME/CFS, such as post-exertional malaise. Using this case definition may classify more patients with ME/CFS compared with more current case definitions, potentially resulting in misclassification and misleading results.

Inadequate number of trials. Most interventions and comparisons were evaluated in small numbers of trials and estimates were frequently imprecise. As a result, meta-analysis was restricted to the most frequently evaluated interventions, exercise and CBT.,.

Applicability limitations. There were few trials of ME/CFS treatment in adolescents; most of the available trials in this age group addressed CBT. Severely affected ME/CFS patients were under-represented. The trials often failed to report important patient characteristics (such as prevalence of post-exertional malaise) and few trials evaluated the effects of important patient characteristics such as age, duration of symptoms, or severity of symptoms on outcomes.

Inadequate data on harms. Harms of therapy were poorly reported in most trials.

Limitations in Methods Used to Conduct review

Challenges with high statistical heterogeneity. A number of analyses were characterized by high statistical heterogeneity. To address this, we utilized random effects models and conducted sensitivity and stratified analyses. However, sensitivity and stratified analyses were limited by small numbers of trials and imprecision, and statistical heterogeneity remained present despite stratification.

Inability to pool results across active comparators. For meta-analyses of exercise or CBT versus other active therapies, we did not pool results across active comparators, given differences in mechanisms of action and potential therapeutic effects. Therefore, findings for each of these comparisons were based on small numbers (one or two) trials.

Pooled analyses across inactive therapies. We pooled analyses across different "inactive" therapies (placebo, wait list, usual care, attention control) and then stratified analyses by the type of inactive therapy because each could potentially impact treatment estimates. , The findings in the stratified analyses were generally consistent and similar across the variety of inactive therapy controls but the small number of trials precludes strong conclusions regarding the impact of type of inactive therapy on findings.

Restricted inclusion to English language articles.

Potential under-analysis of small sample effects. We did not perform graphical or statistical tests for small sample effects, a potential marker for publication bias, due to the small numbers of trials.

Future Research

Research is needed to clarify, further quantify, and understand the effectiveness of exercise, CBT, and other treatments for ME/CFS in patients diagnosed using more specific ME/CFS case definitions that are used in clinical settings and include standardized methods of application. Future research should address the limitations identified above, with improvements in methodological design and conduct that will reduce the risk of bias and improve the strength of the evidence.

Study populations should be more well-defined and reported, specifically characteristics that are important in understanding the effects of treatments for ME/CFS. Use of common data elements (CDE) to characterize important illness characteristics in standardized ways will help study interpretation.[134] Trials should be designed and adequately powered to evaluate subgroup effects based on the severity of symptoms, duration of symptoms, type of onset, demographic factors, biomarkers, and presence of post-exertional malaise and other key symptoms. Studies of adolescents are needed, and additional studies are needed to corroborate the findings for interventions with potential for benefit.

The development of standardized, clinically relevant criteria to define recovery, improvement in fatigue, improvement in function, and other dichotomous outcomes is needed, and such outcomes should be measured in future trials. Measurement of more objective measures of function (e.g., activity trackers or the 6-minute walk test) could help interpret effects based on scales of fatigue or function. Outcomes should be both measured and reported more consistently across studies, with standardized definitions for outcomes such as "improvement," "worsening," and "recovery." Trials should be designed to rigorously assess harms including worsening in function and post-exertional malaise. Research is needed to determine effects of treatments on specific symptoms and conditions associated with ME/CFS, such as cognitive difficulties,

autonomic dysfunction, gastrointestinal disturbance, pain, orthostatic intolerance, and multiple chemical sensitivity.

For interventions which cannot be blinded, expertise-based trial designs may help reduce bias related to provider preferences and enthusiasm regarding the treatment evaluated.[135] In this design, instead of having the same providers deliver all of the interventions in a trial, patients are randomized to clinicians with expertise in intervention A or clinicians with expertise in intervention B. However, the expertise-based design does not address potential biases related to patient expectations and preferences regarding treatment.

Future trials should assess patient expectations and preferences regarding treatments and determine effects on treatment outcomes. Research is also needed to understand benefits and harms of ME/CFS diagnosis versus non-diagnosis, optimal sequencing and combinations of treatments.

Conclusions

Evidence on effective treatments for ME/CFS remains limited. The strength of evidence supporting the use of graded exercise and CBT was small to moderate, with inadequate evaluation in patients diagnosed with more current case definitions, limited reporting of harms, and inadequate evaluation in severely affected patients. Methodological and other limitations (imprecision, inconsistency, uncertain generalizability) precluded strong conclusions. Other therapies were not shown to be effective or require additional evidence to determine effectiveness. Guidance for design and conduct of much needed larger clinical trials is provided.

References

1. Smith MEB, Nelson HD, Haney E, et al. Diagnosis and treatment of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. Evid Rep Technol Assess. 2014 (219):1-433. doi: 10.23970/ahrqepcerta219. PMID: 30313001.

2. Carruthers BM, Jain AK, de Meirleir KL, et al. Myalgic encephalomyelitis/chronic fatigue syndrome: clinical working case definition, diagnostic and treatment protocols. J Chronic Fatigue Syndr. 2003;11(1):7-115. doi: 10.1300/J092v11n01_02.

3. Carruthers BM, van de Sande MI, De Meirleir KL, et al. Myalgic encephalomyelitis: International Consensus Criteria. J Intern Med. 2011;270(4):327-38. doi: 10.1111/j.1365-2796.2011.02428.x. PMID: 21777306.

4. Fukuda K, Straus SE, Hickie I, et al. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med. 1994;121(12):953-9. PMID: 7978722.

5. Institute of Medicine. Beyond myalgic encephalomyelitis/chronic fatigue syndrome: Redefining an illness. Washington, DC: National Academies Press; US; 2015.

6. Chu L, Valencia IJ, Garvert DW, et al. Onset patterns and course of myalgic encephalomyelitis/chronic fatigue syndrome. Front Pediatr. 2019;7(12)doi: 10.3389/fped.2019.00012.

7. Levine PH. Epidemic neuromyasthenia and chronic fatigue syndrome: epidemiological importance of a cluster definition. Clin Infect Dis. 1994;18 Suppl 1:S16-20. PMID: 8148446.

8. Hooper M. Myalgic encephalomyelitis: a review with emphasis on key findings in biomedical research. J Clin Pathol. 2007;60(5):466-71. PMID: 16935967.

9. Jason LA, Brown A, Clyne E, et al. Contrasting case definitions for chronic fatigue syndrome, myalgic encephalomyelitis/chronic fatigue syndrome and myalgic encephalomyelitis. Eval Health Prof. 2012;35(3):280-304. doi: 10.1177/0163278711424281. PMID: 22158691.

10. Holmes GP, Kaplan JE, Gantz NM, et al. Chronic fatigue syndrome: a working case definition. Ann Intern Med. 1988;108(3):387-9. PMID: 2829679.

11. Ramsay AM. Postviral fatigue syndrome: the saga of Royal Free disease: Gower Medical for the Myalgic Encephalomyelitis Association; 1986.

12. A new clinical entity? Lancet. 1956;270(6926):789-90. PMID: 13320887.

13. Lim EJ, Son CG. Review of case definitions for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). J Transl Med. 2020;18(1):289. doi: 10.1186/s12967-020-02455-0. PMID: 32727489.

14. Sharpe MC, Archard LC, Banatvala JE, et al. A report-chronic fatigue syndrome: guidelines for research. J R Soc Med. 1991;84(2):118-21. PMID: 1999813.

15. National Institute for Health Clinical Excellence. Chronic Fatigue Syndrome/myalgic Encephalomyelitis (or Encephalopathy). Diagnosis and Management of CFS/ME in Adults and Children. Quick Reference Guide. NICE Clinical Guideline 53. NICE; 2007.

16. Brurberg KG, Fonhus MS, Larun L, et al. Case definitions for chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review. BMJ Open. 2014;4(2):e003973. doi: 10.1136/bmjopen-2013-003973. PMID: 24508851.

17. Wessely S, Sharpe M. Letter to the Editor. Fatigue: Biomedicine, Health & Behavior. 2018;6(1):61-2. doi: 10.1080/21641846.2017.1364148.

18. Afari N, Buchwald D. Chronic fatigue syndrome: a review. Am J Psychiatry. 2003;160(2):221-36. PMID: 12562565.

19. Rangel L, Garralda ME, Levin M, et al. The course of severe chronic fatigue syndrome in childhood. J R Soc Med. 2000;93(3):129-34. PMID: 10741312.

20. Crawley EM, Emond AM, Sterne JA. Unidentified Chronic Fatigue Syndrome/myalgic encephalomyelitis (CFS/ME) is a major cause of school absence: surveillance outcomes from school-based clinics. BMJ Open. 2011;1(2):e000252. doi: 10.1136/bmjopen-2011-000252. PMID: 22155938.

21. Centers for Disease Control and Prevention. Symptoms and Diagnosis of ME/CFS in Children. 2018. https://www.cdc.gov/me-cfs/me-cfs-children/children-symptoms-diagnosis.html. Accessed January 1, 2020.

22. Reyes M, Nisenbaum R, Hoaglin DC, et al. Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. Arch Intern Med. 2003;163(13):1530-6. PMID: 12860574.

23. Center for Drug Evaluation and Research. The Voice of the Patient: A series of reports from the U.S. Food and Drug Administration's (FDA's) Patient-Focused Drug Development Initiative: Chronic Fatigue Syndrome and Myalgic Encephalomyelitis. U.S. Food and Drug Administration; 2013 http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM368806.pdf. Accessed April 22, 2019.

24. Dowsett E, Goudsmit E, Macintyre A, et al. Report from the national task force on chronic fatigue syndrome (CFS), post viral fatigue syndrome (PVFS), myalgic encephalomyelitis (ME). Westcare. 1994.

25. Jason LA, Bell DS, Rowe K, et al. A pediatric case definition for myalgic encephalomyelitis and chronic fatigue syndrome. Journal of Chronic Fatigue Syndrome. 2006;13(2-3):1-44. doi: 10.1300/J092v13n02_01.

26. Jason L, Evans M, Porter N, et al. The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. Am J Biochem Biotechnol. 2010;6(2):120-35.

27. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/topics/cer-methods-guide/overview. Accessed April 18, 2019.

28. Chou R, McDonagh M, Griffin J, et al. Diagnosis and treatment of myalgic encephalomyelitis/chronic fatigue syndrome: a systematic evidence review. PROSPERO: National Institute for Health Research; 2019. https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=142805. Accessed January 1, 2020.

29. Nisenbaum R, Jones JF, Unger ER, et al. A population-based study of the clinical course of chronic fatigue syndrome. Health Qual Life Outcomes. 2003;1:49. PMID: 14613572.

30. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20(3 Suppl):21-35. doi: 10.1016/s0749-3797(01)00261-6. PMID: 11306229.

31. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1198-207. doi: 10.1016/j.jclinepi.2010.11.021. PMID: 21463926.

32. Goligher EC, Pouchot J, Brant R, et al. Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. J Rheumatol. 2008;35(4):635-42. PMID: 18322987.

33. Wyrwich KW, Tierney WM, Babu AN, et al. A comparison of clinically important differences in health-related quality of life for patients with chronic lung disease, asthma, or heart disease. Health Serv Res. 2005;40(2):577-91. doi: 10.1111/j.1475-6773.2005.00373.x. PMID: 15762908.

34. Lemay KR, Tulloch HE, Pipe AL, et al. Establishing the minimal clinically important difference for the hospital anxiety and depression scale in patients with cardiovascular disease. J Cardiopulm Rehabil Prev. 2019;39(6):E6-e11. doi: 10.1097/hcr.0000000000000379. PMID: 30489438.

35. Cohen J. Statistical power analysis for the behavioural sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

36. Sutton AJ, Abrams KR, Jones DR, et al. Methods for meta-analysis in medical research: Wiley Chichester; 2000.

37. Fu R, Holmer HK. Change score or follow-up score? Choice of mean difference estimates could impact meta-analysis conclusions. J Clin Epidemiol. 2016;76:108-17. doi: 10.1016/j.jclinepi.2016.01.034. PMID: 26931293.

38. White PD, Goldsmith KA, Johnson AL, et al. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. Lancet. 2011;377(9768):823-36. doi: 10.1016/S0140-6736(11)60096-2. PMID: 21334061.

39. White PD, Sharpe MC, Chalder T, et al. Protocol for the PACE trial: a randomised controlled trial of adaptive pacing, cognitive behaviour therapy, and graded exercise as supplements to standardised specialist medical care versus standardised specialist medical care alone for patients with the chronic fatigue syndrome/ myalgic encephalomyelitis or encephalopathy. BMC Neuro. 2007;7(6) PMID: 17397525.

40. Goldsmith K, White P, Chalder T, et al. The PACE trial: Analysis of primary outcomes using composite measures of improvement. Wolfson Institute of Preventive Medicine, London, UK. 2016.

41. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58. doi: 10.1002/sim.1186. PMID: 12111919.

42. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

43. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. doi: 10.1136/bmj.d4002. PMID: 21784880.

44. AHRQ Methods for Effective Health Care. In: Chang SM, Matchar DB, Smetana GW, Umscheid CA, eds. Methods Guide for Medical Test Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.

45. Stadje R, Dornieden K, Baum E, et al. The differential diagnosis of tiredness: a systematic review. BMC Fam Pract. 2016;17(1):147. PMID: 27765009.

46. Brimmer DJ, Maloney E, Devlin R, et al. A pilot registry of unexplained fatiguing illnesses and chronic fatigue syndrome. BMC Res Notes. 2013;6:309. doi: 10.1186/1756-0500-6-309. PMID: 23915640.

47. Devasahayam A, Lawn T, Murphy M, et al. Alternative diagnoses to chronic fatigue syndrome in referrals to a specialist service: service evaluation survey. JRSM Short Rep. 2012;3(1):4. doi: 10.1258/shorts.2011.011127. PMID: 22299071.

48. Mariman A, Delesie L, Tobback E, et al. Undiagnosed and comorbid disorders in patients with presumed chronic fatigue syndrome. J Psychosom Res. 2013;75(5):491-6. doi: 10.1016/j.jpsychores.2013.07.010. PMID: 24182640.

49. Newton JL, Mabillard H, Scott A, et al. The Newcastle NHS Chronic Fatigue Syndrome Service: not all fatigue is the same. J R Coll Physicians Edinb. 2010;40(4):304-7. doi: 10.4997/JRCPE.2010.404. PMID: 21132135.

50. Nijrolder I, van der Windt D, de Vries H, et al. Diagnoses during follow-up of patients presenting with fatigue in primary care. CMAJ. 2009;181(10):683-7. doi: 10.1503/cmaj.090647. PMID: 19858240.

51. Slomko J, Newton JL, Kujawski S, et al. Prevalence and characteristics of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) in Poland: a cross-sectional study. BMJ Open. 2019;9(3):e023955. doi: 10.1136/bmjopen-2018-023955. PMID: 30850404.

52. Clark LV, Pesola F, Thomas JM, et al. Guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome (GETSET): a pragmatic randomised controlled trial. Lancet. 2017;390(10092):363-73. doi: 10.1016/S0140-6736(16)32589-2. PMID: 28648402.

53. Fulcher KY, White PD. Randomised controlled trial of graded exercise in patients with the chronic fatigue syndrome. BMJ. 1997;314(7095):1647-52. doi: 10.1136/bmj.314.7095.1647. PMID: 9180065.

54. Jason LA, Torres-Harding S, Friedberg F, et al. Non-pharmacologic interventions for CFS: a randomized trial. J Clin Psychol Med Settings. 2007;14(4):275-96.

55. Moss-Morris R, Sharon C, Tobin R, et al. A randomized controlled graded exercise trial for chronic fatigue syndrome: outcomes and mechanisms of change. J Health Psychol. 2005;10(2):245-59. PMID: 15723894.

56. Powell P, Bentall RP, Nye FJ, et al. Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome. BMJ. 2001;322(7283):387-90. PMID: 11179154.

57. Wallman KE, Morton AR, Goodman C, et al. Randomised controlled trial of graded exercise in chronic fatigue syndrome. Med J Aust. 2004;180(9):444-8. PMID: 15115421.

58. Wearden AJ, Dowrick C, Chew-Graham C, et al. Nurse led, home based self help treatment for patients in primary care with chronic fatigue syndrome: randomised controlled trial. BMJ. 2010;340:c1777. doi: 10.1136/bmj.c1777. PMID: 20418251.

59. Wearden AJ, Morriss RK, Mullis R, et al. Randomised, double-blind, placebo-controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome. Br J Psychiatry. 1998;172:485-90. PMID: 9828987.

60. Windthorst P, Mazurak N, Kuske M, et al. Heart rate variability biofeedback therapy and graded exercise training in management of chronic fatigue syndrome: an exploratory pilot study. J Psychosom Res. 2017;93:6-13. doi: 10.1016/j.jpsychores.2016.11.014. PMID: 28107894.

61. Powell P, Bentall RP, Nye FJ, et al. Patient education to encourage graded exercise in chronic fatigue syndrome. 2-year follow-up of randomised controlled trial. Br J Psychiatry. 2004;184:142-6. PMID: 14754826.

62. Dougall D, Johnson A, Goldsmith K, et al. Adverse events and deterioration reported by participants in the PACE trial of therapies for chronic fatigue syndrome. J Psychosom Res. 2014;77(1):20-6. doi: 10.1016/j.jpsychores.2014.04.002. PMID: 24913337.

63. Sharpe M, Goldsmith KA, Johnson AL, et al. Rehabilitative treatments for chronic fatigue syndrome: long-term follow-up from the PACE trial. Lancet Psychiatry. 2015;2(12):1067-74. doi: 10.1016/S2215-0366(15)00317-X. PMID: 26521770.

64. Bourke JH, Johnson AL, Sharpe M, et al. Pain in chronic fatigue syndrome: response to rehabilitative treatments in the PACE trial. Psychol Med. 2014;44(7):1545-52. doi: 10.1017/S0033291713002201. PMID: 23967878.

65. Wilshire C, Kindlon T, Matthees A, et al. Can patients with chronic fatigue syndrome really recover after graded exercise or cognitive behavioural therapy? A critical commentary and preliminary re-analysis of the PACE trial. Fatigue. 2017;5(1):43-56. doi: 10.1080/21641846.2017.1259724.

66. Wilshire CE, Kindlon T, Courtney R, et al. Rethinking the treatment of chronic fatigue syndrome-a reanalysis and evaluation of findings from a recent major trial of graded exercise and CBT. BMC Psychol. 2018;6(1):6. doi: 10.1186/s40359-018-0218-3. PMID: 29562932.

67. Sutcliffe K, Gray J, Tan MP, et al. Home orthostatic training in chronic fatigue syndrome-a randomized, placebo-controlled feasibility study. Eur J Clin Invest. 2010;40(1):18-24. doi: 10.1111/j.1365-2362.2009.02225.x. PMID: 19912315.

68. Burgess M, Andiappan M, Chalder T. Cognitive behaviour therapy for chronic fatigue syndrome in adults: face to face versus telephone treatment: a randomized controlled trial. Behav Cogn Psychother. 2012;40(2):175-91. doi: 10.1017/S1352465811000543. PMID: 21929831.

69. Deale A, Chalder T, Marks I, et al. Cognitive behavior therapy for chronic fatigue syndrome: a randomized controlled trial. Am J Psychiatry. 1997;154(3):408-14. PMID: 9054791.

70. Janse A, Worm-Smeitink M, Bleijenberg G, et al. Efficacy of web-based cognitive-behavioural therapy for chronic fatigue syndrome: randomised controlled trial. Br J Psychiatry. 2018;212(2):112-8. doi: 10.1192/bjp.2017.22. PMID: 29436329.

71. Knoop H, van der Meer JWM, Bleijenberg G. Guided self-instructions for people with chronic fatigue syndrome: randomised controlled trial. Br J Psychiatry. 2008;193(4):340-1. doi: 10.1192/bjp.bp.108.051292. PMID: 18827302.

72. Lopez C, Antoni M, Penedo F, et al. A pilot study of cognitive behavioral stress management effects on stress, quality of life, and symptoms in persons with chronic fatigue syndrome. J Psychosom Res. 2011;70(4):328-34. doi: 10.1016/j.jpsychores.2010.11.010. PMID: 21414452.

73. O'Dowd H, Gladwell P, Rogers CA, et al. Cognitive behavioural therapy in chronic fatigue syndrome: a randomised controlled trial of an outpatient group programme. Health Technol Assess. 2006;10(37):iii-iv, ix-x, 1-121. PMID: 17014748.

74. Sharpe M, Hawton K, Simkin S, et al. Cognitive behaviour therapy for the chronic fatigue syndrome: a randomized controlled trial. BMJ. 1996;312(7022):22-6. PMID: 8555852.

75. Stubhaug B, Lie SA, Ursin H, et al. Cognitive-behavioural therapy v. mirtazapine for chronic fatigue and neurasthenia: randomised placebo-controlled trial. Br J Psychiatry. 2008;192(3):217-23. doi: 10.1192/bjp.bp.106.031815. PMID: 18310583.

76. Tummers M, Knoop H, van Dam A, et al. Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial. Psychol Med. 2012;42(10):2205-15. doi: 10.1017/S0033291712000232. PMID: 22354999.

77. Wiborg JF, van Bussel J, van Dijk A, et al. Randomised controlled trial of cognitive behaviour therapy delivered in groups of patients with chronic fatigue syndrome. Psychother Psychosom. 2015;84(6):368-76. doi: 10.1159/000438867. PMID: 26402868.

78. Deale A, Husain K, Chalder T, et al. Long-term outcome of cognitive behavior therapy versus relaxation therapy for chronic fatigue syndrome: a 5-year follow-up study. Am J Psychiatry. 2001;158(12):2038-42. PMID: 11729022.

79. Taylor RR. Quality of life and symptom severity for individuals with chronic fatigue syndrome: findings from a randomized clinical trial. Am J Occup Ther. 2004;58(1):35-43. PMID: 14763634.

80. Rimes KA, Wingrove J. Mindfulness-based cognitive therapy for people with chronic fatigue syndrome still experiencing excessive fatigue after cognitive behaviour therapy: a pilot randomized study. Clin Psychol Psychother. 2013;20(2):107-17. doi: 10.1002/cpp.793. PMID: 21983916.

81. Surawy C, Roberts J, Silver A. The effect of mindfulness training on mood and measures of fatigue, activity, and quality of life in patients with chronic fatigue syndrome on a hospital waiting list: a series of exploratory studies. Behav Cogn Psychother. 2005;33(1):103-9. doi: 10.1017/S135246580400181X.

82. Friedberg F, Adamowicz J, Caikauskaite I, et al. Efficacy of two delivery modes of behavioral self-management in severe chronic fatigue syndrome. Fatigue. 2016;4(3):158-74. doi: 10.1080/21641846.2016.1205876.

83. Pinxsterhuis I, Sandvik L, Strand EB, et al. Effectiveness of a group-based self-management program for people with chronic fatigue syndrome: a randomized controlled trial. Clin Rehabil. 2017;31(1):93-103. doi: 10.1177/0269215515621362. PMID: 26672998.

84. Al-Haggar MS, Al-Naggar ZA, Abdel-Salam MA. Biofeedback and cognitive behavioral therapy for Egyptian adolescents suffering from chronic fatigue syndrome. J Pediatr Neurol. 2006;4(3):161-9. doi: 10.1055/s-0035-1557320.

85. Chalder T, Deary V, Husain K, et al. Family-focused cognitive behaviour therapy versus psycho-education for chronic fatigue syndrome in 11- to 18-year-olds: a randomized controlled treatment trial. Psychol Med. 2010;40(8):1269-79. doi: 10.1017/S003329170999153X. PMID: 19891804.

86. Nijhof SL, Bleijenberg G, Uiterwaal CS, et al. Effectiveness of internet-based cognitive behavioural treatment for adolescents with chronic fatigue syndrome (FITNET): a randomised controlled trial. Lancet. 2012;379(9824):1412-8. doi: 10.1016/S0140-6736(12)60025-7. PMID: 22385683.

87. Stulemeijer M, de Jong LW, Fiselier TJ, et al. Cognitive behaviour therapy for adolescents with chronic fatigue syndrome: randomised controlled trial. BMJ. 2005;330(7481):14. PMID: 15585538.

88. Wright B, Ashby B, Beverley D, et al. A feasibility study comparing two treatment approaches for chronic fatigue syndrome in adolescents. Arch Dis Child. 2005;90(4):369-72. PMID: 15781925.

89. Crawley E, Gaunt D, Garfield K, et al. Erratum: clinical and cost-effectiveness of the Lightning Process in addition to specialist medical care for paediatric chronic fatigue syndrome: randomised controlled trial. Arch Dis Child. 2019;104(10):e3. doi: 10.1136/archdischild-2017-313375. PMID: 31296601.

90. Lloyd S, Chalder T, Rimes KA. Family-focused cognitive behaviour therapy versus psycho-education for adolescents with chronic fatigue syndrome: long-term follow-up of an RCT. Behav Res Ther. 2012;50(11):719-25. doi: 10.1016/j.brat.2012.08.005. PMID: 22985998.

91. Nijhof SL, Priesterbach LP, Uiterwaal CS, et al. Internet-based therapy for adolescents with chronic fatigue syndrome: long-term follow-up. Pediatrics. 2013;131(6):e1788-95. doi: 10.1542/peds.2012-2007. PMID: 23669515.

92. Reeves WC, Lloyd A, Vernon SD, et al. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. BMC Health Serv Res. 2003;3(1):25. PMID: 14702202.

93. Bentall RP, Powell P, Nye FJ, et al. Predictors of response to treatment for chronic fatigue syndrome. Br J Psychiatry. 2002;181:248-52. PMID: 12204931.

94. Fluge O, Bruland O, Risa K, 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. doi: 10.1371/journal.pone.0026358. PMID: 22039471.

95. Fluge O, Rekeland IG, Lien K, et al. B-lymphocyte depletion in patients with myalgic encephalomyelitis/chronic fatigue syndrome: a randomized, double-blind, placebo-controlled trial. Ann Intern Med. 2019;170(9):585-93. doi: 10.7326/m18-1451. PMID: 30934066.

96. McKenzie R, O'Fallon A, Dale J, et al. Low-dose hydrocortisone for treatment of chronic fatigue syndrome: a randomized controlled trial. JAMA. 1998;280(12):1061-6. PMID: 9757853.

97. McKenzie R, Reynolds JC, O'Fallon A, et al. Decreased bone mineral density during low dose glucocorticoid administration in a randomized, placebo controlled trial. J Rheumatol. 2000;27(9):2222-6. PMID: 10990237.

98. Montoya JG, Kogelnik AM, Bhangoo M, et al. Randomized clinical trial to evaluate the efficacy and safety of valganciclovir in a subset of patients with chronic fatigue syndrome. J Med Virol. 2013;85(12):2101-9. doi: 10.1002/jmv.23713. PMID: 23959519.

99. See DM, Tilles JG. Alpha-interferon treatment of patients with chronic fatigue syndrome. Immunol Invest. 1996;25(1-2):153-64. PMID: 8675231.

100. Strayer DR, Carter WA, Brodsky I, et al. A controlled clinical trial with a specifically configured RNA drug, poly(I) midline dot poly(C12U), in chronic fatigue syndrome. Clin Infect Dis. 1994;18(SUPPL. 1):S88-S95. PMID: 8148460.

101. Strayer DR, Carter WA, Stouch BC, 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. doi: 10.1371/journal.pone.0031334. PMID: 22431963.

102. Sulheim D, Fagermoen E, Winger A, et al. Disease mechanisms and clonidine treatment in adolescent chronic fatigue syndrome: a combined cross-sectional and randomized clinical trial. JAMA Pediatr. 2014;168(4):351-60. doi: 10.1001/jamapediatrics.2013.4647. PMID: 24493300.

103. Vercoulen JH, Swanink CM, Zitman FG, et al. Randomised, double-blind, placebo-controlled study of fluoxetine in chronic fatigue syndrome. Lancet. 1996;347(9005):858-61. PMID: 8622391.

104. Vollmer-Conna U, Hickie I, Hadzi-Pavlovic D, et al. Intravenous immunoglobulin is ineffective in the treatment of patients with chronic fatigue syndrome. Am J Med. 1997;103(1):38-43. PMID: 9236484.

105. Peterson PK, Shepard J, Macres M, et al. A controlled trial of intravenous immunoglobulin G in chronic fatigue syndrome. Am J Med. 1990;89(5):554-60. PMID: 2239975.

106. Roerink ME, Bredie SJH, Heijnen M, et al. Cytokine inhibition in patients with chronic fatigue syndrome: a randomized trial. Ann Intern Med. 2017;166(8):557-64. doi: 10.7326/M16-2391. PMID: 28265678.

107. Rowe KS. Double-blind randomized controlled trial to assess the efficacy of intravenous gammaglobulin for the management of chronic fatigue syndrome in adolescents. J Psychiatr Res. 1997;31(1):133-47. PMID: 9201655.

108. Blacker CVR, Greenwood DT, Wesnes KA, et al. Effect of galantamine hydrobromide in chronic fatigue syndrome: a randomized controlled trial. JAMA. 2004;292(10):1195-204. doi: 10.1001/jama.292.10.1195. PMID: 15353532.

109. Montoya JG, Anderson JN, Adolphs DL, et al. KPAX002 as a treatment for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): a prospective, randomized trial. Int J Clin Exp Med. 2018;11(3):2890-900.

110. Arnold LM, Blom TJ, Welge JA, et al. 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):242-53. doi: 10.1016/j.psym.2014.12.003. PMID: 25660434.

111. Blockmans D, Persoons P, Van Houdenhove B, et al. Combination therapy with hydrocortisone and fludrocortisone does not improve symptoms in chronic fatigue syndrome: a randomized, placebo-controlled, double-blind, crossover study. Am J Med. 2003;114(9):736-41. PMID: 12829200.

112. Malaguarnera M, Gargante MP, Cristaldi E, et al. Acetyl L-carnitine (ALC) treatment in elderly patients with fatigue. Arch Gerontol Geriatr. 2008;46(2):181-90. PMID: 17658628.

113. Ostojic SM, Stojanovic M, Drid P, et al. Supplementation with guanidinoacetic acid in women with chronic fatigue syndrome. Nutrients. 2016;8(2):72. doi: 10.3390/nu8020072. PMID: 26840330.

114. The GKH, Bleijenberg G, van der Meer JWM. The effect of acclydine in chronic fatigue syndrome: a randomized controlled trial. PLoS Clin Trials. 2007;2(5):e19. PMID: 17525791.

115. Williams G, Waterhouse J, Mugarza J, et al. Therapy of circadian rhythm disorders in chronic fatigue syndrome: no symptomatic improvement with melatonin or phototherapy. Eur J Clin Invest. 2002;32(11):831-7. PMID: 12423324.

116. Öckerman PA. Antioxidant treatment of chronic fatigue syndrome. Clin Pract Alternat Med. 2000;1(2):88-91.

117. Weatherley-Jones E, Nicholl JP, Thomas KJ, et al. A randomised, controlled, triple-blind trial of the efficacy of homeopathic treatment for chronic fatigue syndrome. J Psychosom Res. 2004;56(2):189-97. PMID: 15016577.

118. Hobday RA, Thomas S, O'Donovan A, et al. Dietary intervention in chronic fatigue syndrome. J Hum Nutr Diet. 2008;21(2):141-9. doi: 10.1111/j.1365-277X.2008.00857.x. PMID: 18339054.

119. Vermeulen RCW, Scholte HR. Exploratory open label, randomized study of acetyl- and propionylcarnitine in chronic fatigue syndrome. Psychosom Med. 2004;66(2):276-82. PMID: 15039515.

120. Li DQ, Li ZC, Dai ZY. Selective serotonin reuptake inhibitor combined with dengzhanshengmai capsule improves the fatigue symptoms: a 12-week open-label pilot study. Int J Clin Exp Med. 2015;8(7):11811-7. PMID: 26380022.

121. Chan JSM, Ho RTH, Wang CW, et al. Effects of qigong exercise on fatigue, anxiety, and depressive symptoms of patients with chronic fatigue syndrome-like illness: a randomized controlled trial. Evid Based Complement Alternat Med. 2013 doi: 10.1155/2013/485341. PMID: 23983785.

122. Dybwad M, Frøslie K, Stanghelle J. Work capacity, fatigue and health related quality of life in patients with myalgic encephalopathy or chronic fatigue syndrome, before and after qigong therapy, a randomized controlled study. Nesoddtangen, Norway: Sunnaas Rehabilitation Hospital. 2007.

123. Ho RTH, Chan JSM, Wang C-W, et al. A randomized controlled trial of qigong exercise on fatigue symptoms, functioning, and telomerase activity in persons with chronic fatigue or chronic fatigue syndrome. Ann Behav Med. 2012;44(2):160-70. doi: 10.1007/s12160-012-9381-6. PMID: 22736201.

124. Oka T, Tanahashi T, Chijiwa T, et al. Isometric yoga improves the fatigue and pain of patients with chronic fatigue syndrome who are resistant to conventional therapy: a randomized, controlled trial. Biopsychosoc Med. 2014;14(27):1-9. doi: 10.1186/s13030-014-0027-8. PMID: 25525457.

125. Huanan L, Wang J, Zhang W, et al. Chronic fatigue syndrome treated by the traditional Chinese procedure abdominal tuina: a randomized controlled clinical trial. J Tradit Chin Med. 2017;37(6):819-26. doi: 10.1016/S0254-6272(18)30046-3.

126. Walach H, Bosch H, Lewith G, et al. Effectiveness of distant healing for patients with chronic fatigue syndrome: a randomised controlled partially blinded trial (EUHEALS). Psychother Psychosom. 2008;77(3):158-66. doi: 10.1159/000116609. PMID: 18277062.

127. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. BMJ. 2000;321(7259):504. doi: 10.1136/bmj.321.7259.504. PMID: 10948038.

128. VanNess JM, Stevens SR, Bateman L, et al. Postexertional malaise in women with chronic fatigue syndrome. J Womens Health (Larchmt). 2010;19(2):239-44. doi: 10.1089/jwh.2009.1507. PMID: 20095909.

129. Montgomery J. Re: The PACE trial and the Committee's inquiry on Research Integrity. 2019. https://www.parliament.uk/globalassets/documents/commons-committees/science-technology/Correspondence/190129-Sir-Jonathan-Montgomery-Health-Research-Authority-to-Chair-re-PACE-trial.pdf Accessed October 21, 2021.

130. Whitehead L, Campion P. Can general practitioners manage chronic fatigue syndrome? A controlled trial. J Chronic Fatigue Syndr. 2002;10(1):55-64. doi: 10.1300/J092v10n01_05.

131. Larun L, Brurberg KG, Odgaard-Jensen J, et al. Exercise therapy for chronic fatigue syndrome. Cochrane Database Syst Rev. 2019;10:Cd003200. doi: 10.1002/14651858.CD003200.pub8. PMID: 31577366.

132. Collatz A, Johnston SC, Staines DR, et al. A systematic review of drug therapies for chronic fatigue syndrome/myalgic encephalomyelitis. Clin Ther. 2016;38(6):1263-71.e9. doi: 10.1016/j.clinthera.2016.04.038. PMID: 27229907.

133. Price JR, Mitchell E, Tidy E, et al. Cognitive behaviour therapy for chronic fatigue syndrome in adults. Cochrane Database Syst Rev. 2009 (2).

134. NINDS Common Data Elements: Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. https://www.commondataelements.ninds.nih.gov/Myalgic%20Encephalomyelitis/Chronic%20Fatigue%20Syndrome. Accessed 2-16-2022.

135. Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. BMJ. 2005;330(7482):88. doi: 10.1136/bmj.330.7482.88. PMID: 15637373.

Abbreviations and Acronyms

Abbreviation	Definition
ACT	anaerobic activity therapy
ADL	activities of daily living
AHRQ	Agency for Healthcare Research and Quality
AMD	adjusted mean difference
ANOVA	analysis of variance
AP	anteroposterior
APT	adaptive pacing therapy
ARD	adjusted risk difference
BDI	Beck Depression Inventory
BMI	body mass index
BPI	Brief Pain Inventory
CBT	cognitive behavioral therapy
CDC	Centers for Disease Control and Prevention
CDs	compact discs
CFS	chronic fatigue syndrome
CGI	Clinical Global Impression of Change
CGS-S	Clinical Global Impression Severity Score
CHQ-CF	child health questionnaire-child form
CI	confidence interval
CIBEROBN	Ventro de Investagacion Biomedica en Red de Fisiopatologia de la Obesidad y Nutricion
CNS	central nervous system
COG	cognitive therapy
COPD	Chronic Obstructive Pulmonary Disease
Ctr	Counter
DF	degrees of freedom
DSM-III-R	Diagnostic Statistical Manual third edition revised
DSM-IV	Diagnostic Statistical Manual IV
EPC	Evidence-based Practice Center
ESS	Epworth Sleepiness Scale
FDA	U.S. Food and Drug Administration
FINE	Fatigue Intervention by Nurses Evaluation
FIQ	Fibromyalgia Impact Questionnaire
FITNET	fatigue in teenagers on the internet
FSM	fatigue self-management
FSM:ACT	fatigue self-management with web diaries and actigraphs
FSM:CTR	fatigue self-management with paper diaries and step counters
GAA	guadidinoacetic acid
GES	guided graded exercise self-help
GET	graded exercise therapy
GETSET	guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome
GHQ	general health questionnaire
HADS	Hospital Anxiety and Depression Scale
HADS-A	Hospital Anxiety and Depression Scale-anxiety
HADS-D	Hospital Anxiety and Depression Scale-depression
HAM-D	Hamilton Depression Rating Scale
HHV-6	human herpes virus-6
HRSD	Hamilton Rating Scale
HTA	Health Technology Assessment

iCBT	internet-based cognitive-behavioral therapy
ICD-10	International Statistical Classification of Diseases and Related Health Problems-10th revision
IGF1	insulin-like growth factor-1
IGFBP3	insulin like growth factor binding protein 3
IgG	immunoglobulin G
IOM	Institute of Medicine
IQR	interquartile range
ITT	intention to treat
IV	Instrumental variable
IV	intravenous
KFSS	Krupp Fatigue Severity Scale
KPS	Karnofsky Performance Scale
KSQ	Karolinska Sleep Questionnaire
MBCT	mindfulness-based cognitive therapy
MCT	multi convergent therapy
MD	mean difference
MDD	major depressive disorder
ME	myalgic encephalomyelitis
MFI-20	Multidimensional Fatigue Inventory 20-item
M-H	Mantel-Haenszel test
MOS	Medical Outcome Study
MOS-SF	Medical Outcome Study – Short Form
MRI	magnetic resonance imaging
NAFKAM	Norway's National Research Center in Complementary and Alternative Medicine
NH&MRC	National Health and Medical Research Council
NHS	National Health Service
NIAID	National Institute of Allergy and Infectious Diseases
NICE	National Institute for Health and Care Excellence
NIH	National Institute of Health
NNT	number needed to treat
NR	not reported
NS	not significant
NSAID	nonsteroidal anti-inflammatory drug
OR	odds ratio
PACE	pacing, graded activity, cognitive behavior therapy; a randomized Evaluation
PF	physical function
PHQ	patient health questionnaire
PICOTS	populations, interventions, comparators, outcomes, timing, and setting/study design
POMS	profile of mood states
QLI	quality of life index
QLS	quality of life score
QOL	Quality of Life
QOLI	quality of life inventory
QOL-SF	quality of life short form
RCT	randomized controlled trial
RR	relative risk
SAE	serious adverse event
SCL-90	symptom checklist 90
SCL-90-R	symptom checklist 90-revised
SD	standard deviation
SE	standard error

SEID	systemic exertion intolerance disease
SEM	standard error of the mean
SES	standardized effect sizes
SF-12	12-item Short Form Health Survey
SF-36	36-item Short Form Health Survey
SGR	support the activities of research groups
SIP-8	Sickness Impact Profile 8-item
SMC	specialist medical care
SMD	standardized mean difference
SOE	strength of evidence
SSRI	selective serotonin reuptake inhibitor
Std	standard
VAS	visual analogue scale
WMD	weighted mean difference
ZonMW	ZorgOnderzoek Nederland and Medische wetenschappen

Key Informants

Key Informants representing clinical, research, or patient perspectives in ME/CFS participated in calls with the EPC and CDC during the development of the key questions. Broad expertise and perspectives were sought. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Key Informants were not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants were asked to disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained.

Key Informants:

Lily Chu, MD, MSHS
International Association for CFS/ME

Mary Dimmock, BA
Connecticut CFIDS/ME & FM Association

Stephen Gluckman, MD
University of Pennsylvania

Nancy Klimas, MD
Nova University

Gudrun Lange, PhD
Mount Sinai

Allison Ramiller, MPH
Solve ME/CFS Initiative

Peter Rowe, MD
Johns Hopkins University

Christopher Snell, PhD
Workwell Foundation

Peer Reviewers

Prior to publication of the final evidence report, the EPC sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers. Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained.

Peer Reviewers:

Lily Chu, MD, MSHS
International Association for CFS/ME

Mary Dimmock, BA
Connecticut CFIDS/ME & FM Association

Teresa Dowell, DNP, PT
Four Peaks Healthcare Associates

Stephen Gluckman, MD
University of Pennsylvania

Ben Z. Katz, MD
Northwestern University Feinberg School of Medicine
An & Robert H Lurie Children's Hospital of Chicago

Gudrun Lange, PhD
Mount Sinai

Susan Levine, MD
Cornell Visiting Scientist
Clinician in Private Practice

Christopher Snell, PhD
Workwell Foundation