Drug Class Review

Drugs for Fibromyalgia

Final Original Report

April 2011

The Agency for Healthcare Research and
Quality has not yet seen or approved this report

The purpose of Drug Effectiveness Review Project reports is to make available information regarding the comparative clinical effectiveness and harms of different drugs. Reports are not usage guidelines, nor should they be read as an endorsement of or recommendation for any particular drug, use, or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports.

Beth Smith, DO
Kim Peterson, MS
Rochelle Fu, PhD
Sujata Thakurta, MPA:HA
Allison Low, BA
Marian McDonagh, PharmD

Drug Effectiveness Review Project
Marian McDonagh, PharmD, Principal Investigator

Oregon Evidence-based Practice Center
Mark Helfand, MD, MPH, Director

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

STRUCTURED ABSTRACT

Purpose

We compared the effectiveness and harms of tricyclic antidepressants, serotonin norepinephrine reuptake inhibitors, selective serotonin reuptake inhibitors, selective serotonin and norepinephrine reuptake inhibitors, noradrenergic and specific serotonergic reuptake inhibitor, norepinephrine and dopamine reuptake inhibitor, serotonin receptor antagonist, antiepileptic drugs, and skeletal muscle relaxants in adults with fibromyalgia.

Data Sources

We searched Ovid MEDLINE®, the Cochrane Database of Systematic Reviews®, and the Cochrane Central Register of Controlled Trials® and Database of Abstracts of Reviews of Effects through October 2010. For additional data we also hand searched reference lists, US Food and Drug Administration medical and statistical reviews and dossiers submitted by pharmaceutical companies.

Review Methods

Study selection, data abstraction, validity assessment, grading the strength of the evidence, and data synthesis were all carried out according to standard Drug Effectiveness Review Project review methods.

Results and Conclusions

We found eligible studies of treatment for fibromyalgia with amitriptyline, nortriptyline, citalopram, fluoxetine, paroxetine, cyclobenzaprine, pregabalin, gabapentin, milnacipran, and duloxetine. We found no eligible studies with the other included drugs and no eligible studies of included interventions when used as adjunctive therapy. Head-to-head trials were few, and provided low-strength evidence that short-term treatment with immediate-release paroxetine is superior to amitriptyline in reducing pain and sleep disturbance and provided low-strength evidence there are no significant differences between amitriptyline as compared with cyclobenzaprine and nortriptyline. Although there were some significant differences between drugs in overall adverse events, they did not produce any differences in withdrawals due to adverse events. Additionally, based on indirect comparison meta-analysis, we found low evidence that duloxetine was superior to milnacipran on outcomes of pain, sleep disturbance, depressed mood, and health-related quality of life. We found low evidence that both duloxetine and milnacipran were superior to pregabalin on improvement in depressed mood, whereas pregabalin was superior to milnacipran on improvement in sleep disturbance. Amitriptyline was similar to duloxetine, milnacipran, and pregabalin on outcomes of pain and fatigue, with insufficient data on the other outcomes. Although there were some significant differences between duloxetine, milnacipran, and pregabalin in specific adverse events, they did not produce any differences in overall withdrawals, overall adverse events, and withdrawals due to adverse events. For the remaining drugs, there was only evidence of significant improvements in pain over placebo in 1 trial for gabapentin, in 1 of 3 trials for cyclobenzaprine, and in 1 trial of

fluoxetine. But, no conclusions can be drawn about comparative effectiveness or harms among these drugs because the numbers of trials/patients in placebo-controlled trials were too few to provide meaningful results in indirect comparisons. Duloxetine was not effective on pain reduction in male, nonwhite, and older patients based on a small sample size that was underpowered to detect a difference. Compared with placebo, duloxetine, fluoxetine, controlled-release paroxetine, and pregabalin significantly improved fibromyalgia symptoms regardless of baseline depression but milnacipran was only effective in nondepressed patients. Controlled-release paroxetine and pregabalin significantly improved fibromyalgia symptoms regardless of baseline anxiety.

TABLE OF CONTENTS

TABLES

FIGURES

APPENDIXES

EVIDENCE TABLES

Published in a separate document.

Acknowledgments
We thank Leah Williams, our publications editor, for putting this report into its present form for you to read.

Clinical Advisory Group
We extend our appreciation to the clinical advisors listed below for their thoughtful advice and input during our research process.

Manfred Harth, MD
Professor Emeritus, Medicine
University of Western Ontario
Canadian Rheumatology Association

Mary-Ann Fitzcharles
Associate Professor, McGill University
Director Fibromyalgia Clinic McGill Pain Center, Montreal General Hospital
Quebec Pain Society Executive Committee
Arthritis Society of Canada, medical Advisory Committee

Suggested citation for this report
Smith B, Peterson K, Fu R, Thakurta S, Low A, McDonagh M. Drugs for fibromyalgia. Final original report. Prepared by the Oregon Evidence-based Practice Center for the Drug Effectiveness Review Project. Oregon Health & Science University. Portland, OR. 2011. Available at: http://derp.ohsu.edu/about/final-document-display.cfm

Funding
The Drug Effectiveness Review Project, composed of 12 organizations including 11 state Medicaid agencies, and the Canadian Agency for Drugs and Technology in Health commissioned and funded for this report. These organizations selected the topic of the report and had input into its Key Questions. The content and conclusions of the report were entirely determined by the Evidence-based Practice Center researchers. The authors of this report have no financial interest in any company that makes or distributes the products reviewed in this report.

INTRODUCTION

Fibromyalgia syndrome is a disorder characterized by widespread pain as well as a constellation of other symptoms most commonly including sleep disorders, fatigue, and emotional or cognitive disturbances. The American College of Rheumatology first issued criteria for diagnosing fibromyalgia in 1990, which included widespread pain, defined as axial, bilateral, and both upper and lower segment pain, duration of greater than 3 months, and the presence of ≥ 11 out of 18 tender points (see Appendix A). Prior to this time, terms such as fibrositis and functional somatic syndrome were used with varying criteria including presence of additional symptoms such as depression, sleep disorder, and fatigue, symptoms that are often seen in patients with fibromyalgia but not considered diagnostic. Much controversy has existed over the past 2 decades regarding the validity of fibromyalgia as a clinical entity as well as the validity of the diagnostic criteria used to identify individuals with the disorder. In May 2010, the American College of Rheumatology published their new criteria for diagnosing fibromyalgia, eliminating the requirement of tender points and expanding it to include 3 conditions:[2]

1. Widespread pain index ≥ 7 and symptom severity scale score ≥ 5 or widespread pain index 3-6 and symptom severity scale score ≥ 9
2. Symptoms have been present at a similar level for at least 3 months
3. The patient does not have a disorder that would otherwise explain the pain.

These new criteria were designed to aid in the diagnosis of fibromyalgia particularly in the primary care clinic where tender point examination was either not carried out or not performed reliably, and to recognize the increasing understanding of associated cognitive and somatic symptoms associated with the disorder.[3] It was also designed to aid in monitoring of patients diagnosed with fibromyalgia. Previously, as one improved in their condition, they may no longer satisfy the criteria for diagnosis due to reduction in muscle tenderness. With the new definition, the Symptom Severity scale can serve as a monitoring tool for symptom severity without removing the diagnosis.

Given the diagnostic challenges and the fact that the criteria for diagnosis have recently changed, it is difficult to obtain accurate prevalence data. Two studies, one in Wichita, Kansas in 1993, and the other in Ontario, Canada in 1999, provided data on the prevalence of fibromyalgia.[4, 5] Both populations were primarily Caucasian which is different from the overall North American population. In Kansas, the overall prevalence among adults was 2% (95% CI, 1.4 to 2.7) with the prevalence being higher among women than men (3.4% compared with 0.5%), and with the prevalence rising after age 50 with a peak of 7.4% in women age 70-79 years.[4] The peak prevalence in men was 1% in the same age group. In the Canadian study, the prevalence was slightly higher among adults at 3.3% (95% CI, 3.2 to 3.4); 4.9% women and 1.6% men.[5] The peak prevalence in women was in the 55-64 year age group at 7.9% and for men was in the 45-54 year age group at 2.5%. A cohort study of primarily Utah residents was designed to determine the incidence of fibromyalgia and found that the age-adjusted incidence rates were 6.88 cases per 1000 person-years for males and 11.28 cases per 1000 person-years for females.[6] Unlike the prior studies, the female to male ratio was 1.64.[6]

The underlying etiology of fibromyalgia remains unclear but evidence supports complex functional changes to both the peripheral and central nervous system. Current theories include a peripheral tissue sensitization occurring after injury, a pain amplification syndrome that may

arise from sensitization of the central nervous system and/or peripheral tissue abnormalities, changes in descending noxious inhibitory control, and psychological risk factors that include somatization or increased focus on body symptoms, negative life events, psychological distress, and passive pain coping mechanisms.[7-12] One theory is that nociceptive activity in the peripheral tissue can lead to maintenance of a sustained sensitized state centrally, resulting in chronic pain. The release of substance P, interleukin-1 and 6, tumor necrosis factor, nitric oxide, amino acids, and prostaglandin have all been invoked.[7, 9-12] Given the lack of definitive understanding of the pathophysiology of fibromyalgia and given the constellation of symptoms associated with the disorder, choosing an effective therapy for fibromyalgia has been challenging. A multimodal treatment approach has been recommended, including both pharmacologic and nonpharmacologic therapies.[13] A myriad of pharmacological approaches have been pursued in the hope of finding effective options (Table 1).

The objective of this study is to review evidence on the comparative effectiveness/efficacy and comparative harms of the drugs used to treat fibromyalgia, and to determine if there are any subgroups of patients based on demographics, socioeconomic status, other medications, or comorbidities for which any included drugs are more effective or associated with fewer harms.

Purpose and Limitations of Systematic Reviews

Systematic reviews, also called evidence reviews, are the foundation of evidence-based practice. They focus on the strength and limits of evidence from studies about the effectiveness of a clinical intervention. Systematic reviews begin with careful formulation of research questions. The goal is to select questions that are important to patients and clinicians then to examine how well the scientific literature answers those questions. Terms commonly used in systematic reviews, such as statistical terms, are provided in Appendix B and are defined as they apply to reports produced by the Drug Effectiveness Review Project.

Systematic reviews emphasize the patient's perspective in the choice of outcome measures used to answer research questions. Studies that measure health outcomes (events or conditions that the patient can feel, such as fractures, functional status, and quality of life) are preferred over studies of intermediate outcomes (such as change in bone density). Reviews also emphasize measures that are easily interpreted in a clinical context. Specifically, measures of *absolute risk* or the probability of disease are preferred to measures such as relative risk. The difference in absolute risk between interventions depends on the number of events in each group, such that the difference (absolute risk reduction) is smaller when there are fewer events. In contrast, the difference in relative risk is fairly constant between groups with different baseline risk for the event, such that the difference (relative risk reduction) is similar across these groups. Relative risk reduction is often more impressive than absolute risk reduction. Another useful measure is the *number needed to treat* (or harm). The number needed to treat is the number of patients who would need be treated with an intervention for 1 additional patient to benefit (experience a positive outcome or avoid a negative outcome). The absolute risk reduction is used to calculate the number needed to treat.

Systematic reviews weigh the quality of the evidence, allowing a greater contribution from studies that meet high methodological standards and, thereby, reducing the likelihood of biased results. In general, for questions about the relative benefit of a drug, the results of well-executed randomized controlled trials are considered better evidence than results of cohort, case-

control, and cross-sectional studies. In turn, these studies provide better evidence than uncontrolled trials and case series. For questions about tolerability and harms, observational study designs may provide important information that is not available from controlled trials. Within the hierarchy of observational studies, well-conducted cohort designs are preferred for assessing a common outcome. Case-control studies are preferred only when the outcome measure is rare and the study is well conducted.

Systematic reviews pay particular attention to whether results of *efficacy studies* can be generalized to broader applications. Efficacy studies provide the best information about how a drug performs in a controlled setting. These studies attempt to tightly control potential confounding factors and bias; however, for this reason the results of efficacy studies may not be applicable to many, and sometimes to most, patients seen in everyday practice. Most efficacy studies use strict eligibility criteria that may exclude patients based on their age, sex, adherence to treatment, or severity of illness. For many drug classes, including the antipsychotics, unstable or severely impaired patients are often excluded from trials. In addition, efficacy studies frequently exclude patients who have comorbid disease, meaning disease other than the one under study. Efficacy studies may also use dosing regimens and follow-up protocols that are impractical in typical practice settings. These studies often restrict options that are of value in actual practice, such as combination therapies and switching to other drugs. Efficacy studies also often examine the short-term effects of drugs that in practice are used for much longer periods. Finally, efficacy studies tend to assess effects by using objective measures that do not capture all of the benefits and harms of a drug or do not reflect the outcomes that are most important to patients and their families.

Systematic reviews highlight studies that reflect actual clinical *effectiveness* in unselected patients and community practice settings. Effectiveness studies conducted in primary care or office-based settings use less stringent eligibility criteria, more often assess health outcomes, and have longer follow-up periods than most efficacy studies. The results of effectiveness studies are more applicable to the "average" patient than results from the highly selected populations in efficacy studies. Examples of effectiveness outcomes include quality of life, frequency or duration of hospitalizations, social function, and the ability to work. These outcomes are more important to patients, family, and care providers than surrogate or intermediate measures, such as scores based on psychometric scales.

Efficacy and effectiveness studies overlap. For example, a study might use very narrow inclusion criteria like an efficacy study, but, like an effectiveness study, might examine flexible dosing regimens, have a long follow-up period, and measure quality of life and functional outcomes. For this report we sought evidence about outcomes that are important to patients and would normally be considered appropriate for an effectiveness study. However, many of the studies that reported these outcomes were short-term and used strict inclusion criteria to select eligible patients. For these reasons, it was neither possible nor desirable to exclude evidence based on these characteristics. Labeling a study as either an efficacy or an effectiveness study, although convenient, is of limited value; it is more useful to consider whether the patient population, interventions, time frame, and outcomes are relevant to one's practice or to a particular patient.

Studies anywhere on the continuum from efficacy to effectiveness can be useful in comparing the clinical value of different drugs. Effectiveness studies are more applicable to practice, but efficacy studies are a useful scientific standard for determining whether characteristics of different drugs are related to their effects on disease. Systematic reviews

thoroughly cover the efficacy data in order to ensure that decision makers can assess the scope, quality, and relevance of the available data. This thoroughness is not intended to obscure the fact that efficacy data, no matter how large the quantity, may have limited applicability to practice. Clinicians can judge the relevance of study results to their practice and should note where there are gaps in the available scientific information.

Unfortunately, for many drugs there exist few or no effectiveness studies and many efficacy studies. Yet clinicians must decide on treatment for patients who would not have been included in controlled trials and for whom the effectiveness and tolerability of the different drugs are uncertain. Systematic reviews indicate whether or not there exists evidence that drugs differ in their effects in various subgroups of patients, but they do not attempt to set a standard for how results of controlled trials should be applied to patients who would not have been eligible for them. With or without an evidence report, these decisions must be informed by clinical judgment.

In the context of development of recommendations for clinical practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of an intervention are based on strong evidence from clinical studies. By themselves, they do not say what to do. Judgment, reasoning, and applying one's values under conditions of uncertainty must also play a role in decision making. Users of an evidence report must also keep in mind that *not proven* does not mean *proven not*; that is, if the evidence supporting an assertion is insufficient, it does not mean the assertion is untrue. The quality of the evidence on effectiveness is a key component, but not the only component, in making decisions about clinical policy. Additional criteria include acceptability to physicians and patients, potential for unrecognized harm, applicability of the evidence to practice, and consideration of equity and justice.

Scope and Key Questions

The goal of this report is to compare the effectiveness and harms of drugs for the treatment of fibromyalgia. The Oregon Evidence-based Practice Center wrote preliminary key questions, identifying the populations, interventions, outcomes of interest, and, based on these, eligibility criteria for studies. A draft of these questions and inclusion and exclusion criteria were posted on the Drug Effectiveness Review Project website for public comment. A group of clinicians specializing in treating patients with fibromyalgia were consulted for clinical insight into the proposed key questions. The draft was reviewed and revised by representatives of the organizations participating in the Drug Effectiveness Review Project. Revision took into consideration input from the public and from clinical advisors and the organizations' desire for the key questions to reflect populations, drugs, and outcome measures of interest to clinicians and patients.

When the scope of the review was originally finalized in July of 2010, the eligibility criteria for populations was limited to only those studies that based their diagnosis of fibromyalgia on either the 1990 or 2010 American College of Rheumatology criteria. However, later in the review process, considering that the 2010 changes to the American College of Rheumatology criteria for diagnosing fibromyalgia involved the removal of tender points and the incorporation of a wider range of somatic symptoms, review authors proposed also broadening our population inclusion criteria. Organizations participating in the Drug Effectiveness Review Project agreed to broaden the population criteria to allow inclusion of studies that based their

diagnosis of fibromyalgia on any other explicit criteria that would now fall under the umbrella of the 2010 definition. Thus, with this expansion of population criteria, inclusion of additional studies of primarily older drugs was permitted (e.g., amitriptyline and cyclobenzaprine), with planned sensitivity analyses to investigate whether variation in diagnostic criteria contributed to differences in outcomes.

The other major change to the planned scope of the review related to the list of included drugs. Originally, the review included the following list of additional drugs: benzodiazepines, dopamine agonists, serotonin receptor antagonists, growth hormone, nonsteroidal antiinflammatory drugs, opioid analgesics, opioid receptor antagonists, sedative hypnotics, selective estrogen receptor modulators, other skeletal muscle relaxants, and synthetic cannabinoids. However, after the review was underway, the organizations participating in the Drug Effectiveness Review Project decided to eliminate the additional drugs listed above as they lacked relevance to their specific programmatic interests. Of note is that, among the eliminated drugs, although opioids may currently still be used in clinical practice for treating fibromyalgia, at least in our initial searching we did not identify any randomized controlled trials of their use in patients with fibromyalgia. Among the eliminated drugs, we only found randomized controlled trials for growth hormone,[14, 15] moclobemide,[16, 17] nabilone,[18, 19] naltrexone,[20] raloxifen,[21] ritanserin,[22] sodium oxybate,[23, 24] terguride,[25] tramadol,[26-28] and zolpidem.[29]

The following Key Questions inclusion criteria reflect the aforementioned revisions and were approved by the organizations participating in the Drug Effectiveness Review Project in December 2010 to guide the review for this report:

1. For adults with fibromyalgia, what is the comparative effectiveness/efficacy of included interventions?
    a. When used as monotherapy?
    b. When used as adjunctive therapy?
2. For adults with fibromyalgia, what are the comparative harms of included interventions?
    a. When used as monotherapy?
    b. When used as adjunctive therapy?
3. Are there subgroups of patients based on demographics (age, racial or ethnic groups, and gender), socioeconomic status, other medications, or comorbidities for which any included drugs are more effective or associated with fewer harms?
    a. When used as monotherapy?
    b. When used as adjunctive therapy?

METHODS

Inclusion Criteria

Populations

Included were adult outpatient populations with fibromyalgia or fibromyalgia syndrome as diagnosed by the 1990 or 2010 American College of Rheumatology diagnostic criteria for fibromyalgia.[2, 30] Studies of patients with fibromyalgia, fibromyalgia syndrome, or fibrositis based on diagnostic criteria other than those established by American College of Rheumatology

(1990 or 2010 versions) were also included, with planned sensitivity analyses to investigate whether variation in diagnostic criteria contributed to differences in outcomes.

Interventions

Table 1 below lists the interventions that are included in this report. Black box warnings for the included interventions are listed in Appendix C.

Table 1. Included interventions

Generic name	Trade name	Approved for treatment of fibromyalgia
Tricyclic antidepressants		
Amitriptyline	Elavil® [a] *Generic only*	
Desipramine	Norpramin® [b]	
Imipramine	Tofranil®, Tofranil-PM®, Impril [a]	
Nortriptyline	Aventyl®, Pamelor®	
Serotonin norepinephrine reuptake inhibitors		
Desvenlafaxine	Pristiq®	
Venlafaxine	Effexor®, Effexor XR®	
Selective serotonin reuptake inhibitors		
Citalopram	Celexa®	
Escitalopram	Lexapro®, Cipralex® [c]	
Fluoxetine	Prozac®, Prozac weekly™ [b]	
Fluvoxamine	Luvox®, Luvox CR® [b]	
Paroxetine	Paxil®, Paxil CR®, Pexeva®	
Sertraline	Zoloft®	
Selective serotonin and norepinephrine reuptake inhibitors		
Duloxetine	Cymbalta®	X
Milnacipran	Savella®	X
Noradrenergic and specific serotonergic reuptake inhibitor		
Mirtazapine	Remeron®, Remeron Soltab®	
Norepinephrine and dopamine reuptake inhibitor		
Bupropion	Wellbutrin® [b], Wellbutrin SR®, Wellbutrin XL®	
Serotonin receptor antagonist		
Nefazodone [b]	*Generic only* [b]	
Antiepileptic drugs		
Carbamazepine	Tegretol®, Tegretol XR®, Carbatrol®, Equetro®, Mazepine [a]	
Divalproex [d]	Depakote® [b], Depakote ER® [b], Epival® [a]	
Ethotoin [b]	Peganone® [b]	
Gabapentin	Neurontin®	
Lacosamide	Vimpat®	
Lamotrigine	Lamictal®, Lamictal ODT®,	

Generic name	Trade name	Approved for treatment of fibromyalgia
	Lamictal XR®, Lamictal CD®	
Levetiracetam	Keppra®, Keppra XR™	
Oxcarbazepine	Trileptal®	
Phenytoin	Dilantin®	
Pregabalin	Lyrica®	X
Tiagabine [b]	Gabitril® [b]	
Topiramate	Topamax®	
Valproic acid [d]	Depakene®, Depacon® [b] Stavzor®	
Zonisamide [b]	Zonegran® [b]	
Skeletal muscle relaxants		
Cyclobenzaprine	Amrix®, Flexeril®	

Abbreviations: CD, chewable dispersible; CR, controlled-release; ER, extended-release; HP, high potency; ODT, orally disintegrating tablet; PM, pamoate; SR, sustained-release; XL, extended-release; XR, extended-release.
[a] Available in Canada, not available in the United States.
[b] Available in the United States, not available in Canada.
[c] Canadian trade name.
[d] Also known as valproate.

Comparators

- Direct comparisons of included drugs in head-to-head trials were preferred
- For indirect comparisons, only placebo-controlled trials were considered.

Effectiveness/Efficacy Outcomes

- Pain – primary outcome, including tender points, as based on all types of assessments and at all time points
- Functional capacity (e.g., work productivity, days missed from work, etc.)
- Health-related quality of life
- Response (e.g., proportion achieving, speed of, duration of, etc.)
- Fatigue, depressiveness, sleep, global status.

Harms

- Overall adverse events
- Withdrawals due to adverse events
- Specific adverse events (e.g., hepatic, renal, hematologic, dermatologic, sedation/drowsiness, and other neurologic side effects).

Study Designs

1. For effectiveness, controlled clinical trials and good-quality systematic reviews
2. For harms, in addition to controlled clinical trials, observational studies were included

a. Observational studies were defined as comparative cohort and case-control studies with a well defined fibromyalgia population
b. Noncomparative observational studies were included only if the duration of follow-up was 1 year or longer, and if serious harms were reported. A serious harm is one that results in long-term health effects or mortality.

Literature Search

We searched Ovid MEDLINE® (1947 to September Week 3 2010), the Cochrane Database of Systematic Reviews® (2005 to September 2010), and the Cochrane Central Register of Controlled Trials® (3rd Quarter 2010) and Database of Abstracts of reviews of Effects (3rd Quarter 2010) using included drugs, indications, and study designs as search terms. (See Appendix D for complete search strategies). We attempted to identify additional studies through hand searches of reference lists of included studies and reviews. In addition, we searched the US Food and Drug Administration Center for Drug Evaluation and Research website for medical and statistical reviews of individual drug products. Finally, we requested dossiers of published and unpublished information from the relevant pharmaceutical companies for this review. All received dossiers were screened for studies or data not found through other searches. All citations were imported into an electronic database (Endnote® X2, Thomson Reuters).

Study Selection

Selection of included studies was based on the inclusion criteria created by the Drug Effectiveness Review Project participants, as described above. Titles and abstracts were first assessed by one reviewer for inclusion using the criteria described above and then checked by a second reviewer. Full-text articles of potentially relevant citations were retrieved and again were assessed for inclusion by one reviewer and checked by a second reviewer. Disagreements were resolved by consensus. Results published *only* in abstract form were not included because inadequate details were available for quality assessment.

Data Abstraction

The following data were abstracted from included trials: eligibility criteria; interventions (dose and duration); population characteristics, including sex, age, ethnicity, and diagnosis; numbers randomized, withdrawn, lost to follow-up and analyzed; and results for each included outcome. We recorded intention-to-treat results when reported. If true intention-to-treat results were not reported, but loss to follow-up was very small, we considered these results to be intention-to-treat results. In cases where only per protocol results were reported, we calculated intention-to-treat results if the data for these calculations were available. Data abstraction was performed by one reviewer and was independently checked by a second reviewer. Differences were resolved by consensus.

Validity Assessment

We assessed the internal validity (quality) of trials based on the predefined criteria (see www.ohsu.edu/drugeffectiveness). These criteria are based on the US Preventive Services Task

Force and the National Health Service Centre for Reviews and Dissemination (United Kingdom) criteria.[31, 32] We rated the internal validity of each trial based on the methods used for randomization, allocation concealment, and blinding; the similarity of compared groups at baseline; maintenance of comparable groups; adequate reporting of dropouts, attrition, crossover, adherence, and contamination; loss to follow-up; and the use of intention-to-treat analysis. Trials that had a fatal flaw were rated poor quality; trials that met all criteria were rated good quality; the remainder were rated fair quality. As the fair-quality category is broad, studies with this rating vary in their strengths and weaknesses: the results of some fair-quality studies are *likely* to be valid, while others are only *possibly* valid. A poor-quality trial is not valid; the results are at least as likely to reflect flaws in the study design as a true difference between the compared drugs. A fatal flaw is reflected by failure to meet combinations of items of the quality assessment checklist. A particular randomized trial might receive 2 different ratings, one for effectiveness and another for adverse events.

The criteria used to rate observational studies of adverse events reflect aspects of the study design that are particularly important for assessing adverse event rates. We rated observational studies as good quality for adverse event assessment if they adequately met 6 or more of the 7 predefined criteria, fair quality if they met 3 to 5 criteria, and poor quality if they met 2 or fewer criteria.

Included systematic reviews were also rated for quality. We rated the internal validity based on a clear statement of the questions(s); reporting of inclusion criteria; methods used for identifying literature (the search strategy), validity assessment, and synthesis of evidence; and details provided about included studies. Again, these studies were categorized as good when all criteria were met.

Quality assessment was performed by one reviewer and independently checked by a second reviewer and differences were resolved by consensus.

Grading the Strength of Evidence

We graded strength of evidence based on the guidance established for the Evidence-based Practice Center Program of the Agency for Healthcare Research and Quality.[33] Developed to grade the overall strength of a body of evidence, this approach incorporates 4 key domains: risk of bias (includes study design and aggregate quality), consistency, directness, and precision of the evidence. It also considers other optional domains that may be relevant for some scenarios, such as a dose-response association, plausible confounding that would decrease the observed effect, strength of association (magnitude of effect), and publication bias.

Table 2 describes the grades of evidence that can be assigned. Grades reflect the strength of the body of evidence to answer key questions on the comparative effectiveness, efficacy, and harms of drugs for fibromyalgia. Grades do not refer to the general efficacy or effectiveness of pharmaceuticals. Grading the strength of the evidence was first performed by one reviewer and independently checked by a second reviewer and differences were resolved by consensus.

Among the multitude of outcomes assessed in trials of drugs for fibromyalgia, we focused on rating the strength of evidence for only a subset of 6 that we judged to represent the most clinically important and reliable: pain, fatigue, proportion of patients with a 50% or greater improvement in symptoms, mean change in Fibromyalgia Impact Questionnaire Total Score, overall adverse events, and withdrawals due to adverse events.

Table 2. Definitions of the grades of overall strength of evidence[34]

Grade	Definition
High	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate.
Low	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate.
Insufficient	Evidence either is unavailable or does not permit estimation of an effect.

Data Synthesis

We constructed evidence tables showing the study characteristics, quality ratings, and results for all included studies. We reviewed studies using a hierarchy of evidence approach, where the best evidence is the focus of our synthesis for each question, population, intervention, and outcome addressed. Studies that evaluated one drug for fibromyalgia against another provided direct evidence of comparative effectiveness and adverse event rates. Where possible, these data were the primary focus. Direct comparisons were preferred over indirect comparisons; similarly, effectiveness and long-term safety outcomes were preferred to efficacy and short-term tolerability outcomes.

In theory, trials that compare an included drug for fibromyalgia with any other nonincluded treatment or with placebos can also provide evidence about effectiveness. This is known as an indirect comparison and can be difficult to interpret for a number of reasons, primarily heterogeneity of trial populations, interventions, and outcomes assessment. Data from indirect comparisons are used to support direct comparisons, where they exist, and are used as the primary comparison where no direct comparisons exist. Indirect comparisons should be interpreted with caution.

Meta-analyses were conducted to summarize data and obtain more precise estimates on outcomes for which studies were homogeneous enough to provide a meaningful combined estimate. In order to determine whether meta-analysis could be meaningfully performed, we considered the quality of the studies and the heterogeneity among studies in design, patient population, interventions, and outcomes. When meta-analysis could not be preformed, the data were summarized qualitatively.

For continuous outcomes, we used the mean difference between treatment and placebo groups as the effect measure, which we estimated based on mean change scores and standard errors from baseline to follow up for each group from each study. Hedge's *g*, one of the measures for standardized mean differences, was used if different instruments (scales) were used by different studies for the same outcome. For dichotomous outcomes, relative risk was used as the effect measure. All combined effects were estimated using random-effects models.[35] The Q statistic and the I^2 statistic (the proportion of variation in study estimates due to heterogeneity) were calculated to assess heterogeneity in effects between studies.[36, 37] Due to the small number of studies, it was not feasible to use subgroup analysis and meta-regression to explore heterogeneity. We conducted sensitivity analyses to check the impact of dosage, length of follow-up, and definitions of outcome on the results.

Because head-to-head evidence was sparse, we used the method described by Bucher, et al. to perform indirect comparison meta-analysis to evaluate the difference between drugs based

on data from placebo-controlled trials, as the trials were generally comparable in patient population and clinical and methodological characteristics. The magnitude of difference was characterized using relative risk ratio for relative risks and difference of mean difference for mean differences. Negative (−) difference of mean differences were interpreted as suggesting that drug A is associated with a greater reduction in fibromyalgia symptoms than drug B. Relative risk ratios greater than 1.0 were interpreted as suggesting that drug A is associated with a higher relative benefit compared to drug B for efficacy outcomes and higher relative risk for adverse events. All analyses were performed using Stata 11.0 (StataCorp, College Station, TX, 2009).

Peer Review

We requested and received peer review of the report from 4 content experts. Their comments were reviewed and, where possible, incorporated into the final document. All comments and the authors' proposed actions were reviewed by representatives of the participating organizations of the Drug Effectiveness Review Project before finalization of the report. Names of peer reviewers for the Drug Effectiveness Review Project are listed at www.ohsu.edu/drugeffectiveness.

Public Comment

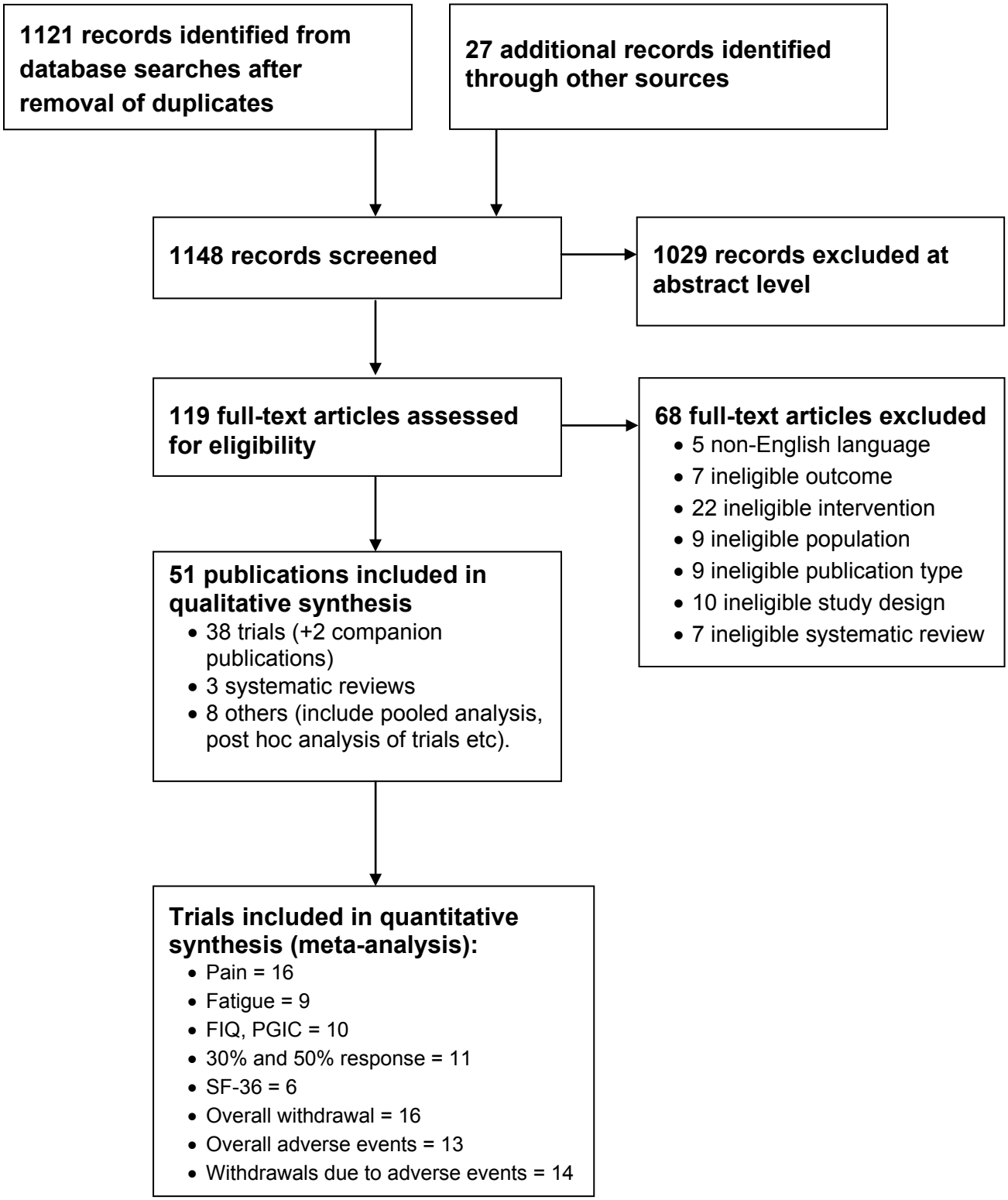
This report was posted to the Drug Effectiveness Review Project website for public comment. We received comments from 3 persons, representing 2 pharmaceutical companies.

RESULTS

Overview

Literature searches identified 1148 citations. We received dossiers from 2 pharmaceutical manufacturers: Eli Lilly and Company and Forest Laboratories Inc. By applying the eligibility and exclusion criteria to titles and abstracts of all identified citations, we obtained full-text copies of 119 citations. After re-applying the criteria for inclusion, we ultimately included 51publications. Five included studies were identified after expanding the population inclusion criteria to include a broadened definition of fibromyalgia or fibrositis.[38-42] See Appendix E for a list of excluded studies and reasons for exclusion at this stage. Figure 1 shows the flow of study selection.

Figure 1. Results of literature search[a]



[a] The Drug Effectiveness Review Project uses a modified PRISMA flow diagram.[1]

Key Question 1. For adults with fibromyalgia, what is the comparative effectiveness/efficacy of included interventions?

1a. When used as monotherapy?

1b. When used as adjunctive therapy?

Summary of Findings

General

- We found no eligible studies of treatment for fibromyalgia with desipramine, imipramine, desvenlafaxine, venlafaxine, escitalopram, fluvoxamine, sertraline, mirtazapine, bupropion, nefazodone, carbamazepine, divalproex, ethotoin, lacosamide, lamotrigine, levetiracetam, oxcarbazepine, phenytoin, tiagabine, topiramate, valproic acid, or zonisamide
- We found no eligible studies of included interventions when used as adjunctive therapy.

Direct evidence

- There was low-strength evidence that immediate-release paroxetine is superior to amitriptyline in reducing pain (−28% compared with −1%; z= −5.64; $P<0.001$) and sleep disturbance (−39% compared with −13%; z= −4.62; $P<0.001$) based on an 100-mm visual analog scale, but there were no differences in fatigue, tender points, depression, or in the numbers of patients with moderate or marked improvement based on clinical global assessment
- There was low-strength evidence of no significant differences between amitriptyline and cyclobenzaprine or nortriptyline in any efficacy outcomes.

Indirect evidence

Pooled analysis

- All trials included used the drugs as monotherapy and no trial evaluated the effectiveness of the drugs as adjunctive therapy
- Pain, 50% response rate, Patient Global Impression of Change: Pooled analysis of placebo-controlled trials of amitriptyline, pregabalin, milnacipran, and duloxetine found that all drugs were superior to placebo
- Fatigue: Pooled analysis of placebo-controlled trials of amitriptyline, pregabalin, and milnacipran found that these drugs were superior to placebo for short-term results, but not in longer-term trials of 24 to 28 weeks in duration.

Indirect meta-analysis

- Pain: Indirect meta-analysis of short-term trials (8-15 weeks) of amitriptyline, pregabalin, milnacipran, and duloxetine on measures of pain found that duloxetine was superior to milnacipran (mean difference, −0.45; 95% CI, −0.80 to −0.08); no differences existed between the other drugs and no differences existed between the drugs at 28 weeks
- Response rate: Indirect meta-analysis of pregabalin, milnacipran, and duloxetine on response rate of 50% or 30% improvement in pain found all drugs superior to placebo

with no difference between the drugs; there was insufficient evidence to report on this measure for amitriptyline.
- Function: Indirect meta-analysis found no difference between of amitriptyline, pregabalin, milnacipran, and duloxetine on function as measured by the Fibromyalgia Impact Questionnaire and there was no difference between duloxetine, milnacipran, or pregabalin on the physical and mental components of the Medical Outcomes Study 36-item Short-Form Health Survey
- Duloxetine and pregabalin were superior to milnacipran on measures of sleep disturbance and health-related quality of life; duloxetine and milnacipran were superior to pregabalin on improvement in depressed mood.

Comparisons to placebo

- Gabapentin significantly improved pain severity and response, overall impact of fibromyalgia, global status, and sleep, but not tender point pain threshold, depression or overall quality of life
- Compared with placebo, a significant reduction in pain severity was only found with cyclobenzaprine in 1 of 3 trials
- Among selective serotonin reuptake inhibitors, only fluoxetine, at a higher dose (45 mg), resulted in significantly greater improvements than placebo in pain, fatigue, and Fibromyalgia Impact Questionnaire Total Score
- Controlled-release paroxetine did not significantly decrease pain, disability, or depressiveness or increase the number of patients with a 50% or greater response, but did significantly decrease the Fibromyalgia Impact Questionnaire Total Score, fatigue, and improved global status
- Citalopram did not significantly improve pain or fatigue and only reduced depression and improved sleep in 1 of 2 trials.

Detailed Assessment

Direct evidence

Direct evidence regarding the comparative effectiveness among included interventions was limited and only available from 4 small randomized controlled trials that compared amitriptyline to cyclobenzaprine (N=208),[43] fluoxetine (N=31),[44] nortriptyline (N=118),[45] and immediate-release paroxetine (N=68).[46] All patients met the American College of Rheumatology 1990 criteria for classification of fibromyalgia. Three trials reported duration of fibromyalgia, which ranged from 36 months[46] to 101 months.[43] Participants were 95% female with mean ages ranging from 36 years[46] to 53.4 years.[45] Race was 100% Caucasian in the fluoxetine trial conducted in Massachusetts,[44] 62% Caucasian and 38% non-Caucasian in the Brazilian trial of nortriptyline,[45] and was not reported in the Canadian study of cyclobenzaprine[43] or the Turkish study of immediate-release paroxetine.[46] Trial settings included outpatient rheumatology clinics[43, 45] and a tertiary referral center,[44] but was not well described in the Turkish study.[46] Trial durations ranged from 6 weeks[44, 46] to 6 months.[43] Amitriptyline dosage ranged from 25 mg[44, 45] to 100 mg.[46] Mean dosages for the comparator drugs were 20 mg for cyclobenzaprine, 20 mg for fluoxetine, 20 mg for immediate-release paroxetine, and 25 mg for nortriptyline.

Three trials were rated fair quality[43, 45, 46] and the third was rated poor quality.[44] The main limitation of the poor-quality trial was that its analyses excluded a large proportion of the data –

over one-third; consequently, its results will not be discussed here, but can be found in Evidence Tables 1 and 2.[44]

Only 1[46] of 3 trials[43, 45, 46] found any significant differences between drugs in efficacy, in which there was low-strength evidence that immediate-release paroxetine 20 mg was superior to amitriptyline 100 mg over 6 weeks in reducing pain (−28% compared with −1%; z= −5.64; $P<0.001$) and sleep problems (−39% compared with −13%; z= −4.62; $P<0.001$), as measured based on a 100-mm visual analog scale.[46] However, differences between immediate-release paroxetine and amitriptyline were not significant for change in fatigue (−9% compared with −5%; z=0.74), tender points, depression, or in the numbers of patients with moderate or marked improvement based on clinical global assessment (22% compared with 17%; relative risk, 1.27; 95% CI, 0.47 to 3.47).

Otherwise, the remaining 2 trials provided low-strength evidence of no significant differences between amitriptyline and cyclobenzaprine[43] or nortriptyline[45] in any efficacy outcomes. The 2 head-to-head trials were inconsistent in their methods for assessing all efficacy outcomes. Compared with amitriptyline, similar reductions were found for cyclobenzaprine in visual analog scores for pain (−33% compared with −28%) and fatigue (−33% compared with −32%), McGill Pain Questionnaire Pain Rating Index scores (−32% compared with −31%), Depression Scale scores from the Arthritis Impact Measurement Scale (−25% compared with −20%), and in the Health Assessment Questionnaire Disability Index (−15% for both drugs). Similar proportions of patients were classified as responders based on meeting at least 4 of the following 6 criteria: 50% improvement in pain, sleep, fatigue, patient global assessment, or physician global assessment, and increase of 1 kg in mean total myalgic score (33% compared with 36%).[43] Also compared with amitriptyline, similar reductions were found for nortriptyline in mean number of tender points (−3 compared with −2.2) and in the Fibromyalgia Impact Questionnaire total score (27% compared with 37%), and similar proportions of patients met criteria for slight to great improvement based on the Verbal Evaluation Scale for Global Improvement (72% compared with 86%).[45]

We found 5 systematic reviews that assessed multiple drugs for the treatment of fibromyalgia.[47-51] Four reviews were graded as good quality[47-50] and 1 was graded as fair quality.[47] Only 1 of these reviews performed an indirect meta-analysis to compare the effectiveness between duloxetine, milnacipran, and pregabalin in fibromyalgia from placebo-controlled trials of individual drugs.[49] Although it was a recent review, we chose to update the indirect meta-analysis to include evidence from 2 new trials for milnacipran that are now available.[52, 53] We have reported results from the Hauser, et al. analysis where there was no new data or the new data did not change the results.

The remaining reviews had limited usefulness for this report as none provided any additional direct or indirect comparative data. One systematic review reported evidence on antidepressant drugs used for the treatment of fibromyalgia.[51] No meta-analysis was performed and the only individual drug data reported was the means across studies for amitriptyline. We conducted a meta-analysis comparing amitriptyline to other drugs to treat fibromyalgia so did not use the data reported in this review.[51] One systematic review reported a calculated effect size for amitriptyline and tramadol but only a class effect for antidepressants.[47] Given the new criteria for diagnosis of fibromyalgia released in 2010, we extended our inclusion criteria and included earlier studies of fibromyalgia not included in their review so did not report on their outcomes. One systematic review of amitriptyline reported on 10 randomized controlled trials but did not perform a meta-analysis due to large clinical variability and statistical heterogeneity.[50] The

results are considered as appropriate below. One systematic review pooled data of pregabalin and gabapentin and reported on a class effect size rather than comparative data which limited its usefulness for this report.[48]

Indirect evidence

Indirect meta-analysis

Of the 8 drugs included in this review, we found that only 4 drugs had multiple trials with sample sizes adequate to perform a comparative analysis. All trials used the drugs as monotherapy and no trial evaluated the effectiveness of the drugs as adjunctive therapy. We performed a meta-analysis of 6 placebo-controlled trials of amitriptyline,[39, 43, 45, 54-56] 4 placebo-controlled trials of pregabalin,[57-60] 5 placebo-controlled trials of milnacipran,[52, 53, 61-64] and 4 placebo-controlled trials of duloxetine.[65-69] Two trials of amitriptyline were identified by expanding our definition of fibromyalgia to include criteria for fibrositis that would fall under the umbrella of the updated definition of fibromyalia in 2010.[39, 70] Data from Scudds, et al. was excluded from the 8-15 week meta-analysis due to its short duration of 4 weeks. One additional trial of pregabalin was included but did not contribute data to our analysis due to significant methodological variance from the other trials and given that its outcome was loss of effectiveness in responders. The trial results are reported below.[71] No trials were of long duration (8 to 28 weeks). Sample size was small for all of the amitriptyline trials (N=22 to 126) but generally moderate to large for the rest (N=125 to 1025). All of the trials were rated as fair quality. The trials of duloxetine, milnacipran, and pregabalin were industry sponsored whereas only 1 of the amitriptyline trials were industry sponsored.[56] The baseline demographics for all of the trials were similar, with the majority of patients being middle aged (mean age range 40-53 years), white (65% to 93%), and female (82% to 100%). For pregabalin, we excluded data on the lowest dose of 150 mg once daily as only 1 trial included this dose, N=131, and a prior meta-analysis found this dose to be no different than placebo.[58, 72] Otherwise we combined data on different daily dosage groups including amitriptyline 25-50 mg, pregabalin 300-600 mg, milnacipran 100-200 mg, and duloxetine 40-120 mg. We reported on outcomes of clinical importance and with adequate data to perform the analysis. We performed our analysis using short-term data (8-15 weeks) as this was consistent across most studies. We performed a sensitivity analysis to determine if differences in outcome were seen using long-term trials. We excluded trials shorter than 8 weeks from our analysis.[42, 44, 73] The results and quality of these trials are summarized in Evidence Tables 1 and 2.

Pain

Multiple measures of pain were used across studies and at variable intervals. We performed our analysis based on a 10-point derivative scale as this was most consistently reported across trials and is considered a valid method of pain assessment. All drugs were effective in improving average pain or 24 hour recall of pain compared with placebo, with amitriptyline showing the most effectiveness (−1.68; 95% CI, −3.3 to −0.05), followed by duloxetine (−0.94; 95% CI, −1.3 to −0.61), pregabalin (−0.67; 95% CI, −0.94 to −0.40), and milnacipran (−0.49; 95% CI, −0.65 to −0.33) (Table 3). Significant heterogeneity was seen with the amitriptyline (I^2=72.7%) and pregabalin trials (I^2=62.6%). Given that the Ginsberg 1996 trial used a sustained-release formulation of amitriptyline, we repeated the analysis excluding the data from this trial and found that although the result was in the direction of improvement, the significance for amitriptyline was lost with a pooled mean difference of −0.94 (95% CI, −2.18 to 0.31),

I^2=58.3%. A similar effect was seen when we repeated the analysis excluding the data from Carette 1986 which did not use the American College of Rheumatology criteria for population inclusion, with a pooled mean difference of −1.82 (95% CI, −3.99 to 0.35), I^2=81%. Nishishinya, et al. performed a systematic review of amitriptyline and identified 10 placebo-controlled trials studying the effectiveness of amitriptyline in treating fibromyalgia.[50] They did not perform a pooled analysis due to clinical and statistical heterogeneity but did find significant improvement in pain for the trials of amitriptyline 25 mg compared with placebo, consistent with our analysis.[50]

Indirect meta-analysis of all placebo-controlled trials found that there was no difference between the drugs except that duloxetine was superior to milnacipran (difference in mean difference, −0.45; 95% CI, −0.80 to −0.08). This finding held true when the analysis was repeated, excluding the Ginsberg 1996 and the Carette 1986 trial data for amitriptyline. The recent meta-analysis performed by Hauser[49] was consistent with the finding that duloxetine was superior to milnacipran but unlike our analysis, they found that duloxetine and milnacipran were also superior to pregabalin (standardized mean difference, 1.22; 95% CI, 1.10 to 1.80 and standardized mean difference, 0.70; 95% CI, 0.58 to 0.82, respectively). One of the limitations of the trials was that there were multiple ways in which pain was reported in each trial, including daily and weekly symptoms as well as multiple different pain scales between trials. We analyzed average or 24-hour daily pain score, converting it to a 0-10 scale, and restricted our analysis to data 8-15 weeks in duration. It is unclear how Hauser combined the multiple reports of pain. Additionally, the median duration of the randomized phase of the trials from the Hauser meta-analysis was 24 (range 6-28) weeks.[49] When we repeated our analysis including all trial data (duration range 6-28 weeks), the superiority of duloxetine over milnacipran was lost (−0.32; 95% CI, −0.68 to 0.04). In summary, for short-term use, there was low evidence that all drugs are superior to placebo in pain response, with no difference between the drugs except that duloxetine was superior to milnacipran. Head-to-head trials are needed to confirm these findings.

Table 3. Pooled effectiveness of amitriptyline, pregabalin, milnacipran, and duloxetine compared with placebo (8-15 weeks)

Outcome measure	Amitriptyline	Pregabalin	Milnacipran	Duloxetine
Pain				
Mean difference in 24-hour daily pain score converted to a 0-10 scale Pooled mean difference (95% CI) vs. placebo	−1.68 (−3.3 to −0.05) [39, 43, 54, 55]	−0.67 (−0.94 to −0.40) [57-60]	−0.49 (−0.65 to −0.33) [52, 53, 61-64]	−0.94 (−1.3 to −0.61) [65, 66, 68, 69]
Mean difference in 24-hour daily pain score converted to a 0-10 scale Pooled mean difference (95% CI) vs. placebo (excluding Ginsberg 1996)[55]	−0.94 (−2.18 to 0.31) [39, 43, 54]	---	---	---

Outcome measure	Amitriptyline	Pregabalin	Milnacipran	Duloxetine
Mean difference in 24-hour daily pain score converted to a 0-10 scale Pooled mean difference (95% CI) vs. placebo (excluding Carette 1986)[39]	-1.82 (-3.99 to 0.35) [38, 50, 51]			
Response rate				
50% improvement in pain RR (95% CI) vs. placebo	18.76 (2.61 to 134.83) [54, 55]	1.81 (1.41 to 2.33) [57, 58, 60]	1.46 (1.23 to 1.74) [52, 62-64]	1.65 (1.32 to 2.06) [65, 66, 68, 69]
30% improvement in pain RR (95% CI) vs. placebo	NR	1.39 (1.09 to 1.77) [57-60]	1.37 (1.26 to 1.50) [52, 53, 61-64]	1.50 (1.25 to 1.80) [66, 69]
PGII or PGIC Any improvement RR (95% CI) vs. placebo	1.59 (1.13 to 2.20) [45]	1.33 (1.11 to 1.58) [57-60]	1.47 (1.22 to 1.67) [52, 53, 61-64]	---
Fatigue				
MFI, MAF, VAS 0-10 Pooled mean difference (95% CI) vs. placebo	−0.46 (−0.91 to −0.01) [43, 54, 55]	−2.27 (−4.39 to −0.15) [57, 58]	−1.91 (−2.87 to −0.95) [52, 53, 61, 63]	−0.09 (−0.25 to 0.06) [66, 69]
VAS 0−10 Pooled SMD (95% CI) vs. placebo (excluding Ginsberg, 1996)[55]	−0.23 (−0.55 to 0.08) [43, 54]	---	---	---
Fibromyalgia Impact Questionnaire				
FIQ 0-80 scale Pooled mean difference (95% CI) vs. placebo	−7.43 (−19.13 to 4.27) [45]	−3.26 (−5.82 to −0.70) [57, 59]	−3.61 (−5.03 to −2.18) [52, 53, 61, 63]	−6.24 (−8.52 to −3.96) [65, 66, 69]
SF-36 Physical Component				
SF-36 0-100 Pooled mean difference (95% CI) vs. placebo	---	0.41 (−0.83 to 1.65) [57]	1.08 (0.54 to 1.61) [52, 53, 61, 59]	1.84 (−0.20 to 3.88) [65, 69]
SF-36 Mental Component				
SF-36 0-100 Pooled mean difference (95% CI) vs. placebo	---	1.67 (−0.12 to 3.45) [57]	1.47 (0.74 to 2.19) [52, 53, 61, 59]	0.27(0.95 to 4.44) [65, 69]

Abbreviations: CI, confidence interval; FIQ, Fibromyalgia Impact Questionnaire total score; MAF, multidimensional assessment of fatigue; MFI, multidimensional fatigue inventory; PGIC, Patient Global Impression of Change; PGII, Patient Global Impression of Improvement; SF-36 MC, Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) mental summaries component; SF-36 PC, Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) physical summaries component; SMD, standardized mean difference.

Table 4. Indirect analysis of placebo-controlled trials in fibromyalgia

	Duloxetine vs. milnacipran	Duloxetine vs. pregabalin	Duloxetine vs. amitriptyline	Milnacipran vs. pregabalin	Milnacipran vs. amitriptyline	Pregabalin vs. amitriptyline
Pain						
Difference of mean differences in 24-hour daily pain score, 0-10 scale (95%CI)	−0.45 (−0.80, to −0.08)	−0.27 (−0.68 to 0.15)	0.74 (−0.92 to 2.41)	0.18 (−0.13 to 0.49)	1.19 (−0.45, to 2.83)	1.01 (−0.64 to 2.66)
Difference of mean differences in 24−hour daily pain score, 0-10 scale (95%CI) (excluding Ginsberg 1996)[55]	---	---	−0.01 (−1.29 to 1.29)	---	0.44 (−0.81 to 1.70)	0.26 (−1.01 to 1.54)
Difference of mean differences in 24−hour daily pain score, 0-10 scale (95%CI) (excluding Carette1986)[39]	---	---	0.88 (-1.3 to 0.49)	---	1.33 (-0.85 to 3.5)	1.15 (-1.04 to 3.33)
Response						
50% improvement in pain Ratio of RR (95% CI)	1.13 (0.85 to 1.50)	0.91 (0.65 to 1.28)	---	0.81 (0.59 to 1.10)	---	---
30% improvement in pain Ratio of RR (95% CI)	1.09 (0.89 to 1.33)	1.07 (0.79 to 1.45)	---	0.98 (0.76 to 1.27)	---	---
PGII or PGIC Any improvement Ratio of RR (95% CI)	---	---	---	1.07 (0.85 to 1.36)	0.90 (0.62 to 1.30)	0.84 (0.57 to 1.22)
Fatigue						
MFI, MAF, VAS 0-10 Difference of SMD (95% CI)	0.035 (−1.34 to 0.21)	0.13 (−0.12 to 0.38)	0.36 (−0.12 to 0.84)	0.09 (−0.11 to 0.30)	0.33 (−0.13 to 0.79)	0.24 (−0.26 to 0.73)

	Duloxetine vs. milnacipran	Duloxetine vs. pregabalin	Duloxetine vs. amitriptyline	Milnacipran vs. pregabalin	Milnacipran vs. amitriptyline	Pregabalin vs. amitriptyline
MFI, MAF, VAS 0-10 Difference of SMD (95% CI), (Excluding Ginsberg 1996)	---	---	0.14 (−0.22 to 0.49)	---	0.10 (−0.22 to 0.43)	0.01 (−0.37 to 0.38)
Function						
FIQ 0-80 scale Difference of mean difference (95% CI)	−2.63 (−5.32 to 0.06)	−2.97 (−6.40 to 0.46)	1.19 (−10.73 to 13.11)	−0.34 (−3.27 to 2.59)	3.82 (−7.96 to 15.61)	4.17 (−7.81 to 16.14)
SF-36 MC Difference of mean difference (95% CI)	1.23 (-0.66 to 3.11)	1.03 (-1.46 to 3.52)	---	-0.20 (-2.12 to 1.72)	---	---
SF-36 PC Difference of mean difference (95% CI)	0.77 (-1.34 to 2.89)	1.43 (−0.96 to 3.82)	---	0.66 (-0.69 to 2.02)	---	---

Abbreviations: CI, confidence interval; FIQ, Fibromyalgia Impact Questionnaire total score; MAF, multidimensional assessment of fatigue; MFI, multidimensional fatigue inventory; PGIC, Patient Global Impression of Change; PGII, Patient Global Impression of Improvement; SF-36 MC, Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) mental summaries component; SF-36 PC, Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) physical summaries component; SMD, standardized mean difference; VAS, visual analogue scale.

Response

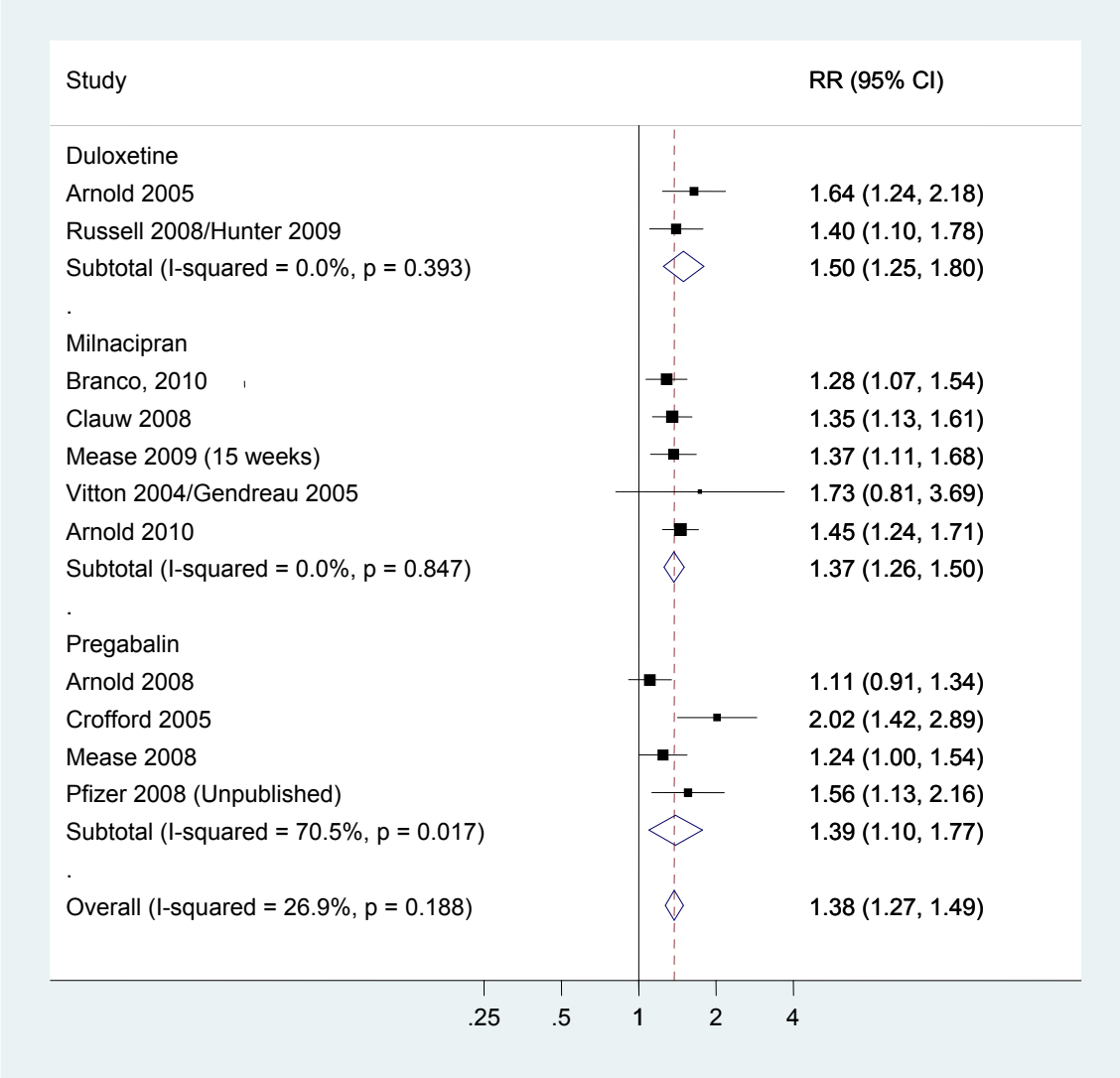
Response was defined differently in all of the trials, with many of the trials having a composite response that included multiple outcome measures such as 30% or 50% reduction in pain and improvement on Patient Global Impression of Improvement or Change. Many of the trials also reported this data separately, allowing us to perform a pooled analysis and indirect meta-analysis on pain response rate. Pooled analysis of placebo-controlled trials found that all drugs were effective in achieving a 50% improvement in pain (Table 4, Figure 2). The data for amitriptyline was insufficient to determine its validity given the small number of patients from 2 trials, N=68, one of which used a sustained release formulation of amitriptyline.[54, 55] Two systematic reviews of pregabalin compared with placebo in fibromyalgia used the same 4 trials, pooled data based on dose, and found similar results on 30% or 50% pain response and for "much or very much improved" on Patient Global Impression of Improvement or Change.[72, 74] Indirect meta-analysis of the placebo-controlled trials of duloxetine, milnacipran, and pregabalin found that there was no significant difference between the drugs on ability to achieve a 50% reduction in pain (Table 4). The data for amitriptyline was too sparse for indirect comparison. We had no data on amitriptyline for the outcome of 30% improvement in pain but all of the other drugs had a small but significant improvement in 30% pain response compared with placebo and no significant difference between the drugs was found (Tables 3 and 4). This finding was consistent with the Hauser, et al. analysis.[49]

For the outcome of any improvement on the Patient Global Impression of Improvement or Change score, a significant benefit over placebo was found for milnacipran, pregabalin, and amitriptyline. However, I^2 statistics revealed substantial heterogeneity within the milnacipran (74.4%) and pregabalin (73.6%) trials. When we used meta-regression to explore reasons for the heterogeneity, we found a significant association for placebo group response rate (P=0.008), but not for definition of improvement (any compared with much or very much). Findings from the meta-regression indicated that trials with higher rates of improvement in the placebo group had smaller benefits with milnacipran or pregabalin. In our indirect meta-analysis, no significant differences were found between milnacipran, pregabalin, and amitriptyline. However, the data on amitriptyline was insufficient to make any conclusions given that there was only 1 small trial that reported on this outcome (N=80). When we repeated our analysis considering only those who reported much or very much improvement on the Patient Global Impression of Improvement or Change score, no difference was found between milnacipran and pregabalin (ratio of relative risk, 0.85; 95% CI, 0.59 to 1.24) (Table 4). Trials of duloxetine that reported on this outcome provided only the mean change rather than the absolute number of patients reporting improvement and thus could not be included in our analysis. Arnold, et al. performed a pooled analysis of the same 4 duloxetine trials and reported a significant difference between duloxetine and placebo with 38.4% of duloxetine-treated patients reporting at least much improved compared with 21.7% of placebo patients, P<0.001.[75]

One additional trial of pregabalin was reviewed but not included in our analysis due to significant methodological variance from the other trials.[71] In a 26-week placebo-controlled randomized discontinuation trial of patients with fibromyalgia who had achieved at least 50% reduction on the visual analogue scale and much or very much improvement on the Patient Global Impression of Improvement or Change score after a 6-week open-label treatment period (n=566), the time to loss of response (<30% reduction in pain) was longer for pregabalin than for placebo (34 days compared with 7 days; P<0.0001).[71]

In summary, all drugs were superior to placebo in 50% response rate and Patient Global Impression of Improvement or Change score. There was low evidence that no differences exist between pregabalin, duloxetine, or milnacipran on pain response rate with insufficient evidence to report on this outcome for amitriptyline. There was insufficient evidence to draw conclusions of the comparative effectiveness on the Patient Global Impression of Improvement or Change score.

Figure 2. Response rate 50% improvement in pain



Fatigue

Fatigue was defined differently between the drugs. We considered the general fatigue score of the multidimensional fatigue inventory, the fatigue score of the Fibromyalgia Impact Questionnaire, the Visual Analogue Scale, and the global fatigue index of the Multidimensional Assessment of Fatigue score, and determined a standardized mean difference based on available short-term data (8-15 weeks). Milnacipran, pregabalin, and amitriptyline were superior to placebo in short-term trials of 8-15 weeks, but not in longer-term trials of 24-28 weeks. There was high heterogeneity noted in the pregabalin trials (I^2=62.8%). In their meta-analysis of the same 4 pregabalin trials, Straube, et al. found pregabalin superior to placebo at doses of 300 mg (P<0.01) and 450 mg (P<0.01), but not at 600 mg.[74] There was no difference between duloxetine and placebo, with 2 trials reporting on this outcome (N=727). No difference was found between the drugs (Table 3). Our result contradicts the meta-analysis of Hauser, et al. that found

milnacipran superior to duloxetine (standardized mean difference, 0.77; 95% CI, 0.67 to 0.87), and pregabalin superior to duloxetine (standardized mean difference, 0.62; 95% CI, 0.52 to 0.72) and milnacipran (standardized mean difference, 0.81; 95% CI, 0.71 to 0.91).[49] The difference between our findings and those of Hauser, et al. could be mainly due to the difference in indirect meta-analytic methods. We calculated an absolute difference of standardized mean differences between interventions whereas the Hauser analysis calculated a ratio of standardized mean differences between the drugs, which produces an estimate of the relative effect rather than an absolute difference. The ratio for standardized mean differences is rarely used. When we calculated a ratio of standardized mean differences, we could not replicate the significant value reported by Hauser, et al. We obtained 0.73 (95% CI, 0.13 to 4.19), which indicated that the difference is not significant. Given that the difference is actually small, an analysis based on absolute difference is not significant. In summary, there was low evidence that milnacipran, pregabalin, and amitriptyline are superior to placebo on measures of fatigue, and no differences existed between the drugs.

Function

The Fibromyalgia Impact Questionnaire total score was used to assess overall change in fibromyalgia symptoms and their impact on daily function. It is an instrument designed to reflect the multidimensionality of fibromyalgia by questioning patients about the extent of their symptoms and the effect of these on their activities of daily living.[76] On the measure of total score on the Fibromyalgia Impact Questionnaire, mean change from baseline reached the minimally clinically important difference level proposed by Bennett, et al. in all treatment and placebo groups.[77] Pooled analysis of short-term (8-15 weeks) placebo-controlled trials of duloxetine, milnacipran, and pregabalin found a significant improvement compared with placebo for all drugs (I^2=23.1%) (Table 3). Our results for pregabalin were in agreement with the pooled analysis by Straube, et al. who found pregabalin superior to placebo at all 3 doses (300 mg [P<0.05], 450 mg [P<0.001], and 600 mg [P<0.01]).[74] Indirect meta-analysis of placebo-controlled trials of duloxetine, milnacipran, and pregabalin found no significant difference between the drugs (Table 4). The data on amitriptyline was insufficient to make a statement on this outcome as only 1 small trial, N=80, reported data on Fibromyalgia Impact Questionnaire total score, although it did not find a significant difference compared with placebo.[45]

We also considered the Medical Outcomes Study 36-item Short-Form Health Survey physical and mental component summaries to assess therapeutic response of the included drugs on overall physical and mental function. No differences were found in mean differences between duloxetine, milnacipran or pregabalin on either of these measures based on 6 trials that reported data.[52, 53, 57, 61, 65, 68, 69] The pooled analysis of placebo-controlled trials (8-15 weeks) of milnacipran and duloxetine found a small but significant mean difference compared to placebo for the Medical Outcomes Study 36-item Short-Form mental component summary (milnacipran 1.47; 95% CI, 0.74 to 2.19 and duloxetine 0.27; 95% CI, 0.95 to 4.44), whereas only milnacipran found a significant mean difference compared to placebo on the physical component summary (1.08; 95% CI, 0.54 to 1.61)[52, 53, 61, 63] When longer-term data (6-28 weeks) were added to the analysis, milnacipran was still found to be significantly superior to placebo on the physical component summary score (standard mean difference, 1.06; 95% CI, 0.52 to 1.60) and no significant benefits were found for duloxetine or pregabalin. In the pooled analysis of pregabalin by Straube, et al., the authors found a significant improvement in the mental component score for both the 450 mg (P<0.01) and the 600 mg (P<0.05) doses of pregabalin.[74] Combining the data

for all 3 doses may explain the difference in our results. It is important to recognize that although significant, the absolute mean differences noted between the active drug and placebo ranged between 1.0 to 1.5 on a scale of 0-100, which is small and questions the clinical significance of this change. This may explain why the significant difference noted with milnacipran and duloxetine compared to placebo did not translate into any difference between the drugs in the meta-analysis. Our results differed from the analysis by Arnold on the 4 placebo-controlled trials of duloxetine which found a statistically significant difference on the physical component score as well (1.08; 95% CI, 0.12 to 2.03; $P<0.05$). In reviewing the Arnold, et al. pooled analysis of the 4 placebo-controlled duloxetine trials,[65-67, 69] it appeared that the authors used unpublished individual patient data which was not provided in their report. This study was given a poor quality rating because of its risk of selection bias due to the failure to perform and report on a systematic search process, not reporting trial details or outcome data of individual trials, and failure to grade the quality of their results.[57] No trial of amitriptyline reported on this outcome.

In summary, there was low to insufficient evidence that all drugs are superior to placebo on the total score of the Fibromyalgia Impact Questionnaire with no difference between drugs. Milnacipran was found to have a small but significant improvement on the Medical Outcomes Study 36-item Short-Form Health Survey physical and mental component compared with placebo and duloxetine was found to have a small but significant improvement on the mental component summary. No difference was found between duloxetine, milnacipran, or pregabalin on this measure.

Other outcomes

Given the significant variability and sparsity of reporting data on outcomes of sleep disturbance, health-related quality of life, and depressed mood in the amitriptyline trials, we did not analyze these outcomes. Hauser, et al. performed a meta-analysis of placebo-controlled trials of duloxetine, milnacipran, and pregabalin, and did perform a pooled analysis and comparative meta-analysis for these outcomes.[49] Since their report was released, 2 additional milnacipran trials have been published,[52, 53] with only 1 providing additional evidence on the outcome of sleep.[53] These new results are consistent with prior studies. Hauser, et al. found that duloxetine was significantly superior to milnacipran on all 3 outcomes but was superior to pregabalin only on the outcome of improvement in depressed mood (Table 5). They found that pregabalin was superior to milnacipran on improvement in sleep disturbance and health-related quality of life, whereas milnacipran was superior to pregabalin on improvement in depressed mood (Table 5).

Table 5. Indirect analysis of placebo-controlled trials of pregabalin, milnacipran, and duloxetine for fibromyalgia[49]

	Duloxetine vs. milnacipran	Duloxetine vs. pregabalin	Milnacipran vs. pregabalin
Sleep disturbance SMD (95%CI)	6.20 (6.05 to 6.35)	0.14 (−0.01 to 0.29)	0.84 (0.69 to 0.99)
Depressed mood SMD (95%CI)	2.45 (2.32 to 2.58)	27.0 (26.83 to 27.17)	11.0 (1.89 to 11.11)
HRQOL SMD (95%CI)	1.47 (1.29 to 1.65)	0 (−0.25 to 0.25)	0.44 (0.28 to 0.60)

Abbreviations: CI, confidence interval; HRQOL, health-related quality of life; SMD, standard mean difference.

Comparisons to placebo

Gabapentin

One randomized, placebo-controlled, 12-week, fair-quality trial of 150 patients found that gabapentin 1800 mg (median) significantly improved pain severity, overall impact of fibromyalgia, global status, and sleep, but not tender point pain threshold, depression, or overall quality of life.[78] Diagnosis of fibromyalgia was based on the 1990 American College of Rheumatology criteria. Patients were 90% female, 97% white, and had a mean age of 48 years. Duration of fibromyalgia was not reported. A total of 19% of patients had a current major depressive disorder and 9% had a current anxiety disorder.

For pain, gabapentin was superior to placebo in reducing average pain severity score (−44% compared with −33%; *P*=0.015) and interference score (−53% compared with −32%; *P*=0.032) from the Brief Pain Inventory and in the proportion of patients who achieved a 30% response rate (51% compared with 31%; *P*=0.014; relative risk 1.65; 95% CI, 1.11 to 2.50; number needed to treat, 5), but not in improvement of the mean tender point pain threshold (+0.2 kg/cm^2 compared with +0.1 kg/cm^2; *P*=0.11). For overall impact fibromyalgia, there was a significantly greater reduction in Fibromyalgia Impact Questionnaire mean total score for gabapentin than placebo (−43% compared with −22%; *P*=0.001). For global status, there was a greater reduction Clinical Global Impression of Severity Scale scores for gabapentin than for placebo (−29% compared with −15%; *P*=0.002). Gabapentin also significantly reduced the Medical Outcomes Study Sleep Problems Index score (−40% compared with −14%; *P*=0.001). The difference between gabapentin and placebo did not reach statistical significance for increasing mean tender point threshold (+11% compared with +6%), reducing depression based on the Montgomery Asberg Depression Rating Scale score (−43% compared with −19%; *P*=0.067) or improving quality of life based on the Medical Outcomes Study, Short-Form 36 (data not reported) for sleep, Montgomery Asberg Depression Rating Scale for depression, and Medical Outcomes Study 36-item Short-Form Health Survey.

Cyclobenzaprine

Compared with placebo, a significant reduction in pain severity was only found with cyclobenzaprine in the largest (N=120)[38] of 3 fair-quality trials.[38, 40, 41] Data on pain could not be pooled across trials due to heterogeneity in outcome assessment. The trials included a total of 172 primarily female patients (range, 83% to 100%) with mean ages ranging from 43 years[41] to 49 years.[38] None of the trials used the 1990 American College of Rheumatology criteria for diagnosing fibromyalgia. In 2 cases, this was because the trials were conducted prior to 1990.[38, 40] History of fibromyalgia ranged widely in the 2 trials that reported this information.[38, 40] One trial enrolled newly diagnosed patients (4.5 months),[38] whereas mean duration of pain was 11.4 years in the second trial.[40] In all 3 trials, cyclobenzaprine dosage was based on a flexible regimen, starting at 10 mg and going up to a maximum dosage of 40 mg per day (means not reported).

The only trial that found a significant reduction in pain severity for cyclobenzaprine involved 120 patients with newly diagnosed fibrositis; 44% of whom had primary fibrositis and 56% had fibrositis considered to be associated with trauma or arthritis.[38] Patients were enrolled both from a university rheumatology clinic in Portland, Oregon and the Center for Arthritis and Back Pain in Philadelphia, Pennsylvania. Based on patient self-assessments (visual analog scale of 0-10), there was a significantly greater improvement with cyclobenzaprine than with placebo

in pain (28% compared with 17%; $P<0.02$) and sleep (34% compared with 18%; $P<0.02$) at 12 weeks. Also, physicians' evaluation of overall response to therapy found a significantly greater proportion of patients in the cyclobenzaprine group to have marked or moderate global improvement (34% compared with 16%; $P<0.012$; relative risk, 2.18; 95% CI, 1.12 to 4.36; number needed to treat, 5). However, at 12 weeks, there were no significant differences between cyclobenzaprine and placebo in duration of stiffness or fatigue, tender points, or physician-assessed muscle tightness or global pain.

In the smaller trials, there were trends favoring cyclobenzaprine over placebo in the number of patients with moderate to marked improvement in muscle pain at 6 weeks (43% compared with 18%; *P* not reported; N=40)[40] and in change in pain severity at 4 weeks (unspecified pain scale, scores ranged 0 to 60) in the evening (−8% compared with +4%; *P* not reported; N=12),[41] but not in the morning (−12% for both). These nonsignificant findings may have been the result of limitations in statistical power due to the small sample sizes in these studies.

Selective serotonin reuptake inhibitors

Among placebo-controlled trials of citalopram (N=82),[79, 80] fluoxetine (N=102),[81, 82] and immediate-release and controlled-release paroxetine (N=40),[83, 84] only fluoxetine, at a higher dose (45 mg), resulted in significantly greater improvements than placebo in pain, fatigue, and Fibromyalgia Impact Questionnaire Total Score.[81]

These trials enrolled 94% to 100% female patients with mean ages ranging from 31 years[83] to 50 years.[82] All but 1 trial[82] used the 1990 American College of Rheumatology criteria for diagnosing fibromyalgia. History of fibromyalgia ranged from 10 years[80] to 13 years.[82] Dosages ranged from 20 mg to 45 mg across all trials. All but 1 trial was rated fair quality. The remaining trial was rated poor quality due to concern over the potential confounding effects of higher numbers of high school graduates (90.5% compared with 61.9%; $P=0.03$) and longer disease duration (16.1 years compared with 9.6 years; $P=0.05$) in the fluoxetine group at baseline and the impact of excluding 43% of patients from the analyses.[82]

Due to small sample sizes and incomplete reporting of outcome data, evidence from these trials was insufficient for indirect meta-analysis and did not permit conclusions about comparative effectiveness. Table 6 summarizes the results for each selective serotonin reuptake inhibitor compared with placebo for the most consistently reported outcomes.

Among selective serotonin reuptake inhibitors, only fluoxetine, at a higher dose (45 mg), resulted in significantly greater improvements than placebo in pain, fatigue, and Fibromyalgia Impact Questionnaire Total Score in 1 small trial of 51 patients (Table 6).[81] Fluoxetine also resulted in a significantly greater reduction than placebo on the Fibromyalgia Impact Questionnaire depression subscore (−33% compared with +44%; $P=0.01$). Effects on sleep and global status were not reported.

As for paroxetine, although the controlled-release formulation (39.1 mg) did not result in a significant reduction in the Fibromyalgia Impact Questionnaire pain subscale (Table 6) and did not significantly increase the number of patients with a 50% reduction in Fibromyalgia Impact Questionnaire Total Score (26% compared with 14%; $P=0.08$), it was superior to placebo on the primary outcome of 25% reduction in Fibromyalgia Impact Questionnaire Total Score and on a few secondary outcome measures.[84] Controlled-release paroxetine was superior to placebo in reducing the Fibromyalgia Impact Questionnaire Total Score (Table 6), the Fatigue Subscore (Table 6), and in improving global status (Clinical Global Impression-Improvement, $P<0.005$),

but not for reducing depression (Beck Depression Inventory) or disability (Sheehan Disability Scale).[84] Evidence from the second, single-blind placebo-controlled trial of immediate-release paroxetine 20 mg (N=40) did not contribute useful information as the analyses focused on evaluating change from baseline in each group, respectively, and did not directly compare immediate-release paroxetine to placebo.[83]

Finally, citalopram demonstrated few advantages over placebo in only 1 (N=40)[79] of 2 studies.[79, 80] Neither trial found a significant improvement with citalopram for pain or fatigue (Table 6). Citalopram was superior to placebo in reducing the total score on the Montgomery Asberg Depression Rating Scale ($P<0.01$, data not reported),[79] but not in reducing the Beck Depression Scale score.[80] For sleep, scores on the Montgomery Asberg Depression Rating Scale sleep item improved significantly in the citalopram group, but not in the placebo group in 1 trial,[79] and there was no difference in Visual Analogue Scale score between the citalopram and placebo groups in the second trial.[80] There was no significant difference between citalopram and placebo in global status in either trial.

Table 6. Selective serotonin reuptake inhibitor compared with placebo: Mean changes in symptom severity

SSRI	Pain	Fatigue	FIQ Total score
Citalopram	−12% to −16% vs. −4% to −10%; NSD	7% vs. −1%, NSD	NR
Fluoxetine	−29% vs. −7%, *P*=0.002	−16% vs. +4%, *P*=0.05	−20% vs. +7%, *P*=0.005
Controlled-release paroxetine	*P*=0.07 (data NR)	*P*<0.05 (data NR)	−36% vs. −27%; *P*=0.015

Abbreviations: FIQ, Fibromyalgia Impact Questionnaire; NR, not reported; NSD, no significant differences; SSRI, selective serotonin reuptake inhibitor.

Key Question 2. For adults with fibromyalgia, what are the comparative harms of included interventions?

2a. When used as monotherapy?

2b. When used as adjunctive therapy?

Summary of Findings

General

- We found no eligible studies of treatment for fibromyalgia with desipramine, imipramine, desvenlafaxine, venlafaxine, escitalopram, fluvoxamine, sertraline, mirtazapine, bupropion, nefazodone, carbamazepine, divalproex, ethotoin, lacosamide, lamotrigine, levetiracetam, oxcarbazepine, phenytoin, tiagabine, topiramate, valproic acid, or zonisamide
- We found no eligible studies of included interventions when used as adjunctive therapy.

Direct evidence

- Although there were some significant differences between drugs in overall adverse events, they did not produce any differences in withdrawals due to adverse events
    - Overall and anticholinergic-type adverse events were significantly more frequent with amitriptyline 100 mg than with immediate-release paroxetine 20 mg, but withdrawals due to adverse events did not differ
    - Overall adverse events were significantly greater for nortriptyline than for amitriptyline, but individual adverse events and withdrawals due to adverse events were not significantly different
    - No significant differences were found between cyclobenzaprine and amitriptyline in any harms outcomes.

Indirect evidence

Indirect meta-analysis

- No difference in overall withdrawal compared with placebo was found for amitriptyline, pregabalin, and duloxetine; milnacipran was found to have a small but significant increase in overall withdrawal and no difference was found between the drugs on overall withdrawal
- Pregabalin, duloxetine, and milnacipran had a significant increase in overall adverse events compared with placebo whereas amitriptyline was found to be no different than placebo on this outcome; there was low evidence that no differences exist between pregabalin, duloxetine, and milnacipran on overall adverse events, with insufficient evidence to report on this outcome for amitriptyline
- Pregabalin, duloxetine, and milnacipran had a significant increase in withdrawal due to adverse events compared to placebo whereas amitriptyline was found to be no different than placebo on this outcome; there was low evidence that no differences exist between pregabalin, duloxetine, and milnacipran on withdrawal due to adverse events, with insufficient evidence to report on this outcome for amitriptyline
- All drugs were generally well-tolerated with greater adverse events reported compared to placebo; pregabalin had significantly less headache, nausea, and diarrhea compared to duloxetine, and significantly less headache and nausea compared to milnacipran
- Duloxetine and milnacipran had significant increase of hyperhidrosis compared with placebo with no differences between the drugs (relative risk, 1.14; 95% CI, −5.08 to 7.36)
- Milnacipran had significant tachycardia compared to placebo (number needed to harm, 20.6; 95% CI, 15.1 to 29.1)
- Pregabalin had significant weight gain and peripheral edema compared with placebo (relative risk, 4.58; 95% CI, 2.44 to 6.82 and relative risk, 3.52; 95% CI, 3.52, 2.01 to 6.18, respectively).

Comparisons to placebo

- Gabapentin: Despite greater rates of various individual adverse events (e.g., dizziness, sedation, lightheadedness, and weight gain), withdrawals due to adverse events were not significantly greater with gabapentin than placebo

- Cyclobenzaprine: Dry mouth, drowsiness, and overall adverse events were more frequent with cyclobenzaprine than placebo, but withdrawals due to adverse events did not differ between treatment groups
- Selective serotonin reuptake inhibitors: Compared with placebo, significant increases in adverse events were only found in the largest trial of controlled-release paroxetine, including drowsiness, dry mouth, and ejaculatory problems.

Detailed Assessment

Direct evidence

Direct evidence regarding the comparative harms among included interventions was limited to 3 fair-quality randomized controlled trials that evaluated amitriptyline with cyclobenzaprine,[43] nortriptyline,[45] and immediate-release paroxetine.[46] A detailed description of the trial design and patient characteristics can be found above in Key Question 1.

Although there were some significant differences between drugs in overall adverse events, they did not produce any differences in withdrawals due to adverse events (Table 7).[43, 45, 46] Among individual adverse events, the only significant difference came from the trial that compared immediate-release paroxetine 20 mg and amitriptyline 100 mg, in which anticholinergic-type adverse events (e.g., dry mouth, constipation, urinary retention) were more frequent with such a high dose of amitriptyline (41% compared with 9%; $P<0.004$).[46] Otherwise, there were no significant differences between amitriptyline and either cyclobenzaprine[43] or nortriptyline.[45] In the trial that compared amitriptyline to cyclobenzaprine, dry mouth, somnolence, dizziness, and weight gain were cited as being the most frequently reported adverse events, but incidence rates were not reported.[43] The most frequent adverse events for amitriptyline and nortriptyline were dry mouth (10% compared with 16%), abdominal pain (10% compared with 18%), and dizziness (5% compared with 10%).[45]

Table 7. Adverse events in head-to-head trials

Comparison Sample size Trial duration	Overall adverse events	Withdrawals due to adverse events
Cyclobenzaprine 30 mg vs. amitriptyline 50 mg[43] N=184 6 months	98% vs. 95% RR, 1.02 (95% CI, 0.96 to 1.11)	16% vs. 8% RR, 1.90 (95% CI, 0.82 to 4.44)
Nortriptyline 25 mg vs. amitriptyline 25 mg[45] N=106 8 weeks	82% vs. 40% RR, 2.04 (95% CI, 1.53 to 2.72)	3% vs. 0 RR, 3.15 (95% CI, 0.28 to infinity
Immediate-release paroxetine 20 mg vs. amitriptyline 100 mg[46] N=68 6 weeks	37% vs. 93% RR, 0.40 (0.24 to 0.60)	6% vs. 17% RR, 0.36 (0.08 to 1.49)

Abbreviations: CI, confidence interval; RR, relative risk.

Indirect evidence

We found 3 good-quality systematic reviews that assessed multiple drugs for the treatment of fibromyalgia and analyzed harm.[48, 50, 85] One of these reviews performed an indirect meta-analysis to compare the effectiveness and harms between duloxetine, milnacipran, and pregabalin in fibromyalgia from placebo-controlled trials of individual drugs.[85] Although it was a recent review, we chose to update the indirect meta-analysis to include evidence from 2 new trials for milnacipran that are now available.[52, 53] We have compared our results to the Hauser, et al. analysis and discussed where applicable below. One systematic review of amitriptyline reported on 10 randomized controlled trials but did not perform a meta-analysis due to large clinical variability and statistical heterogeneity.[50] The results are considered as appropriate below. One systematic review pooled data of pregabalin and gabapentin and reported on a class effect size rather than comparative data, which limited its usefulness for this report.[48]

Of the 8 drugs included in this review, we found that only 4 drugs had multiple trials with sample sizes adequate to perform a comparative analysis. All trials used the drugs as monotherapy and no trial evaluated the drugs as adjunctive therapy. We performed a meta-analysis of 6 placebo-controlled trials of amitriptyline,[16, 39, 43, 45, 54, 55] 4 placebo-controlled trials of pregabalin,[57-60] 5 placebo-controlled trials of milnacipran,[52, 53, 61-64] and 4 placebo-controlled trials of duloxetine [65-69] to determine comparative harm. Given that Hauser, et al. performed a recent meta-analysis to determine comparative harm, and given that the harms data from the 2 additional milnacipran trials were consistent with prior results, we elected to study the differences between the drugs on measures of overall withdrawal, overall adverse events, and withdrawal due to adverse events.[49, 52, 53] None were of long duration (8 to 28 weeks) with the amitriptyline trials being of the shortest duration (mean 9 weeks). Sample size was small for all of the amitriptyline trials (N=22 to 126) but generally moderate to large for the rest (N=125 to 1025). All of the trials were rated fair quality. One poor-quality trial of amitriptyline was excluded from our analysis.[73] The trials of duloxetine, milnacipran, and pregabalin were industry sponsored whereas only 1 of the amitriptyline trials were industry sponsored.[16] The baseline demographics for all of the trials were similar, with the majority of patients being middle aged (mean age range 40-53 years), white (65% to 93%), and female (82% to 100%). For amitriptyline, we excluded data on the lowest dose of 12.5 mg once daily from 1 trial, and for pregabalin we excluded data on the lowest dose of 150 mg once daily. Otherwise we combined data on different daily dosage groups including amitriptyline 25-50 mg, pregabalin 300-600 mg, milnacipran 100-200 mg, and duloxetine 40-120 mg. We reported the outcomes of clinical importance and with adequate data to perform the analysis. We performed our analysis using short-term data (8-15 weeks) as this was consistent across most studies. We excluded trials shorter than 8 weeks from our analysis.[42, 44] We performed a sensitivity analysis to determine if differences in outcome were seen using both short- and long-term data. The results and quality of these trials are summarized in Evidence Tables 1 and 2.

Overall withdrawal

Pooled analysis of short-duration (8-15 weeks) placebo-controlled trials found no difference in overall withdrawal of amitriptyline, pregabalin, and duloxetine compared with placebo whereas milnacipran was found to have a small but significant increase in overall withdrawal compared with placebo (Table 8). There was low evidence that no differences exist between the drugs on overall withdrawal when we included all of the trials (Table 9). This result held true for our sensitivity analysis using all trials (6-28 weeks) of pregabalin, duloxetine, and milnacipran.

There was insufficient long-term data to include amitriptyline in this analysis. When we excluded the Carette 1986 trial that was conducted prior to the development of the 1990 American College of Rheumatology criteria for fibromyalgia, amitriptyline appeared to have less overall withdrawal, however it is felt that this may be reflective of the lower sample size (N=208).

Overall adverse events

Pooled analysis of placebo-controlled trials of amitriptyline, pregabalin, duloxetine, and milnacipran found a significant increase in overall adverse events of pregabalin, duloxetine, and milnacipran compared with placebo, whereas amitriptyline was found to be no different than placebo on this outcome (Table 8). This held true when we excluded the Ginsberg trial, which used a sustained-release formulation of amitriptyline, and the Carette 1986 trial.[39, 55] Given the small sample size in the pooled analysis of amitriptyline, we felt that the data was insufficient to make any conclusive statements. There was low evidence that no differences exist between the other drugs on overall adverse events (Table 9). This result held true for our sensitivity analysis using all trials (6-28 weeks) of pregabalin, duloxetine, and milnacipran. There was insufficient long-term data to include amitriptyline in this analysis.

Withdrawal due to adverse events

Pooled analysis of placebo-controlled trials of amitriptyline, pregabalin, duloxetine, and milnacipran found a significant increase in withdrawal due to adverse events of pregabalin, duloxetine, and milnacipran compared with placebo whereas amitriptyline was found to be no different than placebo on this outcome (Table 8). Given the small sample size in the pooled analysis of amitriptyline, we felt that the data was insufficient to make any conclusive statements regarding this outcome. There was low evidence that no differences exist between the other drugs on withdrawal due to adverse events (Table 9). This result held true for our sensitivity analysis using all trials (6-28 weeks) of pregabalin, duloxetine, and milnacipran. There was insufficient long-term data to include amitriptyline in this analysis.

Table 8. Pooled effectiveness of amitriptyline, pregabalin, milnacipran, and duloxetine compared with placebo (8-15 weeks)

Outcome measure	Amitriptyline	Pregabalin	Milnacipran	Duloxetine
Overall withdrawal RR (95% CI) vs. placebo	0.64 (0.35 to1.15) [39, 43, 45, 54, 55]	1.18 (0.94 to 1.32) [58-60, 86]	1.27 1.02 to 1.58) [52, 53, 61, 62, 64]	0.94 (−.76 to 1.16) [65, 66, 69]
Overall withdrawal RR (95% CI) vs. placebo (excluding Ginsberg 1996)[55]	1.08 (0.94 to 1.24) [39, 43, 45, 54]	---	---	---
Overall withdrawal RR (95% CI) vs. placebo (excluding Carette 1986)[39]	0.48 (0.28 to 0.83) [38, 40, 50, 51]	---	---	---
Overall adverse events RR (95% CI) vs. placebo	1.68 (0.78 to 3.62) [39, 43, 45, 55]	1.18 (1.13 to 1.24) [58-60, 86]	1.12 (1.08 to 1.16) [52, 53, 61]	1.18 (1.09 to 1.27) [65, 66]
Overall adverse events RR (95% CI) vs. placebo (excluding Ginsberg 1996)	1.45 (0.68 to 3.12) [39, 43, 45]	---	---	---
Overall adverse events RR (95% CI) vs. placebo (excluding Carette 1986)[39]	1.02 (0.42 to 2.49) [38, 40]	---	---	---
Withdrawal due to adverse event RR (95% CI) vs. placebo	0.99 (0.32 to 3.11) [39, 43, 45]	1.95 (1.54 to 2.47) [58-60, 86]	1.92 (1.31 to 2.82) [52, 53, 61, 62, 64]	1.61 (1.16 to 2.23) [65, 66, 69]
Withdrawal due to adverse events RR (95% CI) vs. placebo (excluding Carette 1986)[39]	0.96 (0.23 to 3.99) [38, 40]	---	---	---

Abbreviations: CI, confidence interval; RR, relative risk.

Table 9. Indirect analysis of placebo-controlled trials in fibromyalgia

	Duloxetine vs. milnacipran	Duloxetine vs. pregabalin	Duloxetine vs. amitriptyline	Milnacipran vs. pregabalin	Milnacipran vs. amitriptyline	Pregabalin vs. amitriptyline
Overall Withdrawal Ratio of RR (95% CI)	0.74 (0.55 to 1.00)	0.84 (0.64 to 1.10)	1.47 (0.79 to 2.75)	1.14 (0.86 to 1.50)	1.99 (1.06 to 3.72)	1.75 (0.95 to 3.23)
Overall Withdrawal Ratio of RR (95% CI) Excluding Ginsberg 1996[55]			1.45 (0.66 to 3.18)		1.96 (0.89 to 4.30)	1.72 (0.79 to 3.74)
Overall withdrawal Ratio of RR (95% CI) Excluding Carette 1986[39]			1.94 (1.09 to 3.45)		2.62 (1.47 to 4.67)	2.31 (1.32 to 4.04)
Overall adverse events Ratio of RR (95% CI)	1.04 (0.96 to 1.14)	0.99 (0.90 to 1.09)	0.70 (0.32 to 1.51)	0.95 (0.89 to 1.01)	0.67 (0.31 to 1.44)	0.71 (0.33 to 1.52)
Overall adverse events Ratio of RR (95% CI) Excluding Ginsberg 1996			0.81 (0.37 to 1.74)		0.77 (0.36 to 1.66)	0.81 (0.38 to 1.75)
Overall adverse events Ratio of RR (95% CI) Excluding Carette1986[39]			0.94 (0.39 to 2.25)		0.90 (0.38 to 2.14)	0.94 (0.40 to 2.26)
Withdrawal due to adverse events Ratio of RR (95% CI)	0.84 (0.51 to 1.39)	0.83 (0.56 to 1.24)	1.62 (0.50 to 5.32)	0.99 (0.63 to 1.55)	1.94 (0.58 to 6.46)	1.96 (0.61 to 6.30)
Withdrawal due to adverse events Ratio of RR (95% CI) Excluding Carette 1986[39]			1.68 (0.39 to 1.55)		2.01 (0.46 to 8.80)	2.04 (0.48 to 8.64)

Abbreviations: CI, confidence interval; RR, relative risk.

Other adverse events

Hauser, et al. performed a good-quality systematic review with an indirect meta-analysis to compare the harms between duloxetine, milnacipran, and pregabalin in fibromyalgia from placebo-controlled trials of individual drugs.[49] No drug-related deaths were reported and all drugs were generally well tolerated. They found that duloxetine and milnacipran had significantly greater reporting of headache and nausea compared with pregabalin but no difference between each other (Table 10). Duloxetine also had increased diarrhea compared with milnacipran and pregabalin and no difference between the later 2 drugs (Table 10). All drugs had a significant increase in dry mouth, constipation, and dizziness compared with placebo and no difference between the drugs.[49] Both duloxetine and pregabalin had a significant increase in fatigue and somnolence compared with placebo and no difference between the 2 drugs.[49]

Reporting of hyperhidrosis was unique to duloxetine and milnacipran with no differences between the drugs on this outcome (relative risk, 1.14; 95% CI, −5.08 to 7.36).[49]

Milnacipran was the only drug that reported tachycardia (number needed to harm, 20.6; 95% CI, 15.1 to 29.1).[49] Pregabalin was the only drug that reported weight gain and peripheral edema, both of which were significantly greater than placebo (relative risk, 4.58; 95% CI, 2.44 to 6.82; and relative risk, 3.52; 95% CI, 2.01 to 6.18 respectively).[49] All drugs reported rare serious adverse events including risk of suicide (duloxetine 1.1% based on one 1-year trial, milnacipran 1.3% in those with depression at baseline, and pregabalin <1%) and hepatic-related adverse events.[49]

Table 10. Indirect analysis of harms from placebo-controlled trials of pregabalin, milnacipran, and duloxetine for fibromyalgia[49]

	Duloxetine vs. milnacipran	Duloxetine vs. pregabalin	Milnacipran vs. pregabalin
Headache	1.24	2.24	1.81
RR (95% CI)	(0.71 to 1.75)	(1.83 to 2.65)	(1.48 to 2.14)
Nausea	1.72	3.25	1.90
RR (95% CI)	(0.60 to 2.84)	(2.13 to 4.37)	1.37 to 2.43)
Diarrhea	2.21	2.01	1.03
RR (95%CI)	(1.64 to 2.78)	(1.23 to 2.99)	(0.10 to 1.96)

Abbreviations: CI, confidence interval; RR, relative risk.

Nishishinya, et al. conducted a systematic review of placebo-controlled trials of amitriptyline but did not perform a pooled analysis. Of 6 trials reporting adverse events, they found that the mean adverse event rate for amitriptyline was 51.84% (2.8% to 95%) compared with 36.63% (2.8% to 80%) for placebo.[50] The high placebo event rate questioned the validity of the results given that 2 trials had higher adverse event rates in the placebo arm compared with the amitriptyline arm.[50] Amitriptyline was generally well tolerated in all the trials with no severe or life threatening events reported.[50] Somnolence, dry mouth, gastrointestinal symptoms, and weight gain were the most frequently reported adverse events and there were no differences in withdrawal due to adverse events compared with placebo.[50]

Comparisons to placebo

Gabapentin
Over 12 weeks, dizziness (25% compared with 9%; $P<0.05$; relative risk, 2.86; 95% CI, 1.3 to 6.29, number needed to harm, 6), sedation (24% compared with 4%; $P<0.001$; relative risk, 6.33; 95% CI, 2.12 to 19.54; number needed to harm, 5), lightheadedness (15% compared with 1%; $P<0.01$; relative risk, 11.00; 95% CI, 1.91 to 65.55; number needed to harm, 7), and weight gain (8% compared with 0; $P<0.05$; relative risk, 13.00; 95% CI, 1.61 to infinity; number needed to harm, 12) were more frequent with gabapentin 1800 mg than placebo.[78] But, withdrawals due to adverse events did not differ significantly in the gabapentin and placebo groups (16% compared with 9%; $P=0.34$).

Cyclobenzaprine
Data on harms were inconsistently reported across placebo-controlled trials of cyclobenzaprine. Although the incidence of dry mouth (pooled rates, 56% compared with 20%; pooled relative risk, 2.75; 95% CI, 1.71 to 4.43; number needed to harm, 3),[38, 40] drowsiness (34% compared with 17%; $P<0.01$; relative risk, 1.96; 95% CI, 1.04 to 3.82; number needed to harm, 6),[38] and overall adverse events (89% compared with 64%; relative risk, 1.39; 95% CI, 1.14 to 1.76; number needed to harm, 4)[38] was greater with cyclobenzaprine than placebo, withdrawals due to adverse events did not differ between treatment groups (pooled rates: 6% compared with 2%; pooled relative risk, 2.56; 95% CI, 0.48 to 13.59).[38, 40, 41]

Selective serotonin reuptake inhibitors
Adverse events were sparsely reported in placebo-controlled trials of selective serotonin reuptake inhibitors.[79-84] Compared with placebo, withdrawals due to adverse events were similar for citalopram (14% compared with 0%; *P* not reported; N=40)[79] and for controlled-release paroxetine (7% compared with 2%; *P* not reported; N=116),[84] but were not reported for fluoxetine. Overall adverse events were only reported in 1 trial of controlled-release paroxetine, and were not significantly different from placebo (65% compared with 59%; *P* not reported).[84] Compared with placebo, in the largest trial (N=116), controlled-release paroxetine was the only selective serotonin reuptake inhibitor to significantly increase any specific adverse events, including drowsiness (26% compared with 7%; *P* not reported), dry mouth (36% compared with 9%; *P* not reported), and ejaculatory problems (66% compared with 2%; *P* not reported).[84]

Key Question 3. Are there subgroups of patients based on demographics (age, racial or ethnic groups, and gender), socioeconomic status, other medications, or comorbidities for which any included drugs are more effective or associated with fewer harms?

Summary of Findings

Direct evidence

- Extremely limited direct evidence exists regarding treatment of fibromyalgia in subgroup populations

- Response to either amitriptyline or cyclobenzaprine did not differ on the basis of age.

Indirect evidence

- Nine individual trials performed subgroup analysis
- The majority of patients in all individual trials were middle-aged, white (84% to 91%), and female (89% to 100%), and there was a high prevalence of baseline anxiety and depression
- Duloxetine was no different than placebo in pain response in male patients (−1.24; standard error, 0.4 compared with −1.25; standard error, 0.5; *P*=0.969), those ≥ 65 years, (−1.92; standard error, 0.3 compared with −1.50; standard error, 0.4; *P*=0.374), and nonwhite patients (−1.70; standard error, 0.3 compared with −1.37; standard error, 0.3, *P*=0.386); this result should be interpreted with caution given the small number of patients in each group
- Duloxetine: Based on 2 trials, effect on pain was estimated to be 61% to 86% after accounting for the indirect effect on depression
- Pregabalin: Based on 1 trial, effect on pain was estimated to be 75% after accounting for the indirect effect on anxiety and depression
- Milnacipran: Greater improvement in pain reduction was seen in nondepressed patients compared with depressed patients; however, the small number of patients in each group limits the interpretation of these results
- Cyclobenzaprine: Race, age, nor sex were found to influence effectiveness as compared to placebo
- Selective serotonin reuptake inhibitors: History of depression was not associated with response to either fluoxetine or controlled-release paroxetine, nor was anxiety associated with response to controlled-release paroxetine.

Detailed Assessment

We identified 10 trials that performed subgroup analysis of the included drugs.[38, 43, 62, 65, 66, 78, 81, 87-89] The baseline demographics for all of the trials were similar, with the majority of patients being middle aged (mean age range 44-50 years), white (84% to 91%), and female (89% to 100%). Analysis of comparative effectiveness in men, nonwhites, and older patients was limited by the small number of patients in these subgroups and the corresponding lack of reporting of data in the individual trials. In the 3 trials that reported time since diagnosis of fibromyalgia, the duration ranged from 49 months to 129 months.[43, 62, 87]

Direct evidence

Only 1 head-to-head trial addressed drugs used for fibromyalgia in subgroups of the population, and only assessed the impact of age.[43] In a fair-quality randomized controlled trial of 208 patients who met the 1990 American College of Rheumatology criteria for fibromyalgia, variation in age did not differentiate response at 6 months to either amitriptyline 50 mg or cyclobenzaprine 30 mg. Response was defined as meeting at least 4 of the following 6 criteria: 50% improvement in pain, sleep, fatigue, patient global assessment, or physician global assessment, and increase of 1 kg in mean total myalgic score.

Indirect evidence

The small number of patients in each subgroup and the limited reporting of subgroup data did not allow for indirect meta-analysis of the outcome measures for comparative effectiveness.

Duloxetine

One randomized double-blind fair-quality placebo-controlled trial restricted inclusion of patients to females only (N=354).[66] The results for all outcome measures did not differ from the indirect meta-analysis revealing that duloxetine either 60 mg once daily or 60 mg twice daily is more effective than placebo at 12 weeks on outcomes of pain, Fibromyalgia Impact Questionnaire, and patient global impression of improvement. The 50% response rate in pain was significant and ranged from 41% for duloxetine compared with 23% for placebo (P=0.003). A pooled subgroup analysis of female patients in all 4 duloxetine placebo-controlled trials (n=1262) found a statistically significant difference in pain reduction compared to placebo with mean least squares change in 24-hour pain measures on the brief pain inventory, a 10-point scale of −1.74 (standard error, 0.1) for the duloxetine group and −1.10 (standard error, 0.1) for the placebo group (P<0.001).[65-69]

A pooled analysis of all 4 duloxetine placebo-controlled trials evaluated the response of male patients, those age 65 or older, and nonwhite patients.[75] Unlike the results of the total population, the difference in pain response between duloxetine and placebo was not significant in male patients, n=70, (−1.24; standard error, 0.4 compared with −1.25; standard error, 0.5; P=0.969), those ≥ 65 years, number of patients not reported, (−1.92; standard error, 0.3; compared with −1.50; standard error, 0.4; P=0.374), and nonwhite patients, n=147, (−1.70; standard error, 0.3 compared with −1.37; standard error, 0.3; P=0.386). Despite the pooling of 4 moderate sized trials, there remained a small number of patients in each subgroup resulting in the analysis being underpowered to detect a true difference if a difference exists. Further analysis demonstrated that treatment by subgroup interaction was not significant in their overall analysis of results on pain measures (sex, P=0.320; age, P=0.362; race, P=0.180).

Two trials of duloxetine in women performed an analysis to determine whether the treatment effect of duloxetine in fibromyalgia was independent of its effect on major depressive disorder or anxiety.[65, 66] A regression model was used to determine if the treatment has an effect on pain reduction after accounting for the indirect effect through improvement of depressive symptoms. In Arnold 2004, the direct effect of duloxetine on reduction of pain accounted for 61.1% to 83.3% of the total treatment effect, with an indirect effect through improvement in depressive symptoms accounting for 15.3% to 38.5%, and improvement in anxiety accounting for 0.5% to 1.5% of the total effect.[65] In Arnold 2005, the direct effect of duloxetine on reduction of pain accounted for 75.6% to 85.9% of the total treatment effect with an indirect effect through improvement in depressive symptoms accounting for 13.1% to 24.4%. In a subgroup of 91 patients with a current diagnosis of major depressive disorder, a significant reduction in the Brief Pain Inventory average pain severity score was demonstrated between duloxetine and placebo (60 mg once daily, P=0.005 and 60 mg twice daily, P=0.003).

Milnacipran

One 12-week randomized placebo-controlled fair-quality trial evaluated the effect of depression comorbidity on response to once or twice-daily milnacipran in patients diagnosed with fibromyalgia.[62, 64] The mean dose was 174 mg for the once daily group and 191 mg for the twice daily group. There was statistically less baseline depression in the once daily milnacipran group

(n=3; 7%) compared with the milnacipran twice daily group (n=8; 16%) and placebo group (n=9; 32%). Greater improvement in pain reduction was seen in nondepressed patients compared with depressed patients treated with milnacipran; however, the small number of patients in each group and the high placebo response rate in depressed patients limited the interpretation of these results. The authors attributed this difference to the higher placebo response rate among depressed patients. For twice daily milnacipran, 25% of depressed patients and 37% of nondepressed patients had 50% improvement in daily pain score compared with 44% and 0% respectively for placebo.

Pregabalin

One 8-week randomized double-blind placebo-controlled trial (N=529) compared pregabalin 450 mg once daily to placebo and evaluated the impact of anxiety and/or depression on the effectiveness of pregabalin in reducing pain compared with placebo.[87] At baseline patients completed the Hospital Anxiety and Depression Scale which consists of two 7-item subscales, 1 for depression and 1 for anxiety. A regression model was used to determine if the treatment has an effect on pain reduction after accounting for the indirect effect through improvement of depressive symptoms. Baseline anxiety and depression fell in the mild range for both, with 29% of patients being free of anxiety and 44% of patients being free of depression at baseline indicating significantly higher baseline anxiety ($P<0.0001$). Baseline mean pain was 7.0 (standard deviation, 1.3). There was a statistically significant reduction in pain in the pregabalin group with a mean score of 4.94 compared with 5.88 in the placebo group ($P=0.001$). Using a path analysis based on linear regression models, they estimated that improvement in anxiety and depression each indirectly attributed 12.4% of pain relief, with 75.3% attributable to the direct treatment effect of pregabalin. This trial excluded patients with clinically significant psychiatric illness at the time of enrollment which may have limited the interpretation of these results.
A pooled post-hoc analysis of the 4 placebo-controlled pregabalin trials was performed by Emir, et al. to evaluated whether clinical characteristics of patients at baseline influenced the magnitude of pain response to pregabalin.[89] They used a covariate interaction to estimate mean pain changes and least square means across different levels of baseline covariates for 2061 patients, including baseline anxiety and depression. Of 2032 patients with this data, they found no significant interactions between treatment effectiveness and baseline anxiety or depressive symptoms (anxiety, $P=0.654$; depression, $P=0.689$).[89] They did find that patients with a higher pain score at baseline, a higher sleep score based on the medical outcomes sleep score index (MOS, 0-100 scale), and older age were all associated with a more robust and significantly greater improvement in pain ($P<0.0001$ for each).[89]

Cyclobenzaprine

In one 12-week trial of 120 patients, race, age, nor sex were found to influence the effectiveness of cyclobenzaprine compared with placebo.[38]

Selective serotonin reuptake inhibitors

Two fair-quality randomized controlled trials consistently found that a history of depression was not associated with response to either fluoxetine[81] or controlled-release paroxetine.[88] In one 12-week trial of 60 patients, analysis of covariance revealed no significant interaction between the effect of fluoxetine 45 mg on pain and history of major depressive disorder, baseline level of depression, or improvement in depression.[81] In a post-hoc logistic regression analysis of data

from a 12-week placebo-controlled trial of controlled-release paroxetine in 116 patients with fibromyalgia, history of depression and/or anxiety did not predict treatment response as measured by at least a 25% reduction in Fibromyalgia Impact Questionnaire score (odds ratio, 0.66; 95% CI, 0.29 to 1.49).[88]

SUMMARY

Strength of Evidence

The results of this review are summarized in Table 11, below, and Appendix F summarizes the strength of the evidence for each key question.

The strength of evidence in patients with fibromyalgia was generally low with limited direct evidence. We did find low evidence that immediate-release paroxetine was more effective in reducing pain than amitriptyline, with low to insufficient evidence that other differences exist between immediate-release paroxetine, cyclobenzaprine, or nortriptyline compared with amitriptyline. Using indirect meta-analysis of placebo-controlled trials, we found low evidence that duloxetine was superior to milnacipran in reduction of pain, sleep, depression, and health-related quality of life, and low evidence that there is no difference on other measures of effectiveness. We found low evidence that both duloxetine and milnacipran are more effective than pregabalin on depressed mood, and that pregabalin is superior to milnacipran on sleep disturbance. We found low evidence that further differences exist between these 3 drugs. We found low evidence that differences exist between duloxetine, milnacipran, pregabalin, and amitriptyline on other measures of effectiveness.

On measures of harm, all drugs are generally well-tolerated with low evidence that overall adverse events are greater with amitriptyline compared with immediate-release paroxetine and with nortriptyline compared with amitriptyline, and that no significant differences exist between the other drugs on withdrawal and overall adverse event reporting. We found low evidence that response to amitriptyline or cyclobenzaprine does not differ based on age with insufficient evidence to report on other subgroup analyses.

Limitations of this Report

As with other types of research, the limitations of this systematic review are important to recognize. These can be divided into 2 groups, those relating to generalizability of the results and those relating to methodology within the scope of this review. The generalizability of the results are limited by the scope of the Key Questions and inclusion criteria and by the generalizability of the studies included. Most studies included narrowly defined populations of patients who met strict criteria for case definition, had few comorbidities, and used few or no concomitant medications. Minorities, older patients, male patients, and the most seriously ill patients were underrepresented. Most studies excluded patients with major depressive disorder yet mood disorder is a significant component of the spectrum of fibromyalgia.

Methodological limitations of the review within the defined scope included the exclusion of studies published in languages other than English and lack of a specific search for unpublished studies. Measurement of effectiveness outcomes varied considerably across trials by the use of different instruments and different timing of measurements limiting the validity of combining

scores to allow for comparison between drugs. Few direct head-to-head comparisons of the included drugs have been conducted, limiting our conclusions to indirect comparison of placebo-controlled trials for many of the outcomes. This limited the strength of the evidence due to heterogeneity of trial populations, interventions, and outcomes assessment.

Applicability

One potential limitation to the applicability of the findings of this review is that they relate to a narrower range of drugs than are available in clinical practice. The selection of drugs included in this review was influenced by the specific programmatic interests of the organizations participating in the Drug Effectiveness Review Project and are not meant to be read as a usage guideline. Of the drugs studied, trials differed with respect to dosing regimens limiting any conclusions about optimal dose. Additionally, most trials excluded patients with major depressive disorder and some trials excluded patients who had failed to respond to other antidepressant medications or were unable to tolerate assigned stable doses, thus limiting the applicability to an actual clinical practice. Given that fibromyalgia is a chronic disease, the applicability of results from short-term trials such as those included in this report may be limited. In clinical practice, a multimodal treatment approach is often invoked involving multiple drugs and multiple nonpharmacological interventions. Although we planned to review a multimodal approach of the included drugs, we found no eligible studies that included interventions when used as adjunctive therapy.

Studies Pending Review

We identified no trials in progress that would meet inclusion criteria for this review and would potentially change conclusions.

Table 11. Summary of the evidence by key question

Key question	Comparison	Strength of evidence	Conclusion
1. For adults with fibromyalgia, what is the comparative effectiveness/efficacy of included interventions?			
1a. When used as monotherapy?			
Direct evidence	Immediate-release paroxetine vs. amitriptyline	Low	Pain: Significantly greater reduction with immediate-release paroxetine
		Low	Fatigue: No significant difference
		Insufficient	50% response, FIQ mean change: No data available
	Cyclobenzaprine vs. amitriptyline	Low	Pain and fatigue: No significant differences
		Insufficient	50% response, FIQ mean change: No data available
	Nortriptyline vs. amitriptyline	Low	Pain and FIQ: No significant differences
		Insufficient	50% response, FIQ mean change: No data available
Indirect evidence	Duloxetine vs. milnacipran	Low	Pain, sleep disturbance, depressed mood, and HRQOL: Significantly greater improvement with duloxetine 50% response, Fatigue and FIQ mean change: No significant difference
	Duloxetine vs. pregabalin	Low	Depressed mood: Significantly greater improvement with duloxetine Pain, 50% response, Fatigue, FIQ mean change, SF-36 physical and mental components, sleep disturbance, and HRQOL: No significant difference
	Duloxetine vs. amitriptyline	Low	Pain and Fatigue: No significant difference
		Insufficient	50% response and FIQ mean change: No significant difference
	Milnacipran vs. pregabalin	Low	Depressed mood: Significantly greater improvement with milnacipran Sleep disturbance: Significantly greater improvement with pregabalin Pain, 50% response, 30% response, Fatigue, FIQ, and SF-36 physical and mental components: No significant difference
	Milnacipran vs. pregabalin	Insufficient	PGII or PGIC: No significant difference
	Milnacipran vs. amitriptyline	Low	Pain, Fatigue: No significant difference
		Insufficient	50% response, FIQ, and PGII or PGIC: Insufficient data
	Pregabalin vs. amitriptyline	Low	Pain, Fatigue: No significant difference
		Insufficient	50% response, FIQ, and PGII or PGIC: Insufficient data
	Gabapentin, cyclobenzaprine, citalopram, fluoxetine, controlled-release paroxetine vs. other drugs	Insufficient	No conclusions can be drawn about comparative effectiveness/efficacy because the numbers of trials/patients were too few to provide meaningful results in indirect comparisons
1b. When used as adjunctive therapy?			
	All	Insufficient	No evidence found

Key question	Comparison	Strength of evidence	Conclusion
2. For adults with fibromyalgia, what are the comparative harms of included interventions?			
2a. When used as monotherapy?			
Direct evidence	Immediate-release paroxetine vs. amitriptyline	Low	Overall AE: Significantly greater with amitriptyline
		Low	Withdrawals due to adverse events: No significant difference
	Cyclobenzaprine vs. amitriptyline	Moderate	Overall AE: No significant difference
		Low	Withdrawals due to adverse events: No significant difference
	Nortriptyline vs. amitriptyline	Moderate	Overall AE: Significantly greater with nortriptyline
		Low	Withdrawals due to adverse events: No significant difference
Indirect evidence	Duloxetine vs. milnacipran	Low	Overall withdrawal, overall adverse events, and withdrawal due to adverse events: No significant difference Headache and Nausea: No significant difference Diarrhea: Significantly greater with duloxetine
	Duloxetine vs. pregabalin	Low	Overall withdrawal, overall adverse events, and withdrawal due to adverse events: No significant difference Headache, Nausea, and Diarrhea: Significantly greater with duloxetine
	Duloxetine vs. amitriptyline	Low	Overall withdrawal: No significant difference
		Insufficient	Overall adverse events and withdrawal due to adverse events: No significant difference
	Milnacipran vs. pregabalin	Low	Overall withdrawal, overall adverse events, withdrawal due to adverse events, and diarrhea: No significant difference Headache and nausea: Significantly greater with milnacipran
	Milnacipran vs. amitriptyline	Low	Overall withdrawal: No significant difference
		Insufficient	Overall adverse events and withdrawal due to adverse events: No significant difference
	Pregabalin vs. amitriptyline	Low	Overall withdrawal: No significant difference
		Insufficient	Overall adverse events and withdrawal due to adverse events: No significant difference
	Gabapentin, cyclobenzaprine, citalopram, fluoxetine, controlled-release paroxetine vs. other drugs	Insufficient	No conclusions can be drawn about comparative harms because the numbers of trials/patients were too few to provide meaningful results in indirect comparisons
2b. When used as adjunctive therapy?			
	All	Insufficient	No evidence found

Key question	Comparison	Strength of evidence	Conclusion
3. Are there subgroups of patients based on demographics (age, racial or ethnic groups, and gender), socioeconomic status, other medications, or comorbidities for which any included drugs are more effective or associated with fewer harms?			
	Amitriptyline, cyclobenzaprine	Low	**Age:** Response to amitriptyline or cyclobenzaprine did not differ based on age.
	Others	Insufficient	**Sex:** Efficacy findings (not specified) for cyclobenzaprine were not influenced by sex. However, effect of duloxetine on pain was no longer significant in males. **Race:** Race did not influence efficacy for cyclobenzaprine, but pain reduction with duloxetine was significant in white but not nonwhite patients based on a small sample size. **Comorbidities:** Compared with placebo, duloxetine, fluoxetine, controlled-release paroxetine, and pregabalin significantly improved fibromyalgia symptoms regardless of baseline depression. Controlled-release paroxetine and pregabalin significantly improved fibromyalgia symptoms regardless of baseline anxiety.

Abbreviations: AE, adverse event; FIQ, Fibromyalgia Impact Questionnaire total score; HRQOL, health-related quality of life; PGII, Patient Global Impression of Improvement; PGIC, Patient Global Impression of Change.

CONCLUSIONS

We found eligible studies of treatment for fibromyalgia with amitriptyline, nortriptyline, citalopram, fluoxetine, paroxetine, cyclobenzaprine, pregabalin, gabapentin, milnacipran, and duloxetine. We found no eligible studies with the other included drugs and no eligible studies of included interventions when used as adjunctive therapy. Head-to-head trials were few, and provided low-strength evidence that short-term treatment with immediate-release paroxetine is superior to amitriptyline in reducing pain and sleep problems and provided low-strength evidence there are no significant differences between amitriptyline as compared to cyclobenzaprine and nortriptyline. Although there were some significant differences between drugs in overall adverse events, they did not produce any differences in withdrawals due to adverse events. Additionally, based on indirect comparison meta-analysis, we found low evidence that duloxetine was superior to milnacipran on outcomes of pain, sleep disturbance, depressed mood, and health-related quality of life. We found low evidence that both duloxetine and milnacipran were superior to pregabalin on improvement in depressed mood, whereas pregabalin was superior to milnacipran on improvement in sleep disturbance. Amitriptyline was similar to duloxetine, milnacipran, and pregabalin on outcomes of pain and fatigue with insufficient data on the other outcomes. Although there were some significant differences between duloxetine, milnacipran, and pregabalin in specific adverse events, they did not produce any differences in overall withdrawals, overall adverse events, and withdrawals due to adverse events. Amitriptyline was no different than duloxetine, milnacipran, and pregabalin in overall withdrawals with insufficient evidence to report on comparative overall adverse events and

withdrawals due to adverse events. For the remaining drugs, there was only evidence of significant improvements in pain over placebo in 1 trial for gabapentin, 1 of 3 trials for cyclobenzaprine, and in 1 trial of fluoxetine. But, no conclusions can be drawn about comparative effectiveness or harms among these drugs because the numbers of trials/patients in placebo-controlled trials were too few to provide meaningful results in indirect comparisons. There was a small body of evidence suggesting that duloxetine was not effective on pain reduction in male, nonwhite, and older patients based on a small sample size that was underpowered to detect a difference. Compared with placebo, duloxetine, fluoxetine, controlled-release paroxetine, and pregabalin significantly improved fibromyalgia symptoms regardless of baseline depression but a significant milnacipran effect compared with placebo was only observed in nondepressed patients. Controlled-release paroxetine and pregabalin significantly improved fibromyalgia symptoms regardless of baseline anxiety.

REFERENCES

1. Liberati A, Altman D, Tetzlaff J, Mulrow C, al. e. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med.* 2009;151(4):W65-W94.
2. Wolfe F, Clauw DJ, Fitzcharles M-A, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care Res.* May 2010;62(5):600-610.
3. Wolfe F. New American College of Rheumatology criteria for fibromyalgia: a twenty-year journey. *Arthritis Care Res.* 2010;62(5):583-584.
4. Lawrence R, Felson D, Helmick C, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum.* 2008;58(1):26-35.
5. White KP, Speechley M, Harth M, Ostbye T. The London Fibromyalgia Epidemiology Study: the prevalence of fibromyalgia syndrome in London, Ontario. *J Rheumatol.* 1999;26:1570-1576.
6. Weir P, Harlan G, Nkoy F, et al. The incidence of fibromyalgia and its associated comorbidities: a population-based retrospective cohort study based on International Classification of Diseases, 9th Revision codes. *Journal of Clinical Rheumatology.* 2006;12(3):124-128.
7. Burgmer M, Pogatzki-Zahn E, Gaubitz M, Wessoleck E, Heuft G, Pfleiderer B. Altered brain activity during pain processing in fibromyalgia. *Neuroimage.* 2009;44(2):502-508.
8. Staud R. Biology and therapy of fibromyalgia: pain in fibromyalgia syndrome. *Arthritis Research & Therapy.* 2006;8(3):208.
9. Verdu B, Decosterd I, Buclin T, Stiefel F, Berney A. Antidepressants for the treatment of chronic pain. *Drugs.* 2008;68(18):2611-2632.
10. Abeles AM, Pillinger MH, Solitar BM, Abeles M. Narrative review: the pathophysiology of fibromyalgia. *Ann Intern Med.* May 15 2007;146(10):726-734.
11. Arnold LM. The pathophysiology, diagnosis and treatment of fibromyalgia. *Psychiatr Clin North Am.* Jun 2010;33(2):375-408.
12. Clauw DJ. Fibromyalgia: An Overview. *Am J Med.* 2009;122(12):S3-S13.
13. Burckhardt CS, Goldenberg D, Crofford L, et al. Guideline for the Management of Fibromyalgia Syndrome Pain in Adults and Children *APS Clinical Practice Guidelines Series* 2005:1-107.
14. Bennett RM, Clark SC, Walczyk J. A randomized, double-blind, placebo-controlled study of growth hormone in the treatment of fibromyalgia. *Am J Med.* Mar 1998;104(3):227-231.
15. Cuatrecasas G, Riudavets C, Guell MA, Nadal A. Growth hormone as concomitant treatment in severe fibromyalgia associated with low IGF-1 serum levels. A pilot study. *BMC Musculoskeletal Disorders.* 2007;8:119.
16. Hannonen P, Malminiemi K, Yli-Kerttula U, Isomeri R, Roponen P. A randomized double-blind placebo controlled study of moclobemide and amitriptyline in fibromyalgia. *Scand-J-Rheumatol.* 1998;27(3):246.
17. Yavuzer G, Kucukdeveci A, Arasil T, Elhan A. Moclobemid treatment in primary fibromyalgia syndrome. *European Journal of Physical Medicine and Rehabilitation.* 1998;8(2):35-38.

18. Skrabek RQ, Galimova L, Ethans K, Perry D. Nabilone for the treatment of pain in fibromyalgia. *Journal of Pain.* Feb 2008;9(2):164-173.
19. Ware MA, Fitzcharles M-A, Joseph L, Shir Y. The effects of nabilone on sleep in fibromyalgia: results of a randomized controlled trial. *Anesth Analg.* Feb 2010;110(2):604-610.
20. Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: a pilot study. *Pain Medicine.* May-Jun 2009;10(4):663-672.
21. Sadreddini S, Molaeefard M, Noshad H, Ardalan M, Asadi A. Efficacy of Raloxifene in treatment of fibromyalgia in menopausal women. *European Journal of Internal Medicine.* Jul 2008;19(5):350-355.
22. Olin R, Klein R, Berg PA. A randomised double-blind 16-week study of ritanserin in fibromyalgia syndrome: clinical outcome and analysis of autoantibodies to serotonin, gangliosides and phospholipids. *Clin Rheumatol.* 1998;17(2):89-94.
23. Russell IJ, Perkins AT, Michalek JE, Oxybate SXBFSSG. Sodium oxybate relieves pain and improves function in fibromyalgia syndrome: a randomized, double-blind, placebo-controlled, multicenter clinical trial. *Arthritis Rheum.* Jan 2009;60(1):299-309.
24. Scharf MB, Baumann M, Berkowitz DV. The effects of sodium oxybate on clinical symptoms and sleep patterns in patients with fibromyalgia. *J Rheumatol.* May 2003;30(5):1070-1074.
25. Distler O, Eich W, Dokoupilova E, et al. Evaluation of the efficacy and safety of terguride in patients with fibromyalgia syndrome: results of a twelve-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheum.* Jan 2010;62(1):291-300.
26. Bennett RM, Kamin M, Karim R, Rosenthal N. Tramadol and acetaminophen combination tablets in the treatment of fibromyalgia pain: a double-blind, randomized, placebo-controlled study. *Am J Med.* May 2003;114(7):537-545.
27. Bennett RM, Schein J, Kosinski MR, Hewitt DJ, Jordan DM, Rosenthal NR. Impact of fibromyalgia pain on health-related quality of life before and after treatment with tramadol/acetaminophen. *Arthritis Rheum.* Aug 15 2005;53(4):519-527.
28. Russell J, Kamin M, Bennett RM, Schnitzer TJ, Green JA, Katz WA. Efficacy of tramadol in treatment of pain in fibromyalgia. *JCR: Journal of Clinical Rheumatology.* 2000;6(5):250-257.
29. Moldofsky H, Lue FA, Mously C, Roth-Schechter B, Reynolds WJ. The effect of zolpidem in patients with fibromyalgia: a dose ranging, double blind, placebo controlled, modified crossover study. *J Rheumatol.* Mar 1996;23(3):529-533.
30. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum.* Feb 1990;33(2):160-172.
31. Center for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. CRD Report Number 4 (2nd ed.)*. York, UK: NHS Centre for Reviews and Dissemination; 2001.
32. 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.* Apr 2001;20(3 Suppl):21-35.

33. Owens DK, Lohr KN, Atkins D. Grading the strength of a body of evidence when comparing medical interventions. *Agency for Healthcare Research and Quality. Methods Guide for Comparative Effectiveness Reviews*. Rockville, MD; 2009.
34. Owens DK, Lohr KN, Atkins D, et al. Grading the strength of a body of evidence when comparing medical interventions-Agency for Healthcare Research and Quality and the Effective Health Care Program. *J Clin Epidemiol.* 2010;63(5):513-523.
35. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clinical Trials.* 1986;7(3):177-188.
36. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.
37. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539--1558.
38. Bennett RM, Gatter RA, Campbell SM, Andrews RP, Clark SR, Scarola JA. A comparison of cyclobenzaprine and placebo in the management of fibrositis. A double-blind controlled study. *Arthritis Rheum.* Dec 1988;31(12):1535-1542.
39. Carette S, McCain GA, Bell DA, Fam AG. Evaluation of amitriptyline in primary fibrositis. A double-blind, placebo-controlled study. *Arthritis Rheum.* May 1986;29(5):655-659.
40. Quimby LG, Gratwick GM, Whitney CD, Block SR. A randomized trial of cyclobenzaprine for the treatment of fibromyalgia. *Journal of Rheumatology - Supplement.* Nov 1989;19:140-143.
41. Reynolds WJ, Moldofsky H, Saskin P, Lue FA. The effects of cyclobenzaprine on sleep physiology and symptoms in patients with fibromyalgia. *J Rheumatol.* Mar 1991;18(3):452-454.
42. Scudds RA, McCain GA, Rollman GB, Harth M. Improvements in pain responsiveness in patients with fibrositis after successful treatment with amitriptyline. *Journal of Rheumatology - Supplement.* Nov 1989;19:98-103.
43. Carette S, Bell MJ, Reynolds WJ, et al. Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomized, double-blind clinical trial. *Arthritis Rheum.* Jan 1994;37(1):32-40.
44. Goldenberg D, Mayskiy M, Mossey C, Ruthazer R, Schmid C. A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia. *Arthritis Rheum.* Nov 1996;39(11):1852-1859.
45. Heymann RE, Helfenstein M, Feldman D. A double-blind, randomized, controlled study of amitriptyline, nortriptyline and placebo in patients with fibromyalgia. An analysis of outcome measures. *Clin Exp Rheumatol.* Nov-Dec 2001;19(6):697-702.
46. Ataoglu S, Ataoglu A, Erdogan F, Sarac J. Comparison of paroxetine, amitriptyline in the treatment of fibromyalgia. *Turk J Med Sci.* 1997;27:535-539.
47. Carville SF, Arendt-Nielsen S, Bliddal H, et al. EULAR evidence-based recommendations for the management of fibromyalgia syndrome. *Ann Rheum Dis.* Apr 2008;67(4):536-541.
48. Hauser W, Bernardy K, Uceyler N, Sommer C. Treatment of fibromyalgia syndrome with gabapentin and pregabalin: a meta-analysis of randomized controlled trials (Provisional abstract). *Pain.* 2009;145(1-2):69-81.
49. Hauser W, Petzke F, Sommer C. Comparative Efficacy and Harms of Duloxetine, Milnacipran, and Pregabalin in Fibromyalgia Syndrome. *J. Pain.* 2010;11(6):505-521.

50. Nishishinya B, Urrutia G, Walitt B, et al. Amitriptyline in the treatment of fibromyalgia: a systematic review of its efficacy. *Rheumatology (Oxford).* Dec 2008;47(12):1741-1746.
51. Uceyler N, Hauser W, Sommer C. A systematic review on the effectiveness of treatment with antidepressants in fibromyalgia syndrome. *Arthritis Rheum.* Sep 15 2008;59(9):1279-1298.
52. Arnold L, Gendreau R, Palmer R, Gendreau J, Wang Y. Efficacy and safety of milnacipran 100 mg/day in patients with fibromyalgia: results of a randomized, double-blind, placebo-controlled trial. . *Arthritis Rheum.* 2010;62(9):2745-2756.
53. Branco JC, Zachrisson O, Perrot S, Mainguy Y, Multinational Coordinator Study G. A European multicenter randomized double-blind placebo-controlled monotherapy clinical trial of milnacipran in treatment of fibromyalgia. *J Rheumatol.* Apr 2010;37(4):851-859.
54. Carette S, Oakson G, Guimont C, Steriade M. Sleep electroencephalography and the clinical response to amitriptyline in patients with fibromyalgia. *Arthritis Rheum.* Sep 1995;38(9):1211-1217.
55. Ginsberg F, Mancaux A, Joos E, Vanhove P, Famaey JP. A randomized placebo-controlled trial of sustained-release amitriptyline in primary fibromyalgia. *Journal of Musculoskeletal Pain.* 1996;4(3):37-47.
56. Hannonen P, Malminiemi K, Yli-Kerttula U, Isomeri R, Roponen P. A randomized, double-blind, placebo-controlled study of moclobemide and amitriptyline in the treatment of fibromyalgia in females without psychiatric disorder. *Br J Rheumatol.* Dec 1998;37(12):1279-1286.
57. Arnold LM, Russell IJ, Diri EW, et al. A 14-week, randomized, double-blinded, placebo-controlled monotherapy trial of pregabalin in patients with fibromyalgia. *Journal of Pain.* Sep 2008;9(9):792-805.
58. Crofford LJ, Rowbotham MC, Mease PJ, et al. Pregabalin for the treatment of fibromyalgia syndrome: results of a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* Apr 2005;52(4):1264-1273.
59. Mease PJ, Russell IJ, Arnold LM, et al. A randomized, double-blind, placebo-controlled, phase III trial of pregabalin in the treatment of patients with fibromyalgia. *J Rheumatol.* Mar 2008;35(3):502-514.
60. Pfizer. A 14-Week, Randomized, Double-Blind, Placebo-Controlled Trial of Pregabalin Twice Daily in Patients with Fibromyalgia. *Protocol A0081100.* 2008;PhRMA Web Synopsis: Lyrica/Pregabalin (Unpublished Study).
61. Clauw DJ, Mease P, Palmer RH, Gendreau RM, Wang Y. Milnacipran for the treatment of fibromyalgia in adults: a 15-week, multicenter, randomized, double-blind, placebo-controlled, multiple-dose clinical trial.[Erratum appears in Clin Ther. 2009 Jul;31(7):1617].[Erratum appears in Clin Ther. 2009 Feb;31(2):446]. *Clin Ther.* Nov 2008;30(11):1988-2004.
62. Gendreau RM, Thorn MD, Gendreau JF, et al. Efficacy of milnacipran in patients with fibromyalgia. *J Rheumatol.* Oct 2005;32(10):1975-1985.
63. Mease PJ, Clauw DJ, Gendreau RM, et al. The efficacy and safety of milnacipran for treatment of fibromyalgia. a randomized, double-blind, placebo-controlled trial.[Erratum appears in J Rheumatol. 2009 Mar;36(3):661]. *J Rheumatol.* Feb 2009;36(2):398-409.
64. Vitton O, Gendreau M, Gendreau J, Kranzler J, Rao SG. A double-blind placebo-controlled trial of milnacipran in the treatment of fibromyalgia. *Human Psychopharmacology.* Oct 2004;19 Suppl 1:S27-35.

65. Arnold LM, Lu Y, Crofford LJ, et al. A double-blind, multicenter trial comparing duloxetine with placebo in the treatment of fibromyalgia patients with or without major depressive disorder. *Arthritis Rheum.* Sep 2004;50(9):2974-2984.
66. Arnold LM, Rosen A, Pritchett YL, et al. A randomized, double-blind, placebo-controlled trial of duloxetine in the treatment of women with fibromyalgia with or without major depressive disorder. *Pain.* Dec 15 2005;119(1-3):5-15.
67. Chappell AS, Bradley LA, Wiltse C, Detke MJ, D'Souza DN, Spaeth M. A six-month double blind placebo-controlled, randomized clinical trial of duloxetine for the treatment of fibromyalgia. *international Journal of General Medicine.* 2008;1:91-102.
68. Hunter AM, Leuchter AF, Cook IA, et al. Brain functional changes and duloxetine treatment response in fibromyalgia: a pilot study. *Pain Medicine.* May-Jun 2009;10(4):730-738.
69. Russell IJ, Mease PJ, Smith TR, et al. Efficacy and safety of duloxetine for treatment of fibromyalgia in patients with or without major depressive disorder: Results from a 6-month, randomized, double-blind, placebo-controlled, fixed-dose trial. *Pain.* Jun 2008;136(3):432-444.
70. Straub TA. Endoscopic carpal tunnel release: a prospective analysis of factors associated with unsatisfactory results. *Arthroscopy.* Apr 1999;15(3):269-274.
71. Crofford LJ, Mease PJ, Simpson SL, et al. Fibromyalgia relapse evaluation and efficacy for durability of meaningful relief (FREEDOM): a 6-month, double-blind, placebo-controlled trial with pregabalin. *Pain.* Jun 2008;136(3):419-431.
72. Moore RA, Straube S, Wiffen PJ, Derry S, McQuay HJ. Pregabalin for acute and chronic pain in adults [Systematic Review]. *Cochrane Database of Systematic Reviews.* 2009;1.
73. Goldenberg DL, Felson DT, Dinerman H. A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia. *Arthritis Rheum.* Nov 1986;29(11):1371-1377.
74. Straube S, Derry S, Moore RA, McQuay HJ. Pregabalin in fibromyalgia: meta-analysis of efficacy and safety from company clinical trial reports. *Rheumatology (Oxford).* 2010;49:706–715.
75. Arnold L, Clauw D, Wohlreich M, et al. Efficacy of duloxetine in patients with fibromyalgia: pooled analysis of four placebo-controlled clinical trials. *Prim Care Companion J Clin Psychiatry.* 2009;11(5):237-244.
76. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. *Clin Exp Rheumatol.* 2005;23(Suppl. 39):S154-S162.
77. Bennett RM, Bushmakin AG, Cappelleri JC, Zlateva G, Sadosky AB. Minimal clinically important difference in the fibromyalgia impact questionnaire. *J Rheumatol.* 2009;36(6):1304-1311.
78. Arnold LM, Goldenberg DL, Stanford SB, et al. Gabapentin in the treatment of fibromyalgia: a randomized, double-blind, placebo-controlled, multicenter trial. *Arthritis Rheum.* Apr 2007;56(4):1336-1344.
79. Anderberg UM, Marteinsdottir I, von Knorring L. Citalopram in patients with fibromyalgia--a randomized, double-blind, placebo-controlled study. *European Journal of Pain: Ejp.* 2000;4(1):27-35.
80. Norregaard J, Volkmann H, Danneskiold-Samsoe B. A randomized controlled trial of citalopram in the treatment of fibromyalgia. *Pain.* Jun 1995;61(3):445-449.

81. Arnold LM, Hess EV, Hudson JI, Welge JA, Berno SE, Keck PE, Jr. A randomized, placebo-controlled, double-blind, flexible-dose study of fluoxetine in the treatment of women with fibromyalgia. *Am J Med.* Feb 15 2002;112(3):191-197.
82. Wolfe F, Cathey MA, Hawley DJ. A double-blind placebo controlled trial of fluoxetine in fibromyalgia. *Scand J Rheumatol.* 1994;23(5):255-259.
83. Giordano N, Geraci S, Santacroce C, Mattii G, Battisti E, Gennari C. Efficacy and tolerability of paroxetine in patients with fibromyalgia syndrome: A single-blind study. *Curr Ther Res Clin Exp.* 1999;60(12):696-702.
84. Patkar AA, Masand PS, Krulewicz S, et al. A randomized, controlled, trial of controlled release paroxetine in fibromyalgia. *Am J Med.* May 2007;120(5):448-454.
85. Hauser W, Thieme K, Turk DC. Guidelines on the management of fibromyalgia syndrome - a systematic review. *European Journal of Pain: Ejp.* Jan 2010;14(1):5-10.
86. Arnold BS, Alpers GW, Suss H, et al. Affective pain modulation in fibromyalgia, somatoform pain disorder, back pain, and healthy controls. *European Journal of Pain: Ejp.* Apr 2008;12(3):329-338.
87. Arnold LM, Crofford LJ, Martin SA, Young JP, Sharma U. The effect of anxiety and depression on improvements in pain in a randomized, controlled trial of pregabalin for treatment of fibromyalgia. *Pain Medicine.* Nov-Dec 2007;8(8):633-638.
88. Pae C-U, Masand PS, Marks DM, et al. History of depressive and/or anxiety disorders as a predictor of treatment response: a post hoc analysis of a 12-week, randomized, double-blind, placebo-controlled trial of paroxetine controlled release in patients with fibromyalgia. *Prog Neuropsychopharmacol Biol Psychiatry.* Aug 31 2009;33(6):996-1002.
89. Emir B, Murphy TK, Petersel DL, Whalen E. Treatment response to pregabalin in fibromyalgia pain: effect of patient baseline characteristics. *Expert Opin. Pharmacother.* 2010;11(14):2275-2280.

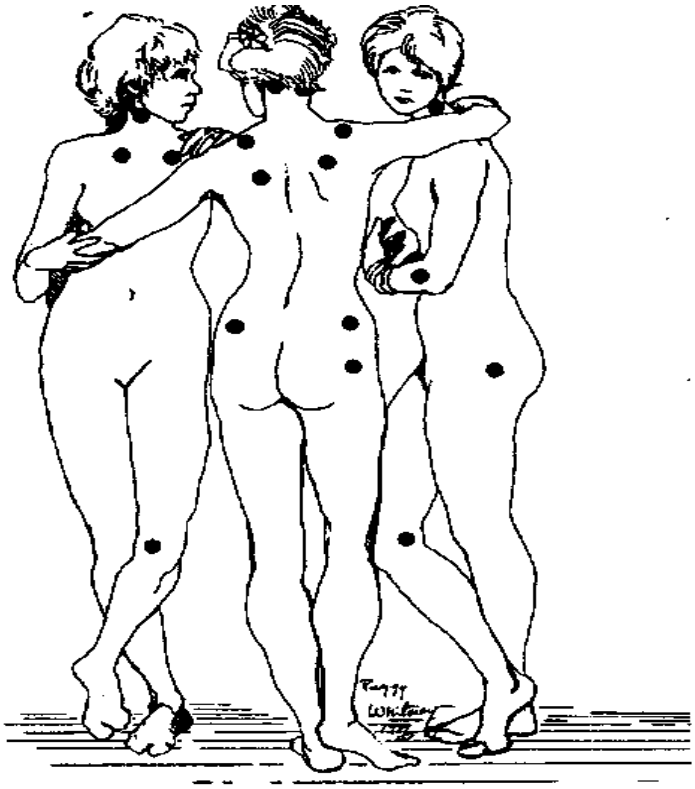
Appendix A. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia[a1]

1. History of widespread pain
   Definition. Pain is considered widespread when all the following are present: pain in the left side of the body, pain in the right side of the body, pain above the waist, and pain below the waist. In addition, axial skeletal pain, (cervical spine or anterior chest or thoracic spine or low back) must be present. In this definition, shoulder and buttock pain is considered as pain for each involved side. "Low back" pain is considered lower segment pain.

2. Pain in 11 of 18 tender point sites on digital palpitation
   Definition. Pain on digital palpitation, must be present in at least 11 of the following 18 tender point sites:
   Occiput: bilateral, at the suboccipital muscle insertions.
   Low cervical: bilateral, at the anterior aspects of the intertransverse spaces C5-C7.
   Trapezius: bilateral, at the midpoint of the upper border.
   Supraspinatus: bilateral, at origins, above the scapula spine near the medial border.
   Second rib: bilateral, at the second costochondral junctions, just lateral to the junctions of upper surfaces.
   Lateral epicondyle: bilateral, 2 cm distal to the epicondyles.
   Gluteal: bilateral, in upper outer quadrants of buttocks in anterior fold of muscle.
   Greater trochanter: bilateral, posterior to the trochanteric prominence.
   Knee: bilateral, at the medial fat pad proximal to the joint line

   Digital palpation should be performed with an approximate force of 4 kg.
   For a tender point to be considered "positive" the subject must state that the palpation was "painful".

[a] For the classification purposes, patients will be said to have fibromyalgia if both criteria are satisfied. Widespread pain must have been present for at least 3 months. The presence of a second clinical disorder does not exclude the diagnosis of fibromyalgia.



The figure specifies tender point locations for the 1990 classification criteria for fibromyalgia (The Three Graces after Baron Jean-Baptiste Regnault, 1793, Louvre Museum, Paris.[1]

Reference:

1. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum. Feb 1990;33(2):160-172.

Appendix B. Glossary

This glossary defines terms as they are used in reports produced by the Drug Effectiveness Review Project. Some definitions may vary slightly from other published definitions.

Absolute risk: The probability or chance that a person will have a medical event. Absolute risk is expressed as a percentage. It is the ratio of the number of people who have a medical event divided by all of the people who could have the event because of their medical condition.

Add-on therapy: An additional treatment used in conjunction with the primary or initial treatment.

Adherence: Following the course of treatment proscribed by a study protocol.

Adverse drug reaction: An adverse effect specifically associated with a drug.

Adverse event: A harmful or undesirable outcome that occurs during or after the use of a drug or intervention but is not necessarily caused by it.

Adverse effect: An **adverse event** for which the causal relation between the intervention and the event is at least a reasonable possibility.

Active-control trial: A trial comparing a drug in a particular class or group with a drug outside of that class or group.

Allocation concealment: The process by which the person determining randomization is blinded to a study participant's group allocation.

Applicability: see *External Validity*

Before-after study: A type nonrandomized study where data are collected before and after patients receive an intervention. Before-after studies can have a single arm or can include a control group.

Bias: A systematic error or deviation in results or inferences from the truth. Several types of bias can appear in published trials, including selection bias, performance bias, detection bias, and reporting bias.

Bioequivalence: Drug products that contain the same compound in the same amount that meet current official standards, that, when administered to the same person in the same dosage regimen result in equivalent concentrations of drug in blood and tissue.

Black box warning: A type of warning that appears on the package insert for prescription drugs that may cause serious adverse effects. It is so named for the black border that usually surrounds the text of the warning. A black box warning means that medical studies indicate that the drug carries a significant risk of serious or even life-threatening adverse effects. The US Food and Drug Administration (FDA) can require a pharmaceutical company to place a black box warning on the labeling of a prescription drug, or in literature describing it. It is the strongest warning that the FDA requires.

Blinding: A way of making sure that the people involved in a research study — participants, clinicians, or researchers —do not know which participants are assigned to each study group. Blinding usually is used in research studies that compare two or more types of treatment for an illness. Blinding is used to make sure that knowing the type of treatment does not affect a participant's response to the treatment, a health care provider's behavior, or assessment of the treatment effects.

Case series: A study reporting observations on a series of patients receiving the same intervention with no control group.

Case study: A study reporting observations on a single patient.

Case-control study: A study that compares people with a specific disease or outcome of interest (cases) to people from the same population without that disease or outcome (controls).

Clinical diversity: Differences between studies in key characteristics of the participants, interventions or outcome measures.

Clinically significant: A result that is large enough to affect a patient's disease state in a manner that is noticeable to the patient and/or a caregiver.

Cohort study: An observational study in which a defined group of people (the cohort) is followed over time and compared with a group of people who were exposed or not exposed to a particular intervention or other factor of interest. A prospective cohort study assembles participants and follows them into the future. A retrospective cohort study identifies subjects from past records and follows them from the time of those records to the present.

Combination Therapy: The use of two or more therapies and especially drugs to treat a disease or condition.

Confidence interval: The range of values calculated from the data such that there is a level of confidence, or certainty, that it contains the true value. The 95% confidence interval is generally used in Drug Effectiveness Review Project reports. If the report were hypothetically repeated on a collection of 100 random samples of studies, the resulting 95% confidence intervals would include the true population value 95% of the time.

Confounder: A factor that is associated with both an intervention and an outcome of interest.

Controlled clinical trial: A clinical trial that includes a control group but no or inadequate methods of randomization.

Control group: In a research study, the group of people who do not receive the treatment being tested. The control group might receive a placebo, a different treatment for the disease, or no treatment at all.

Convenience sample: A group of individuals being studied because they are conveniently accessible in some way. Convenience samples may or may not be representative of a population that would normally be receiving an intervention.

Crossover trial: A type of clinical trial comparing two or more interventions in which the participants, upon completion of the course of one treatment, are switched to another.

Direct analysis: The practice of using data from head-to-head trials to draw conclusions about the comparative effectiveness of drugs within a class or group. Results of direct analysis are the preferred source of data in Drug Effectiveness Review Project reports.

Dosage form: The physical form of a dose of medication, such as a capsule, injection, or liquid. The route of administration is dependent on the dosage form of a given drug. Various dosage forms may exist for the same compound, since different medical conditions may warrant different routes of administration.

Dose-response relationship: The relationship between the quantity of treatment given and its effect on outcome. In meta-analysis, dose-response relationships can be investigated using meta-regression.

Double-blind: The process of preventing those involved in a trial from knowing to which comparison group a particular participant belongs. While double-blind is a frequently used term

in trials, its meaning can vary to include blinding of patients, caregivers, investigators, or other study staff.

Double-dummy: The use of two placebos in a trial that match the active interventions when they vary in appearance or method of administrations (for example, when an oral agent is compared with an injectable agent).

Effectiveness: The extent to which a specific intervention *used under ordinary circumstances* does what it is intended to do.

Effectiveness outcomes: Outcomes that are generally important to patients and caregivers, such as quality of life, responder rates, number and length of hospitalizations, and ability to work. Data on effectiveness outcomes usually comes from longer-term studies of a "real-world" population.

Effect size/estimate of effect: The amount of change in a condition or symptom because of a treatment (compared to not receiving the treatment). It is commonly expressed as a risk ratio (relative risk), odds ratio, or difference in risk.

Efficacy: The extent to which an intervention produces a beneficial result *under ideal conditions* in a selected and controlled population.

Equivalence level: The amount which an outcome from two treatments can differ but still be considered equivalent, as in an equivalence trial, or the amount which an outcome from treatment A can be worse than that of treatment B but still be considered noninferior, as in a noninferiority trial.

Equivalence trial: A trial designed to determine whether the response to two or more treatments differs by an amount that is clinically unimportant. This lack of clinical importance is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence level of clinically acceptable differences.

Exclusion criteria: The criteria, or standards, set out before a study or review. Exclusion criteria are used to determine whether a person should participate in a research study or whether an individual study should be excluded in a systematic review. Exclusion criteria may include age, previous treatments, and other medical conditions. Criteria help identify suitable participants.

External validity: The extent to which results provide a correct basis for generalizations to other circumstances. For instance, a meta-analysis of trials of elderly patients may not be generalizable to children. (Also called generalizability or applicability.)

Fixed-effect model: A model that calculates a pooled estimate using the assumption that all observed variation between studies is due to by chance. Studies are assumed to be measuring the same overall effect. An alternative model is the random-effects model.

Fixed-dose combination product: A formulation of two or more active ingredients combined in a single dosage form available in certain fixed doses.

Forest plot: A graphical representation of the individual results of each study included in a meta-analysis and the combined result of the meta-analysis. The plot allows viewers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centered on each study's point estimate. A horizontal line runs through each square to show each study's confidence interval—usually, but not always, a 95% confidence interval. The overall estimate from the meta-analysis and its confidence interval are represented as a diamond. The center of the diamond is at the pooled point estimate, and its horizontal tips show the confidence interval.

Funnel plot: A graphical display of some measure of study precision plotted against effect size that can be used to investigate whether there is a link between study size and treatment effect.

Generalizability: See *External Validity.*

Half- life: The time it takes for the plasma concentration or the amount of drug in the body to be reduced by 50%.

Harms: See *Adverse Event*

Hazard ratio: The increased risk with which one group is likely to experience an outcome of interest. It is similar to a risk ratio. For example, if the hazard ratio for death for a treatment is 0.5, then treated patients are likely to die at half the rate of untreated patients.

Head-to-head trial: A trial that directly compares one drug in a particular class or group with another in the same class or group.

Health outcome: The result of a particular health care practice or intervention, including the ability to function and feelings of well-being. For individuals with chronic conditions – where cure is not always possible – results include health-related quality of life as well as mortality.

Heterogeneity: The variation in, or diversity of, participants, interventions, and measurement of outcomes across a set of studies.

I^2: A measure of statistical heterogeneity of the estimates of effect from studies. Values range from 0% to 100%. Large values of I^2 suggest heterogeneity. I^2 is the proportion of total variability across studies that is due to heterogeneity and not chance. It is calculated as (Q-(n-1))/Q, where n is the number of studies.

Incidence: The number of new occurrences of something in a population over a particular period of time, e.g. the number of cases of a disease in a country over one year.

Indication: A term describing a valid reason to use a certain test, medication, procedure, or surgery. In the United States, indications for medications are strictly regulated by the Food and Drug Administration, which includes them in the package insert under the phrase "Indications and Usage".

Indirect analysis: The practice of using data from trials comparing one drug in a particular class or group with another drug outside of that class or group or with placebo and attempting to draw conclusions about the comparative effectiveness of drugs within a class or group based on that data. For example, direct comparisons between drugs A and B and between drugs B and C can be used to make an indirect comparison between drugs A and C.

Intention to treat: The use of data from a randomized controlled trial in which data from all randomized patients are accounted for in the final results. Trials often incorrectly report results as being based on intention to treat despite the fact that some patients are excluded from the analysis.

Internal validity: The extent to which the design and conduct of a study are likely to have prevented bias. Generally, the higher the interval validity, the better the quality of the study publication.

Inter-rater reliability: The degree of stability exhibited when a measurement is repeated under identical conditions by different raters.

Intermediate outcome: An outcome not of direct practical importance but believed to reflect outcomes that are important. For example, blood pressure is not directly important to patients but it is often used as an outcome in clinical trials because it is a risk factor for stroke and myocardial infarction (hear attack).

Logistic regression: A form of regression analysis that models an individual's odds of disease or some other outcome as a function of a risk factor or intervention.

Masking: See *Blinding*

Mean difference: A method used to combine measures on continuous scales (such as weight) where the mean, standard deviation, and sample size are known for each group.

Meta-analysis: The use of statistical techniques in a systematic review to integrate the results of included studies. Although the terms are sometimes used interchangeably, meta-analysis is not synonymous with systematic review. However, systematic reviews often include meta-analyses.

Meta-regression: A technique used to explore the relationship between study characteristics (for example, baseline risk, concealment of allocation, timing of the intervention) and study results (the magnitude of effect observed in each study) in a systematic review.

Mixed treatment comparison meta analysis: A meta-analytic technique that simultaneously compares multiple treatments (typical 3 or more) using both direct and indirect evidence. The multiple treatments form a network of treatment comparisons. Also called multiple treatment comparisons, network analysis, or umbrella reviews.

Monotherapy: the use of a single drug to treat a particular disorder or disease.

Multivariate analysis: Measuring the impact of more than one variable at a time while analyzing a set of data.

N-of-1 trial: A randomized trial in an individual to determine the optimum treatment for that individual.

Noninferiority trial: A trial designed to determine whether the effect of a new treatment is not worse than a standard treatment by more than a prespecified amount. A one-sided version of an equivalence trial.

Nonrandomized study: Any study estimating the effectiveness (harm or benefit) of an intervention that does not use randomization to allocate patients to comparison groups. There are many types of nonrandomized studies, including cohort studies, case-control studies, and before-after studies.

Null hypothesis: The statistical hypothesis that one variable (for example, treatment to which a participant was allocated) has no association with another variable or set of variables.

Number needed to harm: The number of people who would need to be treated over a specific period of time before one bad outcome of the treatment will occur. The number needed to harm (NNH) for a treatment can be known only if clinical trials of the treatment have been performed.

Number needed to treat: An estimate of how many persons need to receive a treatment before one person would experience a beneficial outcome.

Observational study: A type of nonrandomized study in which the investigators do not seek to intervene, instead simply observing the course of events.

Odds ratio: The ratio of the odds of an event in one group to the odds of an event in another group. An odds ratio of 1.0 indicates no difference between comparison groups. For undesirable outcomes an odds ratio that is <1.0 indicates that the intervention was effective in reducing the risk of that outcome.

Off-label use: When a drug or device is prescribed outside its specific FDA-approved indication, to treat a condition or disease for which it is not specifically licensed.

Outcome: The result of care and treatment and/ or rehabilitation. In other words, the change in health, functional ability, symptoms or situation of a person, which can be used to measure the

effectiveness of care/treatment/rehabilitation. Researchers should decide what outcomes to measure before a study begins; outcomes are then assessed at the end of the study.

Outcome measure: Is the way in which an outcome is evaluated---the device (scale) used for measuring. With this definition YMRS is an outcome measure, and a patient's outcome after treatment might be a 12-point improvement on that scale.

One-tailed test (one-sided test): A hypothesis test in which the values that reject the null hypothesis are located entirely in one tail of the probability distribution. For example, testing whether one treatment is better than another (rather than testing whether one treatment is either better or worse than another).

Open-label trial: A clinical trial in which the investigator and participant are aware which intervention is being used for which participant (that is, not blinded). Random allocation may or may not be used in open-label trials.

Per protocol: The subset of participants from a randomized controlled trial who complied with the protocol sufficiently to ensure that their data would be likely to exhibit the effect of treatment. Per protocol analyses are sometimes misidentified in published trials as intention-to-treat analyses.

Pharmacokinetics: the characteristic interactions of a drug and the body in terms of its absorption, distribution, metabolism, and excretion.

Placebo: An inactive substance commonly called a "sugar pill." In a clinical trial, a placebo is designed to look like the drug being tested and is used as a control. It does not contain anything that could harm a person. It is not necessarily true that a placebo has no effect on the person taking it.

Placebo-controlled trial: A study in which the effect of a drug is compared with the effect of a placebo (an inactive substance designed to resemble the drug). In placebo-controlled clinical trials, participants receive either the drug being studied or a placebo. The results of the drug and placebo groups are then compared to see if the drug is more effective in treating the condition than the placebo is.

Point estimate: The results (e.g. mean, weighted difference, odds ratio, relative risk or risk difference) obtained in a sample (a study or a meta-analysis) which are used as the best estimate of what is true for the relevant population from which the sample is taken. A confidence interval is a measure of the uncertainty (due to the play of chance) associated with that estimate.

Pooling: The practice of combing data from several studies to draw conclusions about treatment effects.

Power: The probability that a trial will detect statistically significant differences among intervention effects. Studies with small sample sizes can frequently be underpowered to detect difference.

Precision: The likelihood of random errors in the results of a study, meta-analysis, or measurement. The greater the precision, the less the random error. Confidence intervals around the estimate of effect are one way of expressing precision, with a narrower confidence interval meaning more precision.

Prospective study: A study in which participants are identified according to current risk status or exposure and followed forward through time to observe outcome.

Prevalence: How often or how frequently a disease or condition occurs in a group of people. Prevalence is calculated by dividing the number of people who have the disease or condition by the total number of people in the group.

Probability: The likelihood (or chance) that an event will occur. In a clinical research study, it is the number of times a condition or event occurs in a study group divided by the number of people being studied.

Publication bias: A bias caused by only a subset of the relevant data being available. The publication of research can depend on the nature and direction of the study results. Studies in which an intervention is not found to be effective are sometimes not published. Because of this, systematic reviews that fail to include unpublished studies may overestimate the true effect of an intervention. In addition, a published report might present a biased set of results (for example, only outcomes or subgroups for which a statistically significant difference was found).

P value: The probability (ranging from zero to one) that the results observed in a study could have occurred by chance if the null hypothesis was true. A *P* value of ≤0.05 is often used as a threshold to indicate statistical significance.

Q-statistic: A measure of statistical heterogeneity of the estimates of effect from studies. Large values of Q suggest heterogeneity. It is calculated as the weighted sum of the squared difference of each estimate from the mean estimate.

Random-effects model: A statistical model in which both within-study sampling error (variance) and between-studies variation are included in the assessment of the uncertainty (confidence interval) of the results of a meta-analysis. When there is heterogeneity among the results of the included studies beyond chance, random-effects models will give wider confidence intervals than fixed-effect models.

Randomization: The process by which study participants are allocated to treatment groups in a trial. Adequate (that is, unbiased) methods of randomization include computer generated schedules and random-numbers tables.

Randomized controlled trial: A trial in which two or more interventions are compared through random allocation of participants.

Regression analysis: A statistical modeling technique used to estimate or predict the influence of one or more independent variables on a dependent variable, for example, the effect of age, sex, or confounding disease on the effectiveness of an intervention.

Relative risk: The ratio of risks in two groups; same as a risk ratio.

Retrospective study: A study in which the outcomes have occurred prior to study entry.

Risk: A way of expressing the chance that something will happen. It is a measure of the association between exposure to something and what happens (the outcome). Risk is the same as probability, but it usually is used to describe the probability of an adverse event. It is the rate of events (such as breast cancer) in the total population of people who could have the event (such as women of a certain age).

Risk difference: The difference in size of risk between two groups.

Risk Factor: A characteristic of a person that affects that person's chance of having a disease. A risk factor may be an inherent trait, such as gender or genetic make-up, or a factor under the person's control, such as using tobacco. A risk factor does not usually cause the disease. It changes a person's chance (or risk) of getting the disease.

Risk ratio: The ratio of risks in two groups. In intervention studies, it is the ratio of the risk in the intervention group to the risk in the control group. A risk ratio of 1 indicates no difference between comparison groups. For undesirable outcomes, a risk ratio that is <1 indicates that the intervention was effective in reducing the risk of that outcome.

Run-in period: Run in period: A period before randomization when participants are monitored but receive no treatment (or they sometimes all receive one of the study treatments, possibly in a blind fashion). The data from this stage of a trial are only occasionally of value but can serve a valuable role in screening out ineligible or non-compliant participants, in ensuring that participants are in a stable condition, and in providing baseline observations. A run-in period is sometimes called a washout period if treatments that participants were using before entering the trial are discontinued.

Safety: Substantive evidence of an absence of harm. This term (or the term ‘‘safe’’) should not be used when evidence on harms is simply absent or is insufficient.

Sample size: The number of people included in a study. In research reports, sample size is usually expressed as "n." In general, studies with larger sample sizes have a broader range of participants. This increases the chance that the study's findings apply to the general population. Larger sample sizes also increase the chance that rare events (such as adverse effects of drugs) will be detected.

Sensitivity analysis: An analysis used to determine how sensitive the results of a study or systematic review are to changes in how it was done. Sensitivity analyses are used to assess how robust the results are to uncertain decisions or assumptions about the data and the methods that were used.

Side effect: Any unintended effect of an intervention. Side effects are most commonly associated with pharmaceutical products, in which case they are related to the pharmacological properties of the drug at doses normally used for therapeutic purposes in humans.

Standard deviation (SD): A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample.

Standard error (SE): A measure of the variation in the sample statistic over all possible samples of the same size. The standard error decreases as the sample size increases.

Standard treatment: The treatment or procedure that is most commonly used to treat a disease or condition. In clinical trials, new or experimental treatments sometimes are compared to standard treatments to measure whether the new treatment is better.

Statistically significant: A result that is unlikely to have happened by chance.

Study: A research process in which information is recorded for a group of people. The information is known as data. The data are used to answer questions about a health care problem.

Study population: The group of people participating in a clinical research study. The study population often includes people with a particular problem or disease. It may also include people who have no known diseases.

Subgroup analysis: An analysis in which an intervention is evaluated in a defined subset of the participants in a trial, such as all females or adults older than 65 years.

Superiority trial: A trial designed to test whether one intervention is superior to another.

Surrogate outcome: Outcome measures that are not of direct practical importance but are believed to reflect outcomes that are important; for example, blood pressure is not directly important to patients but it is often used as an outcome in clinical trials because it is a risk factor for stroke and heart attacks. Surrogate endpoints are often physiological or biochemical markers that can be relatively quickly and easily measured, and that are taken as being predictive of important clinical outcomes. They are often used when observation of clinical outcomes requires long follow-up.

Survival analysis: Analysis of data that correspond to the time from a well-defined time origin until the occurrence of some particular event or end-point; same as time-to-event analysis.

Systematic review: A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research and to collect and analyze data from the studies that are included in the review.

Tolerability: For therapeutic drugs, it refers a drug's lack of "nuisance side effects," side effects that are thought to have no long-term effect but that are unpleasant enough to the patient that adherence to the medication regimen is affected.

The extent to which a drug's adverse effects impact the patient's ability or willingness to continue taking the drug as prescribed. These adverse effects are often referred to as nuisance side effects, because they are generally considered to not have long-term effects but can seriously impact compliance and adherence to a medication regimen.

Treatment regimen: The magnitude of effect of a treatment versus no treatment or placebo; similar to "effect size". Can be calculated in terms of relative risk (or risk ratio), odds ratio, or risk difference.

Two-tailed test (two-sided test): A hypothesis test in which the values that reject the null hypothesis are located in both tails of the probability distribution. For example, testing whether one treatment is different than another (rather than testing whether one treatment is either better than another).

Type I error: A conclusion that there is evidence that a treatment works, when it actually does not work (false-positive).

Type II error: A conclusion that there is no evidence that a treatment works, when it actually does work (false-negative).

Validity: The degree to which a result (of a measurement or study) is likely to be true and free of bias (systematic errors).

Variable: A measurable attribute that varies over time or between individuals. Variables can be

- *Discrete*: taking values from a finite set of possible values (e.g. race or ethnicity)
- *Ordinal*: taking values from a finite set of possible values where the values indicate rank (e.g. 5-point Likert scale)
- *Continuous:* taking values on a continuum (e.g. hemoglobin A1c values).

Washout period: [In a cross-over trial] The stage after the first treatment is withdrawn, but before the second treatment is started. The washout period aims to allow time for any active effects of the first treatment to wear off before the new one gets started.

Appendix C. Black box warnings

Drug names	Boxed Warnings
Black box warning that is issued for Norpramin® is listed in the right column. Other drugs such as Tofranil-PM®, Aventyl®, Pamelor®, Pristiq®, Effexor®, Effexor XR®, Celexa®, Luvox® CR, Paxil®, Paxil CR®, Pexeva®, Cymbalta®, Savella®, Remeron®, Remeron SolTab® have similar black box warnings. Luvox®, Zoloft®, Prozac®, Prozac® weekly™ and Lexapro have very specific warnings for children, but those are not listed here as the report does not include pediatric population. Wellbutrin®, Wellbutrin SR®, Wellbutrin XL® have additional boxed warnings for use in smoking cessation treatment, but that has not been specified here as well.	**Suicidality and Antidepressant Drugs** Antidepressants increased the risk compared to placebo of suicidal thinking and behavior (suicidality) in children, adolescents, and young adults in short-term studies of major depressive disorder (MDD) and other psychiatric disorders. Anyone considering the use of NORPRAMIN or any other antidepressant in a child, adolescent, or young adult must balance this risk with the clinical need. Short-term studies did not show an increase in the risk of suicidality with antidepressants compared to placebo in adults beyond age 24; there was a reduction in risk with antidepressants compared to placebo in adults aged 65 and older. Depression and certain other psychiatric disorders are themselves associated with increases in the risk of suicide. Patients of all ages who are started on antidepressant therapy should be monitored appropriately and observed closely for clinical worsening, suicidality, or unusual changes in behavior. Families and caregivers should be advised of the need for close observation and communication with the prescriber. NORPRAMIN is not approved for use in pediatric patients. (See WARNINGS: Clinical Worsening and Suicide Risk, PRECAUTIONS: Information for Patients, and PRECAUTIONS: Pediatric Use.)
Black box warning for Tegretol® is listed in the right column. Tegretol XR®, Carbatrol® and Equetro® have similar black box warnings.	**Serious dermatological reactions and HLA-B* 1502 Allele** Serious and sometimes fatal dermatologic reactions, including toxic epidermal necrolysis (Ten) and Stevens-Johnson syndrome (SJS), have been reported during treatment with Tegretol. These reactions are estimated to occur in 1 to 6 per 10,000 new users in countries with mainly Caucasian populations, but the risk in some Asian countries is estimated to be about 10 times higher. Studies in patients of Chinese ancestry have found a strong association between the risk of developing SJS/Ten and the presence of HLA-B*1502, an inherited allelic variant of the HLA-B gene*1502 is found almost exclusively in patients with ancestry across broad areas of Asia. Patients with ancestry in genetically at-risk populations should be screened for the presence of HLA-B*1502 prior to initiating treatment with Tegretol. Patients testing positive for the allele should not be treated with Tegretol unless the benefit carefully outweighs the risk. (See Warnings and Precautions, Laboratory Tests).

Drug names	Boxed Warnings
	Aplastic Anemia and agranulocytosis Aplastic anemia and agranulocytosis have been reported in association with the use of Tegretol. Data from a population-based case control study demonstrate that the risk of developing these reactions is 5-8 times greater than the general population. However, the overall risk of these reactions in the untreated general population is low, approximately six patients per one million population per year for agranulocytosis and two patients per one million population per year for aplastic anemia. Although reports of transient or persistent decreased platelet or white blood cell counts are not uncommon in association with the use of Tegretol, data are not available to estimate accurately their incidence or outcome. However, the vast majority of the cases of leukopenia have not progressed to the more serious conditions of aplastic anemia or agranulocytosis. Because of the very low incidence of agranulocytosis and aplastic anemia, the vast majority of minor hematologic changes observed in monitoring of patients on Tegretol are unlikely to signal the occurrence of either abnormality. Nonetheless, complete pretreatment hematological testing should be obtained as a baseline. If a patient in the course of treatment exhibits low or decreased white blood cell counts, the patient should be monitored closely. Discontinuation of the drug should be considered I any evidence of significant bone marrow depression develops.
Black box warning for Depakote® is listed in the right column. Similar warnings have been used for Depakene®, Depacon® and Stavzor®.	**Hepatotoxicity** Hepatic failure resulting in fatalities has occurred in patients receiving valproic acid and its derivatives. Experience has indicated that children under the age of two years are at a considerably increased risk of developing fatal hepatotoxicity, especially those on multiple anticonvulsants, those with congenital metabolic disorders, those with severe seizure disorders accompanied by mental retardation and those with organic brain disease. When Depakote is used in this patient group, it should be used with extreme caution and as a sole agent. The benefits of the therapy should be weighed against the risks. Above this age group, experience in epilepsy has indicated that the incidence of fetal hepatotoxicity decreases considerably in progressively older patient groups. These incidents usually have occurred during the first six months of treatment. Serious or fatal hepatotoxicity may be preceded by nonspecific symptoms such as malaise, weakness, lethargy, facial edema, anorexia and vomiting. In patients with epilepsy, a loss of seizure control may also

Drug names	Boxed Warnings
	occur. Patients should be monitored closely for appearance of these symptoms. Liver function tests should be performed prior to therapy and at frequent intervals thereafter, especially during the first six months. **Teratogenicity** Valproate can produce teratogenic effects such as neural tube defects (e.g., Spina Bifida). Accordingly, the use of Depakote tablets in women of childbearing potential requires that the benefits of its use be weighed against the risk of injury to the fetus. This is especially important when the treatment of a spontaneously reversible condition not ordinarily associated with permanent injury or risk of death (e.g. migraine) is contemplated. See Warnings, information for patients. Patient information leaflet describing the teratogenic potential of valproate is available for patients. **Pancreatitis** Cases of life threatening pancreatitis have been reported in both children and adults receiving valproate. Some of the cases have been described as hemorrhagic with rapid progression from initial use as well as after several years of use. Patients and guardians should be warned that abdominal pain, nausea, vomiting, and/or anorexia can be symptoms of pancreatitis that require prompt medical evaluation. If pancreatitis is diagnosed, pancreatitis should ordinarily be discontinued. Alternative treatment for the underlying medical condition should be initiated as clinically indicated. (See Warnings and Precautions.)
Black box warning for Lamictal® is listed in the right column. Similar black box warnings have been issued for Lamictal ODT®, Lamictal XR®, and Lamictal CD®.	**Warning: Serious Skin Rashes** LAMICTAL® can cause serious rashes requiring hospitalization and discontinuation of treatment. The incidence of these rashes, which have included Stevens Johnson syndrome, is approximately 0.8% (8 per 1,000) in pediatric patients (2 to 16 years of age) receiving LAMICTAL as adjunctive therapy for epilepsy and 0.3% (3 per 1,000) in adults on adjunctive therapy for epilepsy. In clinical trials of bipolar and other mood disorders, the rate of serious rash was 0.08% (0.8 per 1,000) in adult patients receiving LAMICTAL as initial monotherapy and 0.13% (1.3 per 1,000) in adult patients receiving LAMICTAL as adjunctive therapy. In a prospectively followed cohort of 1,983 pediatric patients (2 to 16 years of age) with epilepsy taking adjunctive LAMICTAL, there was 1 rash-related death. In worldwide postmarketing experience, rare cases of toxic epidermal necrolysis and/or rash-related death have been

Drug names	Boxed Warnings
	reported in adult and pediatric patients, but their numbers are too few to permit a precise estimate of the rate. Other than age, there are as yet no factors identified that are known to predict the risk of occurrence or the severity of rash caused by LAMICTAL. There are suggestions, yet to be proven, that the risk of rash may also be increased by (1) coadministration of LAMICTAL with valproate (includes valproic acid and divalproex sodium), (2) exceeding the recommended initial dose of LAMICTAL, or (3) exceeding the recommended dose escalation for LAMICTAL. However, cases have occurred in the absence of these factors. Nearly all cases of life-threatening rashes caused by LAMICTAL have occurred within 2 to 8 weeks of treatment initiation. However, isolated cases have occurred after prolonged treatment (e.g., 6 months). Accordingly, duration of therapy cannot be relied upon as means to predict the potential risk heralded by the first appearance of a rash. Although benign rashes are also caused by LAMICTAL, it is not possible to predict reliably which rashes will prove to be serious or life-threatening. Accordingly, LAMICTAL should ordinarily be discontinued at the first sign of rash, unless the rash is clearly not drug-related. Discontinuation of treatment may not prevent a rash from becoming life threatening or permanently disabling or disfiguring [see Warnings and Precautions (5.1)].

Appendix D. Search strategies

The searches were repeated in October 2010 to identify additional citations.
Database: Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) <1947 to July Week 1 2010>
Search Strategy:
---
1 fibromyalgia.mp. or Fibromyalgia/ (5623)
2 limit 1 to (english language and humans and (clinical trial, all or clinical trial or controlled clinical trial or randomized controlled trial)) (526)
3 (201004$ or 201005$ or 201006$ or 201007$).ed. (202611)
4 2 and 3 (13)
5 from 4 keep 1-13 (13)

Database: Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) <1947 to July Week 1 2010>
Search Strategy:
---
1 fibromyalgia.mp. or Fibromyalgia/ (5623)
2 limit 1 to (english language and humans) (4571)
3 ae.fs. (1128794)
4 2 and 3 (298)
5 from 4 keep 1-298 (298)

Database: Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) <1947 to July Week 1 2010>
Search Strategy:
---
1 fibromyalgia.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier] (5623)
2 limit 1 to (english language and humans) (4571)
3 (Medline or systematic review or search).tw. or meta-analysis.pt. (146695)
4 2 and 3 (171)
5 from 4 keep 1-171 (171)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <2nd Quarter 2010>
Search Strategy:
---
1 fibromyalgia.mp. or Fibromyalgia/ (568)
2 from 1 keep 1-568 (568)

Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to June 2010>
Search Strategy:
---
1 fibromyalgia.mp. [mp=title, short title, abstract, full text, keywords, caption text] (47)
2 limit 1 to full systematic reviews (32)
3 from 2 keep 1-32 (32)

Database: EBM Reviews - Database of Abstracts of Reviews of Effects <2nd Quarter 2010>
Search Strategy:
---
1 fibromyalgia.mp. [mp=title, full text, keywords] (57)
2 from 1 keep 1-57 (57)

Appendix E. Excluded studies

The following full-text publications were considered for inclusion but failed to meet the criteria for this report. Exclusion codes: 2=ineligible outcome, 3=ineligible intervention, 4=ineligible population, 5=ineligible publication type, 6=ineligible study design

Excluded studies	Exclusion code
Head-to-head trials	
Goldenberg DL, Felson DT, Dinerman H. A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia. Arthritis Rheum. Nov 1986;29(11):1371-1377.	6
Kravitz HM, Katz RS, Helmke N, Jeffriess H, Bukovsky J, Fawcett J. Alprazolam and ibuprofen in the treatment of fibromyalgia - Report of a double-blind placebo-controlled study. J Musculoskelet Pain. 1994;2(1):3-27.	4
Ozerbil O, Okudan N, Gokbel H, Levendoglu F. Comparison of the effects of two antidepressants on exercise performance of the female patients with fibromyalgia. Clin Rheumatol. Jul 2006;25(4):495-497.	2
Quijada-Carrera J, Valenzuela-Castano A, Povedano-Gomez J, et al. Comparison of tenoxicam and bromazepan in the treatment of fibromyalgia: a randomized, double-blind, placebo-controlled trial. Pain. May-Jun 1996;65(2-3):221-225.	3
Russell IJ, Fletcher EM, Michalek JE, McBroom PC, Hester GG. Treatment of primary fibrositis/fibromyalgia syndrome with ibuprofen and alprazolam. A double-blind, placebo-controlled study. Arthritis Rheum. May 1991;34(5):552-560.	4
Sorensen J, Bengtsson A, Ahlner J, Henriksson KG, Ekselius L, Bengtsson M. Fibromyalgia--are there different mechanisms in the processing of pain? A double blind crossover comparison of analgesic drugs. J Rheumatol. Aug 1997;24(8):1615-1621.	3
Sorensen J, Bengtsson A, Backman E, Henriksson KG, Bengtsson M. Pain analysis in patients with fibromyalgia. Effects of intravenous morphine, lidocaine, and ketamine. Scand J Rheumatol. 1995;24(6):360-365.	3
Ware MA, Fitzcharles M-A, Joseph L, Shir Y. The effects of nabilone on sleep in fibromyalgia: results of a randomized controlled trial. Anesth Analg. Feb 2010;110(2):604-610.	3
Placebo-controlled trials	
Bennett RM, Clark SC, Walczyk J. A randomized, double-blind, placebo-controlled study of growth hormone in the treatment of fibromyalgia. Am J Med. Mar 1998;104(3):227-231.	3
Bennett RM, Kamin M, Karim R, Rosenthal N. Tramadol and acetaminophen combination tablets in the treatment of fibromyalgia pain: a double-blind, randomized, placebo-controlled study. Am J Med. May 2003;114(7):537-545.	3
Bennett RM, Schein J, Kosinski MR, Hewitt DJ, Jordan DM, Rosenthal NR. Impact of fibromyalgia pain on health-related quality of life before and after treatment with tramadol/acetaminophen. Arthritis Rheum. Aug 15 2005;53(4):519-527.	3
Biasi G, Manca S, Manganelli S, Marcolongo R. Tramadol in the fibromyalgia syndrome: a controlled clinical trial versus placebo. Int J Clin Pharmacol Res. 1998;18(1):13-19.	3
Branco JC, Danneskiold-Samsoe B, Cherin P, et al. Effect of milnacipran on pain in fibromyalgia: A european, double-blind, randomized, placebo-controlled trial. Ann Neurol. no: M-4, 2009 2009;66(3, Supplement 13):S4.	5
Caruso I, Sarzi Puttini PC, Boccassini L, et al. Double-blind study of dothiepin versus placebo in the treatment of primary fibromyalgia syndrome. J Int Med Res. May-Jun	4

Excluded studies	Exclusion code
1987;15(3):154-159.	
Chappell AS, Ossanna MJ, Liu-Seifert H, et al. Duloxetine, a centrally acting analgesic, in the treatment of patients with osteoarthritis knee pain: a 13-week, randomized, placebo-controlled trial. Pain. Dec 2009;146(3):253-260.	4
Cuatrecasas G, Riudavets C, Guell MA, Nadal A. Growth hormone as concomitant treatment in severe fibromyalgia associated with low IGF-1 serum levels. A pilot study. BMC Musculoskeletal Disorders. 2007;8:119.	3
Distler O, Eich W, Dokoupilova E, et al. Evaluation of the efficacy and safety of terguride in patients with fibromyalgia syndrome: results of a twelve-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. Arthritis Rheum. Jan 2010;62(1):291-300.	3
Drewes AM, Andreasen A, Jennum P, Nielsen KD. Zopiclone in the treatment of sleep abnormalities in fibromyalgia. Scand J Rheumatol. 1991;20(4):288-293.	4
Fossaluzza V, De Vita S. Combined therapy with cyclobenzaprine and ibuprofen in primary fibromyalgia syndrome. Int J Clin Pharmacol Res. 1992;12(2):99-102.	4
Gronblad M, Nykanen J, Konttinen Y, Jarvinen E, Helve T. Effect of zopiclone on sleep quality, morning stiffness, widespread tenderness and pain and general discomfort in primary fibromyalgia patients. A double-blind randomized trial. Clin Rheumatol. Jun 1993;12(2):186-191.	4
Hannonen P, Malminiemi K, Yli-Kerttula U, Isomeri R, Roponen P. A randomized double-blind placebo controlled study of moclobemide and amitriptyline in fibromyalgia. Scand-J-Rheumatol. 1998;27(3):246.	5
Holman AJ, Myers RR. A randomized, double-blind, placebo-controlled trial of pramipexole, a dopamine agonist, in patients with fibromyalgia receiving concomitant medications. Arthritis Rheum. Aug 2005;52(8):2495-2505.	6
Jaeschke R, Adachi J, Guyatt G, Keller J, Wong B. Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials. J Rheumatol. Mar 1991;18(3):447-451.	4
Moldofsky H, Lue FA, Mously C, Roth-Schechter B, Reynolds WJ. The effect of zolpidem in patients with fibromyalgia: a dose ranging, double blind, placebo controlled, modified crossover study. J Rheumatol. Mar 1996;23(3):529-533.	3
Olin R, Klein R, Berg PA. A randomised double-blind 16-week study of ritanserin in fibromyalgia syndrome: clinical outcome and analysis of autoantibodies to serotonin, gangliosides and phospholipids. Clin Rheumatol. 1998;17(2):89-94.	3
Pae C-U, Masand PS, Marks DM, et al. History of early abuse as a predictor of treatment response in patients with fibromyalgia: a post-hoc analysis of a 12-week, randomized, double-blind, placebo-controlled trial of paroxetine controlled release. World Journal of Biological Psychiatry. 2009;10(4 Pt 2):435-441.	2
Palmer R, Forest Laboratories Inc. A Study Of Milnacipran In Patients With Fibromyalgia: Effects On 24 Hour Ambulatory Blood Pressure Monitoring. ClinicalTrials.gov. 2009;Identifier: NCT00618956.	2
Sadreddini S, Molaeefard M, Noshad H, Ardalan M, Asadi A. Efficacy of Raloxifen in treatment of fibromyalgia in menopausal women. European Journal of Internal Medicine. Jul 2008;19(5):350-355.	3
Skrabek RQ, Galimova L, Ethans K, Perry D. Nabilone for the treatment of pain in fibromyalgia. Journal of Pain. Feb 2008;9(2):164-173.	3
Yavuzer G, Kucukdeveci A, Arasil T, Elhan A. Moclobemid treatment in primary fibromyalgia syndrome. European Journal of Physical Medicine and Rehabilitation.	3

Excluded studies	Exclusion code
1998;8(2):35-38.	
Younger J, Mackey S. Fibromyalgia symptoms are reduced by low-dose naltrexone: a pilot study. Pain Medicine. May-Jun 2009;10(4):663-672.	3
Younger JW, Zautra AJ, Cummins ET. Effects of naltrexone on pain sensitivity and mood in fibromyalgia: no evidence for endogenous opioid pathophysiology. PLoS ONE [Electronic Resource]. 2009;4(4):e5180.	2
Yunus MB, Masi AT, Aldag JC. Short term effects of ibuprofen in primary fibromyalgia syndrome: a double blind, placebo controlled trial.[Erratum appears in J Rheumatol 1989 Jun;16(6):855]. J Rheumatol. Apr 1989;16(4):527-532.	4

Appendix F. Strength of evidence

Direct Evidence

Table 1: Paroxetine compared with amitriptyline (Ataoglu 1997)

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
50% response						
No data	NA	NA	NA	NA	NA	Insufficient
Fibromyalgia Impact Questionnaire: Mean change						
No data	NA	NA	NA	NA	NA	Insufficient
Pain: Mean change in number of tender points						
1; N=68	RCT/Fair	NA	Direct	Imprecise	-28% vs. -1%; z=5.64, *P*<0.001	Low
Fatigue						
1; N=68	RCT/Fair	NA	Direct	Imprecise	-9% vs. -5%; z=0.74, NSD	Low
Overall adverse events						
1; N=68	RCT/Fair	NA	Direct	Imprecise	RR 0.40 (0.24 to 0.60)	Low
Withdrawals due to adverse events						
1; N=68	RCT/Fair	NA	Direct	Imprecise	RR 0.36 (0.08 to 1.49)	Low

Table 2: Cyclobenzaprine compared with amitriptyline (Carette 1994)

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
50% response						
No data	NA	NA	NA	NA	NA	Insufficient
Fibromyalgia Impact Questionnaire: Mean change						
No data	NA	NA	NA	NA	NA	Insufficient
Pain: Mean change in visual analogue scale						
1; N=184	RCT/Fair	NA	Direct	Imprecise	NSD, *P* value NR	Low
Fatigue: Mean change in visual analogue scale						
1; N=184	RCT/Fair	NA	Direct	Imprecise	NSD, *P* value NR	Low
Overall adverse events						
1; N=184	RCT/Fair	NA	Direct	Precise	RR 1.02 (0.96 to 1.11)	Moderate
Withdrawals due to adverse events						
1; N=184	RCT/Fair	NA	Direct	Imprecise	RR 1.90 (0.82 to 4.44)	Low

Table 3: Nortriptyline compared with amitriptyline

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
50% response						
No data	NA	NA	NA	NA	NA	Insufficient
Fibromyalgia Impact Questionnaire: Mean change						
1;N=106	RCT/Fair	NA	Direct	Imprecise	ANOVA group effect, *P*=0.071	Low
Pain: Mean change in number of tender points						
1; N=106	RCT/Fair	NA	Indirect	Imprecise	ANOVA group effect, *P*=0.704	Low
Fatigue						
No data	NA	NA	NA	NA	NA	Insufficient
Overall adverse events						
1; N=78	RCT/Fair	NA	Direct	Precise	RR 2.04 (1.53 to 2.72)	Moderate
Withdrawals due to adverse events						
1, N=78	RCT/Fair	NA	Direct	Imprecise	RR 3.15 (0.28 to infinity)	Low

Indirect meta-analysis

Table 4: Duloxetine compared with milnacipran

	Domains pertaining to strength of evidence				Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 Pain: Difference in mean difference						
Duloxetine 3, N=1081 Milnacipran 5, N=4118	RCT/Fair	Consistent	Indirect	Precise	-0.48 (-0.80 to -0.08)	Low
Outcome 2 50% Response in reduction of pain: Ratio of RR						
Duloxetine 3, N=1081 Milnacipran 3, N=2038	RCT/Fair	Consistent	Indirect	Imprecise	1.13 (0.85 to 1.50)	Low
Outcome 3 Fibromyalgia Impact Questionnaire: Difference in mean difference						
Duloxetine 3, N=1081 Milnacipran 4, N=3993	RCT/Fair	Consistent	Indirect	Imprecise	-2.63 (-5.32 to 0.06)	Low
Outcome 4 Fatigue: Difference in standardized mean difference						
Duloxetine 2, N=727 Milnacipran 4, N= 3993	RCT/Fair	Consistent	Indirect	Imprecise	0.04 (-1.34 to 0.21)	Low
Outcome 5 Overall Withdrawal: Ratio of RR (95% CI)						
Duloxetine 3, N=1081 Milnacipran 4, N= 3230	RCT/Fair	Consistent	Indirect	Precise	-.74 (0.55 to 1.00)	Low
Outcome 6 Overall Adverse events: Ratio of RR (95% CI)						
Duloxetine 2, N=561 Milnacipran 3, N= 3105	RCT/Fair	Consistent	Indirect	Precise	1.04 (0.96 to 1.14)	Low
Outcome 7 Withdrawal due to adverse events: Ratio of RR (95% CI)						
Duloxetine 3, N=1081 Milnacipran 4, N= 3230	RCT/Fair	Consistent	Indirect	Imprecise	0.84 (0.51 to 1.39)	Low

Table 5: Duloxetine compared with pregabalin

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 **Pain: Difference in mean difference**						
Duloxetine 3, N=1081 Pregabalin 4, N= 2774	RCT/Fair	Consistent	Indirect	Precise	-0.27 (-0.68 to -0.15)	Low
Outcome 2 **50% Response in reduction of pain: Ratio of RR**						
Duloxetine 3, N=1081 Pregabalin 3, N=2026	RCT/fair	Consistent	Indirect	Precise	0.91 (0.65 to 1.28)	Low
Outcome 3 **Fibromyalgia Impact Questionnaire: Difference in mean difference**						
Duloxetine 3, N=1081 Pregabalin 2, N=1498	RCT/fair	Consistent	Indirect	Imprecise	-2.97 (-6.40 to 0.46)	Low
Outcome 4 **Fatigue: Difference in standardized mean difference**						
Duloxetine 2, N=727 Pregabalin 2, 1279	RCT/fair	Consistent	Indirect	Precise	0.13 (-0.12 to 0.38)	Low
Outcome 5 **Overall withdrawal: Ratio of RR (95% CI)**						
Duloxetine 3, N=1081 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Precise	0.84 (0.64 to 1.10)	Low
Outcome 6 **Overall adverse events: Ratio of RR (95% CI)**						
Duloxetine 2, N=561 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Precise	0.99 (0.90 to 1.09)	Low
Outcome 7 **Withdrawal due to adverse events: Ratio of RR (95% CI)**						
Duloxetine 3, N=1081 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Imprecise	0.83 (0.56 to 1.24)	Low

Table 6: Milnacipran compared with pregabalin

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 **Pain: Difference in mean difference**						
Milnacipran 5, N=4118 Pregabalin 4, N=2774	RCT/fair	Consistent	Indirect	Imprecise	0.18 (-0.13 to 0.49)	Low
Outcome 2 **50% Response in reduction of pain: Ratio of RR**						
Milnacipran 3, N=2038 Pregabalin 3, N=2026	RCT/fair	Consistent	Indirect	Imprecise	0.81 (0.59 to 1.10)	Low
Outcome 3 **Fibromyalgia Impact Questionnaire: Difference in mean difference**						
Milnacipran 4, N=3993 Pregabalin 2, N=1498	RCT/Fair	Consistent	Indirect	Imprecise	-0.34 (-3.27 to 2.59)	Low
Outcome 4 **Fatigue: Difference in standardized mean difference**						
Milnacipran 4, N=3993 Pregabalin 2, N=1279	RCT/Fair	Consistent	Indirect	Imprecise	0.09 (-0.11 to 0.30)	Low
Outcome 5 **Overall withdrawal: Ratio of RR (95% CI)**						
Milnacipran 4, N=3230 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Imprecise	1.14 (0.86 to 1.50)	Low
Outcome 6 **Overall adverse events: Ratio of RR (95% CI)**						
Milnacipran 3, N=3105 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Precise	0.95 (0.89 to 1.01)	Low
Outcome 7 **Withdrawal due to adverse wvents: Ratio of RR (95% CI)**						
Milnacipran 4, N=3230 Pregabalin 4, N=2774	RCT/Fair	Consistent	Indirect	Imprecise	0.99 (0.63 to 1.55)	Low

Table 7: Duloxetine compared with amitriptyline

	Domains pertaining to strength of evidence				Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 **Pain: Difference in mean difference**						
Duloxetine 3, N=1081 Amitriptyline 3, N=207	RCT/Fair	Consistent	Indirect	Imprecise	-0.01 (-1.29 to 1.29)	Low
Outcome 2 **50% Response in reduction of pain: Ratio of RR**						
Duloxetine 3, N=1081 Amitriptyline 2, N=68	RCT/Fair	Consistent	Indirect	Imprecise	Insufficient data for comparison	Insufficient
Outcome 3 **Fibromyalgia Impact Questionnaire: Difference in mean difference**						
Duloxetine 3, N=1081 Amitriptyline 1, 80	RCT/Fair	NA	Indirect	Imprecise	1.19 (-10.73 to 13.11)	Insufficient
Outcome 4 **Fatigue: Difference in standardized mean difference**						
Duloxetine 2, N=727 Amitrityline 2, N=148	RCT/Fair	Consistent	Indirect	Imprecise	0.14 (-0.22 to 0.49)	Low
Outcome 5 **Overall withdrawal: Ratio of RR (95% CI)**						
Duloxetine 3, N=1081 Amitrityline 4, N=287	RCT/Fair	Consistent	Indirect	Imprecise	1.45 (0.66 to 1.51)	Low
Outcome 6 **Overall adverse events: Ratio of RR (95% CI)**						
Duloxetine 2, N=561 Amitrityline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	0.81 (0.37 to 1.74)	Insufficient
Outcome 7 **Withdrawal due to adverse events: Ratio of RR (95%CI)**						
Duloxetine 3, N=1081 Amitrityline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	1.62 (0.50 to 5.32)	Insufficient

Table 8: Milnacipran compared with amitriptyline

	Domains pertaining to strength of evidence				Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 **Pain: Difference in mean difference**						
Milnacipran 5, N=4118 Amitriptyline 3, N=207	RCT/Fair	Consistent	Indirect	Imprecise	0.44 (-0.81 to 1.70)	Low
Outcome 2 **50% Response in reduction of pain: Ratio of RR**						
Milnacipran 3, N=2038 Amitriptyline 2, N=68	RCT/Fair	Consistent	Indirect	Imprecise	Insufficient data for comparison	Insufficient
Outcome 3 **Fibromyalgia Impact Questionnaire: Difference in mean difference**						
Milnacipran 4, N=3993 Amitriptyline 1, N=80	RCT/fair	NA	Indirect	Imprecise	3.82 (-7.96 to 15.61)	Insufficient
Outcome 4 **Fatigue: Difference in standardized mean difference**						
Milnacipran 4, N=3993 Amitriptyline, N=148	RCT/Fair	Consistent	Indirect	Imprecise	0.10 (-0.22 to 0.43)	Low
Outcome 5 **Overall withdrawal: Ratio of RR (95% CI)**						
Milnacipran 4, N=3230 Amitriptyline 4, N=287	RCT/Fair	Consistent	Indirect	Imprecise	1.96 (0.89 to 4.30)	Low
Outcome 6 **Overall adverse events: Ratio of RR (95%CI)**						
Milnacipran 3, N=3105 Amitriptyline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	0.77 (0.36 to 1.66)	Insufficient
Outcome 7 **Withdrawal due to adverse events: Ratio of RR (95% CI)**						
Milnacipran 4, N=3230 Amitriptyline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	1.94 (0.58 to 6.46)	Insufficient

Table 9: Pregabalin compared with amitriptyline

	Domains pertaining to strength of evidence				Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Outcome 1 **Pain: Difference in mean difference**						
Pregabalin 4, N=2774 Amitriptyline 3, N=207	RCT/Fair	Consistent	Indirect	Imprecise	0.26 (-1.01 to 1.54)	Low
Outcome 2 **50% Response in reduction of pain: Ratio of RR**						
Pregabalin 3, N=2026 Amitriptyline 2, N=68	RCT/Fair	Consistent	Indirect	Imprecise	Insufficient data for comparison	Insufficient
Outcome 3 **Fibromyalgia Impact Questionnaire: Difference in mean difference**						
Pregabalin 2, N=1498 Amitriptyline 1, N=80	RCT/fair	NA	Indirect	Imprecise	4.17 (-7.81 to 16.14)	Insufficient
Outcome 4 **Fatigue: Difference in standardized mean difference**						
Pregabalin 4, N=2774 Amitriptyline 2, N=148	RCT/Fair	Consistent	Indirect	Imprecise	0.01 (-0.37 to 0.38)	Low
Outcome 5 **Overall withdrawal: Ratio of RR (95% CI)**						
Pregabalin 4, N=2774 Amitriptyline 4, N=287	RCT/Fair	Consistent	Indirect	Imprecise	1.72 (0.79 to 3.74)	Low
Outcome 6 **Overall adverse events: Ratio of RR (95% CI)**						
Pregabalin 4, N=2774 Amitriptyline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	0.81 (0.38 to 1.75)	Insufficient
Outcome 7 **Withdrawal due to adverse events: Ratio of RR (95% CI)**						
Pregabalin 4, N=2774 Amitriptyline 3, N=265	RCT/Fair	Inconsistent	Indirect	Imprecise	1.96 (0.61 to 6.30)	Insufficient

Comparisons to placebo

Table 10: Gabapentin compared with placebo

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Pain: Change in BPI mean pain severity score (range 0-10)						
1; N=119	RCT/Fair	NA	Indirect	Precise	Difference: -0.92 (-1.75 to -0.71), *P*=0.015	Insufficient
Fibromyalgia Impact Questionnaire Total Score						
1; N=119	RCT/Fair	NA	Indirect	Precise	Difference: -8.4 (-13.0 to -3.3), *P*=0.001	Insufficient
Withdrawals due to adverse events						
1; N=119	RCT/Fair	NA	Indirect	Imprecise	16% vs. 9%, *P*=0.034	Insufficient
Fatigue, 50% response, overall adverse events						
No data						

Table 11: Cyclobenzaprine compared with placebo

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Pain: Various						
3; N=172	RCT/Fair	Inconsistent	Indirect	Imprecise	Superior in 1 of 3 trials	Insufficient
Fatigue: Various						
3; N=172	RCT/Fair	Inconsistent	Indirect	Imprecise	Superior only in evening in 1 of 3 trials	Insufficient
Withdrawals due to adverse events						
3; N=172	RCT/Fair	Consistent	Indirect	Imprecise	Pooled RR 2.56; 95% CI 0.48 to 13.59	Insufficient
Overall adverse events						
1; N=120	RCT/Fair	NA	Indirect	Precise	89% vs. 64%, *P*=0.002	Insufficient
Fibromyalgia Impact Questionnaire Total Score, 50% response						
No data						

Table 12: Fluoxetine compared with placebo

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Pain: Mean change in visual analogue scale						
1; N=51	RCT/Fair	NA	Indirect	Precise	-29% vs. -7%, *P*=0.002	Insufficient
Fibromyalgia Impact Questionnaire Total Score						
1; N=51	RCT/Fair	NA	Indirect	Precise	-20% vs. +7%, *P*=0.005	Insufficient
Fatigue						
1; N=51	RCT/Fair	NA	Indirect	Precise	-16% vs. +4%, *P*=0.05	Insufficient
50% response, overall adverse events, withdrawals due to adverse events						
No data	NA	NA	NA	NA	NA	Insufficient

Table 13: Paroxetine compared with placebo

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
50% Response (Fibromyalgia Impact Questionnaire)						
1; N=116	RCT/Fair	NA	Indirect	Imprecise	26% vs. 14%; *P*=0.08	Insufficient
Fibromyalgia Impact Questionnaire Total						
1; N=116	RCT/Fair	NA	Indirect	Precise	-36% vs. -27%; *P*=0.015	Insufficient
Pain: Fibromyalgia Impact Questionnaire subscale						
1; N=116	RCT/Fair	NA	Indirect	Imprecise	*P*=0.07 (data NR)	Insufficient
Fatigue						
1; N=116	RCT/Fair	NA	Indirect	Precise	*P*<0.05 (data NR)	Insufficient
Overall adverse events						
1; N=116	RCT/Fair	NA	Indirect	Imprecise	65% vs. 59%, NSD	Insufficient
Withdrawals due to adverse events						
1; N=116	RCT/Fair	NA	Indirect	NA	7% vs. 2%, NSD	Insufficient

Table 14: Citalopram compared with placebo

Domains pertaining to strength of evidence					Magnitude of effect	Strength of evidence
Number of studies; number of subjects	**Risk of bias (design/ quality)**	**Consistency**	**Directness**	**Precision**	**Summary effect size (95% confidence interval)**	**High, Moderate, Low, Insufficient**
Pain						
2; N=82	RCT/Fair	Consistent	Indirect	Imprecise	VAS change (range): -12% to -16% vs. -4% to -10%; NSD	Insufficient
Fatigue						
1; N=42	RCT/Fair	NA	Indirect	Imprecise	VAS change: -7% vs. -1%, NSD	Insufficient
Withdrawals due to adverse events						
1; N=40	RCT/Fair	NA	Indirect	Imprecise	14% vs. 0%, NSD	Insufficient
50% response, Fibromyalgia Impact Questionnaire Total Score, overall adverse events						
No data	NA	NA	NA	NA	NA	Insufficient