



## Health Technology Assessment

Volume 28 • Issue 62 • October 2024

ISSN 2046-4924

# Antidepressants for pain management in adults with chronic pain: a network meta-analysis

*Hollie Birkinshaw, Claire Friedrich, Peter Cole, Christopher Eccleston, Marc Serfaty,  
Gavin Stewart, Simon White, Andrew Moore, David Phillippo and Tamar Pincus*





# Antidepressants for pain management in adults with chronic pain: a network meta-analysis

Hollie Birkinshaw<sup>1</sup>, Claire Friedrich<sup>1</sup>, Peter Cole<sup>2</sup>,  
Christopher Eccleston<sup>3</sup>, Marc Serfaty<sup>4</sup>,  
Gavin Stewart<sup>5</sup>, Simon White<sup>6</sup>, Andrew Moore<sup>7</sup>,  
David Phillippo<sup>8</sup> and Tamar Pincus<sup>1\*</sup>

<sup>1</sup>Department of Psychology, University of Southampton, Southampton, UK

<sup>2</sup>Oxford Pain Relief Unit, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK

<sup>3</sup>Centre for Pain Research, University of Bath, Bath, UK

<sup>4</sup>University College London, London, UK

<sup>5</sup>Newcastle University, Newcastle upon Tyne, UK

<sup>6</sup>School of Pharmacy and Bioengineering, Keele University, Keele, UK

<sup>7</sup>Retired, Plymouth, UK

<sup>8</sup>Bristol Medical School, University of Bristol, Bristol, UK

\*Corresponding author

Published October 2024

DOI: 10.3310/MKRT2948

This report should be referenced as follows:

Birkinshaw H, Friedrich C, Cole P, Eccleston C, Serfaty M, Stewart G, *et al.* Antidepressants for pain management in adults with chronic pain: a network meta-analysis. *Health Technol Assess* 2024;**28**(62). <https://doi.org/10.3310/MKRT2948>



# Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hta](http://www.journalslibrary.nihr.ac.uk/hta).

## Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

## This article

The research reported in this issue of the journal was funded by the HTA programme as award number NIHR128782. The contractual start date was in April 2020. The draft manuscript began editorial review in April 2023 and was accepted for publication in March 2024. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Birkinshaw *et al.* This work was produced by Birkinshaw *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Newgen Digitalworks Pvt Ltd, Chennai, India ([www.newgen.co](http://www.newgen.co)).



# Abstract

## Antidepressants for pain management in adults with chronic pain: a network meta-analysis

Hollie Birkinshaw<sup>1</sup>, Claire Friedrich<sup>1</sup>, Peter Cole<sup>2</sup>,  
Christopher Eccleston<sup>3</sup>, Marc Serfaty<sup>4</sup>, Gavin Stewart<sup>5</sup>,  
Simon White<sup>6</sup>, Andrew Moore<sup>7</sup>, David Phillippo<sup>8</sup> and Tamar Pincus<sup>1\*</sup>

<sup>1</sup>Department of Psychology, University of Southampton, Southampton, UK

<sup>2</sup>Oxford Pain Relief Unit, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK

<sup>3</sup>Centre for Pain Research, University of Bath, Bath, UK

<sup>4</sup>University College London, London, UK

<sup>5</sup>Newcastle University, Newcastle upon Tyne, UK

<sup>6</sup>School of Pharmacy and Bioengineering, Keele University, Keele, UK

<sup>7</sup>Retired, Plymouth, UK

<sup>8</sup>Bristol Medical School, University of Bristol, Bristol, UK

\*Corresponding author [t.pincus@soton.ac.uk](mailto:t.pincus@soton.ac.uk)

**Background:** Chronic pain is common and costly. Antidepressants are prescribed to reduce pain. However, there has not been a network meta-analysis examining all antidepressants across all chronic pain conditions, so effectiveness and safety for most antidepressants for pain conditions remain unknown.

**Objective:** To assess the efficacy and safety of antidepressants for chronic pain (except headache) in adults. Our primary outcomes were as follows: substantial pain relief (50%), pain intensity, mood and adverse events. Our secondary outcomes were as follows: moderate pain relief (30%), physical function, sleep, quality of life, Patient Global Impression of Change, serious adverse events and withdrawal.

**Design:** This was a systematic review with a network meta-analysis. We searched CENTRAL, MEDLINE, EMBASE, CINAHL, LILACS, AMED and PsycINFO databases for randomised controlled trials of antidepressants for chronic pain conditions up until 4 January 2022. The review was registered in PROSPERO (CRD42020171855), and the protocol was published in the Cochrane Library (<https://doi.org/10.1002/14651858.CD014682>).

**Setting:** We analysed trials from all settings.

**Participants:** We included trials in which participants had chronic pain, defined as longer than 3 months, from any condition excluding headache.

**Interventions:** We included all antidepressants.

**Main outcome measures:** Our primary outcome was substantial pain relief, defined as a reduction > 50%. We also measured pain intensity, mood and adverse events. Secondary measures included moderate pain relief (above 30% reduction), physical function, sleep, quality of life, Global Impression of Change, serious adverse events, and withdrawal from trial.

**Results:** We identified 176 studies with a total of 28,664 participants. Most studies were placebo-controlled ( $n = 83$ ) and parallel armed ( $n = 141$ ). The most common pain conditions examined were fibromyalgia (59 studies), neuropathic pain (49 studies) and musculoskeletal pain (40 studies). The

average length of randomised controlled trials was 10 weeks. Most studies measured short-term outcomes only and excluded people with low mood and other mental health conditions.

Across efficacy outcomes, duloxetine was consistently the highest-ranked antidepressant with moderate- to high-certainty evidence. Standard dose was equally efficacious as high dose for the majority of outcomes. Milnacipran was often ranked as the next most efficacious antidepressant, although the certainty of evidence was lower than that for duloxetine. There was insufficient evidence to draw robust conclusions for the efficacy and safety of any other antidepressant for chronic pain.

**Limitations:** The evidence for antidepressants other than duloxetine is poor. For duloxetine, it is not clear whether the effect applies to groups with both pain and low mood, since these groups were excluded from trials. There is also insufficient evidence on long-term outcomes and on adverse effects.

**Conclusions:** There is only reliable evidence for duloxetine in the treatment of chronic pain. Duloxetine was moderately efficacious across all outcomes at standard dose. There is also promising evidence for milnacipran, although further high-quality research is needed to be confident in these conclusions. Data for all other antidepressants were of low certainty. However, the findings should not be read as an encouragement to prescribe antidepressants where other non-pharmacological intervention could be equally effective, especially in the absence of good evidence on side effects and safety.

**Future work:** There is a need for large, methodologically sound trials testing the effectiveness of antidepressants for chronic pain. These trials should examine long-term outcomes (> 6 months) and include people with low mood. There should also be better reporting of adverse events, tolerance of drugs, and long-term compliance.

**Study registration:** This study is registered as PROSPERO CRD42020171855.

**Funding:** This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR128782) and is published in full in *Health Technology Assessment*; Vol. 28, No. 62. See the NIHR Funding and Awards website for further award information.

# Contents

List of tables	xi
List of figures	xiii
List of supplementary material	xv
List of abbreviations	xvii
Plain language summary	xix
Scientific summary	xxi
<b>Chapter 1 Background</b>	<b>1</b>
Chronic pain	1
Antidepressants and chronic pain	2
<i>What are they, and how might they work?</i>	2
<i>Guidelines for antidepressants in the treatment of chronic pain</i>	2
<i>Antidepressants in practice</i>	3
<i>Antidepressants and safety</i>	3
<i>The evidence on the efficacy and safety of antidepressants</i>	3
The relationship between pain and low mood	4
Patient and public involvement	5
<b>Chapter 2 Objectives</b>	<b>7</b>
<b>Chapter 3 Methods</b>	<b>9</b>
Criteria for considering studies for this review	9
<i>Types of studies</i>	9
<i>Types of participants</i>	9
<i>Types of interventions</i>	9
<i>Supplementary sets</i>	10
<i>Types of outcome measures</i>	10
Search methods for identification of studies	12
<i>Electronic searches</i>	12
<i>Searching other resources</i>	12
Data collection and analysis	12
<i>Selection of studies</i>	12
<i>Data extraction and management</i>	13
<i>Assessment of risk of bias in included studies</i>	13
<i>Measures of treatment effect</i>	14
<i>Unit of analysis issues</i>	15
<i>Dealing with missing data</i>	15
<i>Assessment of heterogeneity</i>	15
<i>Assessment of the transitivity assumption</i>	15
<i>Assessment of reporting biases</i>	16
Data synthesis	16
<i>Subgroup analysis and investigation of heterogeneity</i>	17
<i>Sensitivity analysis</i>	17
Summary of findings and assessment of the certainty of the evidence	17

<b>Chapter 4 Results</b>	<b>21</b>
Description of studies	21
<i>Results of the search</i>	21
<i>Included studies</i>	21
Excluded studies	23
<i>Risk of bias in included studies</i>	24
<i>Allocation</i>	24
<i>Blinding</i>	24
<i>Incomplete outcome data</i>	24
<i>Selective reporting</i>	24
<i>Other potential sources of bias</i>	25
Effects of interventions	25
<i>Overview</i>	25
<i>Primary outcomes</i>	26
<i>Secondary outcomes</i>	41
<b>Chapter 5 Differences between protocol and review</b>	<b>69</b>
Updating of the background section	69
Methods	69
<b>Chapter 6 Patient and public involvement</b>	<b>71</b>
<b>Chapter 7 Discussion</b>	<b>73</b>
Summary of main results from review	73
<i>Overall</i>	73
<i>Primary efficacy outcomes</i>	73
<i>Secondary efficacy outcomes</i>	73
<i>Safety</i>	73
Overall completeness and applicability of evidence	74
<i>Participants</i>	74
<i>Interventions</i>	74
<i>Study designs and comparisons</i>	74
<i>Mood</i>	75
<i>Follow-up</i>	75
Quality of the evidence	75
<i>Overall quality</i>	75
<i>Risk of bias</i>	75
<i>Heterogeneity</i>	76
<i>Imprecision</i>	76
<i>Inconsistency</i>	76
<i>Publication bias</i>	76
<i>Potential biases in the review process</i>	76
Agreements and disagreements with other studies or reviews	77
Implications	77
<i>Implications for practice</i>	77
<i>Implications for policy-makers and guidelines</i>	78
<i>Implications for funders of the intervention</i>	78
<i>Implications for research</i>	78
Equality, diversity and inclusion	79
<i>Research participants</i>	79
<i>Research team and wider involvement</i>	79
Summary	79
<i>Strengths and limitations</i>	79

<b>Chapter 8 Conclusion</b>	<b>81</b>
<b>Additional information</b>	<b>83</b>
<b>References</b>	<b>87</b>
<b>Appendix 1 Search strategies</b>	<b>97</b>
<b>Appendix 2 Network meta-analysis reporting decisions</b>	<b>105</b>
<b>Appendix 3 Tables of interventions included in network meta-analyses</b>	<b>111</b>
<b>Appendix 4 Characteristics of included studies</b>	<b>123</b>



# List of tables

<b>TABLE 1</b> Antidepressant dose categorisation	<b>11</b>
<b>TABLE 2</b> Substantial pain relief: summary of findings	<b>26</b>
<b>TABLE 3</b> Pain intensity: summary of findings	<b>30</b>
<b>TABLE 4</b> Mood: summary of findings	<b>33</b>
<b>TABLE 5</b> Adverse events: summary of findings	<b>36</b>
<b>TABLE 6</b> Moderate pain relief: summary of findings	<b>41</b>
<b>TABLE 7</b> Physical function: summary of findings	<b>44</b>
<b>TABLE 8</b> Sleep: summary of findings	<b>47</b>
<b>TABLE 9</b> Quality of life: summary of findings	<b>50</b>
<b>TABLE 10</b> Patient Global Impression of Change (responders): summary of findings	<b>53</b>
<b>TABLE 11</b> Patient Global Impression of Change (continuous): summary of findings	<b>56</b>
<b>TABLE 12</b> Serious adverse events: summary of findings	<b>58</b>
<b>TABLE 13</b> Withdrawal: summary of findings	<b>63</b>
<b>TABLE 14</b> Top-ranked antidepressant classes in withdrawal analysis	<b>67</b>
<b>TABLE 15</b> Overview of interventions in substantial pain relief analysis	<b>111</b>
<b>TABLE 16</b> Overview of interventions in pain intensity analysis	<b>112</b>
<b>TABLE 17</b> Overview of interventions in mood analysis	<b>113</b>
<b>TABLE 18</b> Overview of interventions in the adverse events analysis	<b>113</b>
<b>TABLE 19</b> Overview of interventions included in moderate pain relief analysis	<b>115</b>
<b>TABLE 20</b> Overview of all interventions in physical function analysis	<b>116</b>
<b>TABLE 21</b> Overview of interventions in sleep analysis	<b>117</b>
<b>TABLE 22</b> Overview of all interventions in quality of life analysis	<b>117</b>
<b>TABLE 23</b> Overview of all interventions in PGIC much/very much improved analysis	<b>118</b>

## LIST OF TABLES

<b>TABLE 24</b> Overview of all interventions in PGIC continuous analysis	<b>118</b>
<b>TABLE 25</b> Overview of all interventions in serious adverse events analysis	<b>118</b>
<b>TABLE 26</b> Overview of all interventions in withdrawal analysis	<b>120</b>

# List of figures

<b>FIGURE 1</b> Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram	21
<b>FIGURE 2</b> Risk of bias of included studies by domain	24
<b>FIGURE 3</b> Substantial pain relief: network plot	28
<b>FIGURE 4</b> Substantial pain relief: summary forest plot (log OR with Crls)	29
<b>FIGURE 5</b> Pain intensity: network diagram	32
<b>FIGURE 6</b> Pain intensity: summary forest plot (SMD with Crls)	33
<b>FIGURE 7</b> Mood: network plot	35
<b>FIGURE 8</b> Mood: summary forest plot (SMD with Crls)	35
<b>FIGURE 9</b> Adverse events: network plot	38
<b>FIGURE 10</b> Adverse events: summary forest plot (log OR with Crls)	39
<b>FIGURE 11</b> Adverse events (antidepressants only): summary forest plot	40
<b>FIGURE 12</b> Moderate pain relief: network plot	42
<b>FIGURE 13</b> Moderate pain relief: summary forest plot (log OR with Crls)	43
<b>FIGURE 14</b> Physical function: network plot	45
<b>FIGURE 15</b> Physical function: summary forest plot (SMD with Crls)	46
<b>FIGURE 16</b> Sleep: network plot	48
<b>FIGURE 17</b> Sleep: summary forest plot (SMD with Crls)	49
<b>FIGURE 18</b> Quality of life: network plot	51
<b>FIGURE 19</b> Quality of life: summary forest plot (SMD with Crls)	52
<b>FIGURE 20</b> Patient Global Impression of Change much/very much improved: network plot	55
<b>FIGURE 21</b> Patient Global Impression of Change much/very much improved: summary forest plot (log OR with Crls)	55
<b>FIGURE 22</b> Patient Global Impression of Change (continuous): network plot	57
<b>FIGURE 23</b> Patient Global Impression of Change (continuous): summary forest plot (SMD with Crls)	57

## LIST OF FIGURES

<b>FIGURE 24</b> Serious adverse events: network plot	<b>61</b>
<b>FIGURE 25</b> Serious adverse events: summary forest plot (log OR with CIs)	<b>62</b>
<b>FIGURE 26</b> Withdrawal: network plot	<b>65</b>
<b>FIGURE 27</b> Withdrawal: summary forest plot (log OR with CIs)	<b>66</b>

# List of supplementary material

## Report Supplementary Material 1 Additional supplemental files

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/MKRT2948>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.



## List of abbreviations

BOCF	baseline observation carried forward	NMA	network meta-analysis
CI	confidence interval	OR	odds ratio
CINeMA	Confidence in Network Meta-Analysis	PaPaS	Pain, Palliative and Supportive Care
CrI	credible interval	PGIC	Patient Global Impression of Change
DIC	deviance information criterion	PPI	patient and public involvement
GRADE	Grading of Recommendations, Assessment, Development and Evaluations	RCT	randomised controlled trial
IASP	International Association for the Study of Pain	ROB-MEN	Risk Of Bias due to Missing Evidence in Network meta-analysis
IMMPACT	Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials	SARI	serotonin antagonist and reuptake inhibitor
LOCF	last observation carried forward	SF-36	Short Form questionnaire-36 items
MAOI	monoamine oxidase inhibitor	SMD	standardised mean difference
NARI	noradrenaline reuptake inhibitor	SNRI	serotonin norepinephrine reuptake inhibitor
NaSSA	noradrenergic and specific serotonergic antidepressant	SSRI	selective serotonin reuptake inhibitor
NDRI	noradrenaline and dopamine reuptake inhibitor	SUCRA	surface under the cumulative ranking curve
NICE	National Institute for Health and Care Excellence	TCA	tricyclic antidepressant
		TeCA	tetracyclic antidepressant
		UME	unrelated mean effects
		VAS	visual analogue scale



# Plain language summary

## What was the question?

Chronic pain is pain that lasts for more than 3 months. Over one-third of people across the world experience chronic pain. This often has a detrimental impact on people's mood, disability and well-being. Antidepressants are often prescribed to reduce pain, but we are not sure which antidepressants work best for different types of pain, or whether they are safe.

We wanted to find out whether antidepressants were effective and safe for management of chronic pain.

## What did we do?

We searched for studies that had compared any antidepressant with any other treatment for any type of chronic pain (except headache). We compared all the treatments against each other using a statistical method called network meta-analysis. This method allows us to rank the treatments in order of best to worst for each outcome.

## What did we find?

We found 176 studies that included a total of 28,664 people with chronic pain.

Most of the studies (83/176) compared an antidepressant with a placebo (which looks like the real medicine but does not have any medicine in it).

The evidence from our analysis suggests that:

- Duloxetine is the antidepressant that we have the most confidence in. It was the best antidepressant for reducing pain and improving physical function.
- A standard dose of duloxetine was equally as effective for reducing pain as a high dose of duloxetine.
- Milnacipran was also effective at reducing pain, but we are not as confident in this result as in the one for duloxetine because there were fewer studies with fewer people involved.

Aside from duloxetine and milnacipran, we do not have confidence in the results from any other antidepressant included in this review, and even for duloxetine and milnacipran, we do not know the long-term effects.

It is important to recognise that the lack of evidence for the majority of antidepressants in this review does not necessarily equal a lack of benefit. Rather, this means that the large, high-quality trials required for us to be certain of an antidepressant's effectiveness have not been undertaken.

Altogether, although duloxetine and milnacipran are effective, the results of this review should not be read as an encouragement to prescribe antidepressants where other non-pharmacological intervention could be equally effective, especially in the absence of good evidence on side effects and safety. These conclusions were informed by our patient and public involvement group.



# Scientific summary

## Background

Chronic pain is common in adults, and often has a detrimental impact upon physical ability, well-being, and quality of life. Previous reviews have shown that certain antidepressants may be effective in reducing pain with some benefit in improving patients' global impression of change for certain chronic pain conditions. However, there has not been a network meta-analysis examining all antidepressants across all chronic pain conditions.

## Objectives

Our objective was to assess the efficacy and safety of antidepressants for chronic pain (except headache) in adults.

Our primary outcomes were as follows: substantial pain relief (50%), pain intensity, mood and adverse events. Our secondary outcomes were as follows: moderate pain relief (30%), physical function, sleep, quality of life, Patient Global Impression of Change (PGIC), serious adverse events and withdrawal.

## Search methods

We searched CENTRAL, MEDLINE, EMBASE, CINAHL, LILACS, AMED and PsycINFO databases for randomised controlled trials (RCTs) of antidepressants for chronic pain conditions up until 4 January 2022.

## Selection criteria

We included RCTs that examined antidepressants for chronic pain against any comparator. If the comparator was placebo, another medication, another antidepressant or the same antidepressant at different doses, then the study was required to be double-blind. RCTs with active comparators that were unable to be double-blinded (e.g. psychotherapy) were included but rated as at high risk of bias. We excluded RCTs where the follow-up was < 2 weeks and those with < 10 participants in each trial arm. We included any antidepressant at any dose, for any indication but used primarily for treatment of people with chronic pain and compared to placebo or active intervention.

## Participants

We included adults (aged 18 years or older) reporting primary or secondary pain in any part of their body (except headache) as their primary complaint, that matched the International Association for the Study of Pain definition of chronic pain (i.e. at least 3 months' duration). We included all trials regardless of the severity of participants' chronic pain, although we extracted whether severity was part of the inclusion criteria of the individual studies. We excluded studies where the participants' primary pain condition was headache or migraine.

## Data collection and analysis

Two authors separately screened, extracted data and judged risk of bias. We synthesised the data using Bayesian network meta-analysis (NMA) and pairwise meta-analyses for each outcome and ranked the antidepressants in terms of their effectiveness using the surface under the cumulative ranking curve. We primarily used the Confidence in Network Meta-Analysis (CINeMA) framework and 'Risk Of Bias due to Missing Evidence in Network meta-analysis' (ROB-MEN) tool to assess the certainty of the evidence. Where it was not possible to use CINeMA and ROB-MEN due to the complexity of the networks, we used Grading of Recommendations, Assessment, Development and Evaluations (GRADE) to assess the certainty of the evidence.

## Main results

This review and NMA included 176 studies with a total of 28,664 participants. The majority of studies were placebo-controlled ( $n = 83$ ) and parallel-armed ( $n = 141$ ). The most common pain conditions examined were fibromyalgia (59 studies), neuropathic pain (49 studies) and musculoskeletal pain (40 studies). The average length of RCTs was 10 weeks; seven studies provided no useable data and were omitted from the NMAs. The majority of studies measured short-term outcomes only and excluded people with low mood and other mental health conditions.

Across efficacy outcomes, duloxetine was consistently the highest-ranked antidepressant with moderate-to high-certainty evidence. In duloxetine trials, standard dose was equally efficacious as high dose for the majority of outcomes. Milnacipran was often ranked as the next most efficacious antidepressant, although the certainty of evidence was lower than that of duloxetine. There was insufficient evidence to draw robust conclusions for the efficacy and safety of any other antidepressant for chronic pain.

### Primary efficacy outcomes

For pain relief, duloxetine standard dose showed a small to moderate effect for substantial pain relief [odds ratio 0.91, 95% confidence interval (CI) 0.56 to 0.84] and continuous pain intensity [standardised mean difference (SMD)  $-0.31$ , 95% CI  $-0.39$  to  $-0.24$ ]. For pain intensity, milnacipran standard dose also showed a small effect (SMD  $-0.22$ , 95% CI  $-0.39$  to  $0.06$ ) with moderate-certainty evidence. For mood, mirtazapine had a moderate effect (SMD  $-0.5$ , 95% CI  $-0.78$  to  $-0.22$ ), while duloxetine showed a small effect ( $-0.16$ , 95% CI  $-0.22$  to  $-0.1$ ); however, it is important to note that most trials excluded participants with mental health conditions, and so average anxiety and depression scores tended to be in the 'normal' or 'subclinical' ranges at baseline already.

### Secondary efficacy outcomes

Across all secondary efficacy outcomes (moderate pain relief, physical function, sleep, quality of life and PGIC), duloxetine and milnacipran were the highest-ranked antidepressants with moderate-certainty evidence, although effects were small. For both duloxetine and milnacipran, standard doses were equally as efficacious as high doses.

### Safety

There was very low-certainty evidence for all safety outcomes (adverse events, serious adverse events and withdrawal) across all antidepressants. We cannot draw any reliable conclusions from the NMAs for these outcomes.

## Authors' conclusions

Our review and NMAs show that despite studies investigating 25 different antidepressants, there is reliable evidence for only duloxetine in the treatment of chronic pain. Duloxetine was moderately

efficacious across all outcomes at standard dose. There is also promising evidence for milnacipran, although further high-quality research is needed to be confident in these conclusions. Data for all other antidepressants were of low certainty. As RCTs excluded people with low mood, we were unable to establish the effects of antidepressants for people with chronic pain and depression. There is currently no reliable evidence for the long-term efficacy and safety of any antidepressant.

## Study registration

This study is registered as PROSPERO CRD42020171855.

## Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR128782) and is published in full in *Health Technology Assessment*; Vol. 28, No. 62. See the NIHR Funding and Awards website for further award information.



# Chapter 1 Background

Please note that this section duplicates some of the information published in the open-access full Cochrane Review (<https://doi.org/10.1002/14651858.CD014682.pub2>).

## Chronic pain

Chronic pain is defined as pain lasting or recurring for 3 months or longer.<sup>1</sup> It can be a primary condition or can occur in the context of a disease.<sup>2</sup> It is estimated that about one in five adults worldwide experience pain that is moderate or severe in its intensity and lasts 3 months or more;<sup>3</sup> however, estimates vary and may be higher. In the UK, reviews of chronic pain suggest that between one-third and one-half of the population experience chronic pain.<sup>4</sup> Some populations are more likely to experience chronic pain: older adults, women, people not in employment due to ill health and disability and people with comorbidities. Social circumstances are particularly influential; people in low socioeconomic circumstances are not only more likely to experience chronic pain, but also report higher levels of severity and disability.<sup>5</sup> Thus, chronic pain disproportionately affects the poor, facilitating social isolation and increasing all-cause morbidity<sup>6</sup> and mortality.<sup>7,8</sup> Almost one-third (30%) of people living with burdensome chronic pain struggle with productive engagement in society. Policy-makers have called for action.<sup>9</sup>

The impact of chronic pain is similar across conditions, despite the different aetiologies. Globally, chronic pain accounts for the highest number of years lived with disability, and affects individuals' daily lives, society and healthcare services.<sup>10,11</sup> Chronic pain accounts for up to one in five general practice consultations each year in Europe, Africa and Asia.<sup>12-14</sup> Chronic pain is also one of the global leading causes for sickness absence and people being unable to work.<sup>15,16</sup>

There are many different treatments aimed at reducing and managing chronic pain, including analgesic medication, physiotherapy, self-management guidance, exercise, psychological therapy, antidepressants, pain management clinics and surgery. The use of these depends upon the pain condition, severity of pain, individual characteristics, availability of services and national policy and guidelines. NHS England has called for better understanding of the risk of analgesic medication for pain conditions, especially when prescribed long-term. Although there are several non-pharmacological treatments aimed at living well with pain, for patients, the need to reduce pain remains a top priority.

Successful treatment of chronic pain can result in significant improvements in quality of life, including anxiety and depression.<sup>3,17,18</sup>

A systematic review identified that for people with fibromyalgia, reductions in pain intensity of 50% or more are associated with self-reports of sleep, fatigue and depression reverting back to normative values.<sup>3</sup> Therefore, efficacious treatment of the pain condition is essential for improvement of both pain and mood, in addition to potential improvements in sleep, physical function and quality of life. In addition, for many people, engaging effectively with physical exercise depends on reducing daily pain. Thus, effective reduction of pain remains an important aspect of treatment.

Pharmacological approaches directly target pain and are the main treatment available to first-line clinicians when faced with chronic disabling pain. Despite this, upon scrutiny of the evidence, a majority of common medicines have been removed from guideline recommendations for treatment of most chronic pain conditions, including paracetamol, non-steroidal anti-inflammatory drugs, opioids and synthetic cannabis.<sup>8-10</sup>

## Antidepressants and chronic pain

### *What are they, and how might they work?*

Antidepressants are medicines developed and used primarily for the treatment of clinical depression. A network meta-analysis (NMA) of the 21 most common antidepressants has shown that they are efficacious in the treatment of acute major depression, particularly severe depression.<sup>19</sup>

Antidepressants are grouped into different classes based on their chemical structure and presumed mechanism of action. The most common classes are as follows:

1. tricyclic antidepressants (TCAs): amitriptyline, desipramine, imipramine, nortriptyline and others
2. selective serotonin reuptake inhibitors (SSRIs): citalopram, sertraline, fluoxetine and others
3. serotonin norepinephrine reuptake inhibitors (SNRIs): duloxetine, levomilnacipran, milnacipran, venlafaxine

Antidepressants were originally developed to treat depression. Most antidepressants work by targeting monoamine neurotransmitters associated with mood and emotion and their receptors in the nervous system. These receptors, such as 5-hydroxytryptamine receptors, are activated by many neurotransmitters including serotonin, dopamine, adrenaline and noradrenaline.<sup>20</sup> Antidepressants prevent the neurotransmitters from being absorbed into neurons, which prolongs their activity in synapses. While the process by which antidepressants relieve depression is not fully understood, recent theories focus on neurochemical changes and neuroplasticity.<sup>20</sup>

Changes in the pain response systems travelling to and from the brainstem and involving the noradrenergic neurotransmitters have been theorised to explain the analgesic properties of antidepressants, and their proposed ability to reduce pain. By increasing the amount of serotonin and noradrenaline in the nervous system, pain signals are hypothesised to be blocked at the peripheral, spinal and supraspinal levels, reducing perceived pain, particularly in neuropathic pain.<sup>21,22</sup>

In addition, the locus coeruleus in the brain may have an analgesic effect on perceived pain.<sup>23</sup> Signals from this part of the brain are sent when the body reacts to a stimulus, such as pain, and noradrenaline is released into the dorsal horn in the spine to block receptors. Animal studies have shown that when pain signals are continuously received, as is the case in chronic pain, this analgesic response lessens over time, and noradrenaline is then not released.<sup>23,24</sup> However, when antidepressants are given, the analgesic response from the locus coeruleus is restored.<sup>23,25</sup>

### *Guidelines for antidepressants in the treatment of chronic pain*

Antidepressants are one of the few remaining recommended pharmacological interventions for chronic pain, although, to date, the evidence has not allowed the nuance of ranking the prioritisation. Where consideration of the quality of the supporting evidence is reported, it is often unclear.<sup>26</sup> Across guidelines from the USA,<sup>27</sup> Canada<sup>28</sup> and Japan,<sup>29</sup> TCAs (e.g. amitriptyline, nortriptyline) and SNRIs (e.g. duloxetine, milnacipran) are the most common classes recommended. In the UK, the National Institute for Health and Care Excellence (NICE) has produced different sets of recommendations for different pain types: chronic primary pain,<sup>30</sup> neuropathic pain,<sup>31</sup> low back pain and sciatica<sup>32</sup> and osteoarthritis.<sup>33</sup> For chronic primary pain, amitriptyline, citalopram, duloxetine, fluoxetine, paroxetine and sertraline are recommended equally. For neuropathic pain, amitriptyline or duloxetine are recommended. For low back pain and sciatica, the guidelines explicitly advise against use of SNRIs, SSRIs and TCAs; and for osteoarthritis, antidepressants are omitted entirely. The lack of concordance across guidelines can be confusing for clinicians, especially as many patients present with several types of pain concurrently. Furthermore, some of these recommendations are made based on very low-quality evidence. For example, in the chronic primary pain guidelines,<sup>9</sup> citalopram is recommended based upon one trial with 42 participants,<sup>34</sup> sertraline from one crossover trial with 14 participants,<sup>35</sup> and paroxetine from one trial with 46 participants.<sup>36</sup> These small trials are of very low scientific rigour and are not a reliable evidence

base. This is especially alarming given that within the context of the guidelines, antidepressants are the only pharmacological intervention recommended. Although the NICE guidelines for treatment of chronic low back pain published in 2016 recommend not prescribing antidepressants (regardless of mood), NICE guidelines for people with depression and chronic health problems recommend antidepressants in cases of mild depression and physical health problems. No advice exists to help practitioners resolve the contradiction between the two sets of advice.

### **Antidepressants in practice**

Prescriptions of antidepressants are relatively common in patients with chronic pain internationally; for example, 12.3% of people with chronic low back pain in Portugal report taking antidepressants for pain relief.<sup>37,38</sup>

In the UK, amitriptyline has long been the most commonly prescribed antidepressant for chronic pain. Amitriptyline was widely used to treat depression from the 1960s, but has restricted use now due to the risk of taking a lethal overdose. Its use is also characterised by side effects such as dizziness, dry mouth, constipation and weight gain; side effects are more common with higher dosages. Although amitriptyline is only licensed for the treatment of depression and neuropathic pain, it is commonly prescribed 'off licence' at a lower dose to treat any chronic pain. TCAs (of which the most common is amitriptyline) are the 19th most common prescription in primary care, accounting for 1.6% of all prescriptions.<sup>39</sup> Open-source prescribing data recorded over 14.5 million prescriptions for amitriptyline in 2021.<sup>40</sup> It is reasonable to believe that a majority of these prescriptions were for pain: amitriptyline is not recommended to treat depression,<sup>41</sup> and chronic pain is the most common indicator for antidepressant prescription in older adults.<sup>42</sup> A multinational comparison of antidepressant use in older adults found that TCAs were the most common class in the UK prescribed for chronic pain, at 55%. In comparison, SNRI prescriptions were very low (1.5%).<sup>41</sup> Indeed, there were 3.4 million duloxetine prescriptions in 2021, less than one-quarter of the number of amitriptyline prescriptions.

### **Antidepressants and safety**

There are also risks in the prescription of antidepressants. Adverse events such as dizziness, headache, nausea, ejaculation disorder, weight loss, tremor, sweating and insomnia have been found by randomised controlled trials (RCTs) to be more common in people taking antidepressants compared with those taking placebo.<sup>43,44</sup> Studies assessing the safety of antidepressants across a range of adverse outcomes in older people<sup>45,46</sup> and in people aged 20–64 years<sup>47,48</sup> have shown increased risks of falls and fractures associated with most antidepressants, and differences between antidepressants in risks of all-cause mortality, stroke and self-harm or suicide. Antidepressants also increase the risk of onset of seizures,<sup>49</sup> while the potential for gastrointestinal bleeding with SSRIs is widely recognised.<sup>50</sup> Long-term use of antidepressants for pain syndromes is therefore expected to be associated with harms at the population level.

### **The evidence on the efficacy and safety of antidepressants**

At the start of this study, there was no evidence comparing classes of antidepressants to each other in the management of chronic pain, as identified by the recent NICE guidelines.<sup>51</sup> There have been several systematic reviews for specific conditions (detailed below). Therefore, in the absence of any one RCT comparing the efficacy and safety of all antidepressants for chronic pain, a NMA was considered a priority to assess the relative effectiveness of each antidepressant, by dose.

Previous Cochrane Reviews investigated the efficacy of specific antidepressants in improving pain in specific conditions. A summary of their findings indicates that there is no high-quality evidence to support or refute the use of amitriptyline, milnacipran, nortriptyline, venlafaxine, desipramine or imipramine for management of neuropathic pain,<sup>52–57</sup> principally because trials are few and those that exist have small numbers of participants and typically have high risks of bias. The lack of evidence for some antidepressants stands in stark contradiction to guideline recommendations. For example, amitriptyline is recommended as a first-line treatment for neuropathic pain in primary care in guidelines for the UK, Canada and the International Association for the Study of Pain (IASP).<sup>31,58–60</sup>

For fibromyalgia, Cochrane Reviews of antidepressants show that there is no unbiased evidence that amitriptyline, desvenlafaxine, venlafaxine or SSRIs are superior to placebo.<sup>61,62</sup> There is low-quality evidence that duloxetine and milnacipran have some benefit in improving Patient Global Impression of Change (PGIC) scores and providing an improvement in pain relief of 30% or more, but no clinical benefit over placebo for improvement in pain relief of 50% or more, health-related quality of life or fatigue.<sup>62</sup> Similarly for mirtazapine, there is evidence for improvement in pain relief of 30% or more, and reduction of mean pain intensity and sleep problems, but this evidence is of low to medium quality, and there is no benefit for improvement in pain relief of 50% or more, PGIC, 20% improvement of health-related quality of life, reduction of fatigue or reduction in negative mood.<sup>63</sup>

Only one Cochrane Review has investigated the use of antidepressants for low back pain, and it found no clear evidence to support the use of any antidepressants.<sup>64</sup> A more recent systematic review supports these conclusions.<sup>65</sup> However, when analysed using the 'baseline observation carried forward' (BOCF) imputation method for missing data, pooled individual patient data analyses of RCTs have shown duloxetine and etoricoxib to be effective in reducing pain for pain conditions including chronic low back pain.<sup>3,63,66</sup> These distributions were bimodal: participants generally responded very well or very poorly, with few in between.<sup>3</sup>

The current systematic review and NMA allowed us to compare, for the first time, all antidepressants across all chronic pain conditions (bar headache), and identify whether certain classes or doses of antidepressants are useful in the management of pain and mood for people with chronic pain, and for certain chronic pain conditions. As antidepressants are also associated with a number of side effects, the review allowed the comparison of the proportion of adverse events occurring with the use of different antidepressants (including different classes of antidepressants, different types of antidepressants, and different dose regimens) within populations living with chronic pain.

### **The relationship between pain and low mood**

Although antidepressants are typically prescribed as analgesics for chronic pain patients, there is a strong possibility that if effective, they may also be effective at improving patient mood. The prevalence of depression in patients with chronic low back pain has been estimated as three to four times greater than that among the general population.<sup>67,68</sup> Distress and depression have been found to predict the transition to persistent pain states in several reviews.<sup>60,69,70</sup> It is clearly important to provide treatment that improves mood and quality of life in people living with pain. The most common intervention for patients with pain who also present with low mood is the prescription of antidepressant drugs, based on the assumption that these are effective in improving both pain and mood. Patients will often be told by clinicians that these drugs may help in a number of ways: they may have a direct effect on pain reduction, they may help by improving muscle relaxation, they can improve sleep, and they may help by improving mood. Which of these are of most benefit to the patient, if any, and in which patient group has not been established.

People suffering from depression and people living with chronic pain often report similar symptoms, such as low mood, lack of energy, difficulty making decisions and loss of pleasure from activities. Despite this, there appear to be some important differences between the two groups, which might imply that they require distinct interventions. For example, the content of depressive thoughts and the antecedents of feelings of sadness experienced by people in chronic pain may differ to those experienced by people with depression but without pain.<sup>71</sup> It is important to identify differences in pain-related distress (i.e. individuals with chronic pain experiencing low mood because of their pain) and clinical depression, which may reflect on the prevalence statistics reported above. The distinction between pain-related distress and depression is particularly important as primary care practitioners are often given contradictory guidance: they are encouraged to better detect depression,<sup>72,73</sup> while avoiding overmedicalisation of distress and thus overtreatment.<sup>74,75</sup> This is important as antidepressants can be

prescribed for the management of both pain and mood (e.g. clinical depression) in people with chronic pain. This review aimed to clarify this guidance as, unlike previous reviews in this area, we intended to investigate whether there were differences dependent upon whether the antidepressants were prescribed primarily to treat mood or pain.

## Patient and public involvement

We always involve our patient and public network in our research from conception and through the full cycle of research. For this study, in the first instance we met with five people (three females, two males) from our Research User Group at Royal Holloway, University of London, who had experience of chronic musculoskeletal pain and of NHS care for their condition. We discussed whether studying mood in chronic pain would have value, and whether we should know if antidepressants are effective. The group endorsed the general aim of the study strongly, and considered that the design was appropriate. The participants advised that they would personally never take antidepressants, although they had all been offered these at some time for pain or distress. We followed this with a second meeting with seven members of the Research User Group at Keele University. The group informed us that the study of low mood in people with chronic pain was very important, and that they considered the widespread use of antidepressants to be harmful. They wanted to know whether the cost of side effects associated with taking antidepressants outweighed any benefits associated with improving pain or mood. They saw this project as a possible first step in the development of a new, more effective intervention. For a description of the patient and public involvement (PPI) post review, see [Chapter 7](#).



## Chapter 2 Objectives

To assess the comparative efficacy and safety of antidepressants for adults with chronic pain using NMA. We aimed to inform on:

- the efficacy of antidepressants by type, class and dose in improving pain, mood, PGIC, physical functioning, sleep quality, and quality of life
- the safety of antidepressants prescribed for people with chronic pain; specifically, the number of adverse events associated with antidepressants by type, class and dose.



## Chapter 3 Methods

Please note that this section duplicates the information published in the open-access full Cochrane Review (<https://doi.org/10.1002/14651858.CD014682.pub2>).

### Criteria for considering studies for this review

#### *Types of studies*

We included RCTs that compared any antidepressant with any comparator. RCTs are the best design to minimise bias when evaluating the effectiveness of an intervention. We followed the guidance in the *Cochrane Handbook for Systematic Reviews of Interventions* for the inclusion of crossover RCTs, which requires inclusion of this type of study unless there is a justifiable reason not to.<sup>76</sup> The risk in this review was that washout periods between the periods of the study would not be long enough for carry-over effects from the antidepressants or comparators to be sufficiently minimised. Therefore, we only included crossover trials with washout periods of at least five times the length of the antidepressant half-life (this was calculated individually for each antidepressant).

The most common comparators we anticipated finding in the literature were as follows: the same antidepressant at a different dose; a different antidepressant; placebo (both active and inert); other medications for pain management purposes (e.g. pregabalin, gabapentin); analgesics; psychological therapy (e.g. cognitive behavioural therapy, acceptance and commitment therapy); exercise; physiotherapy; multidisciplinary pain programmes; herbal medicines and nutraceuticals (e.g. St John's Wort); and acupuncture. Where the comparator was a placebo, antidepressant, analgesic or other medication for pain management purposes, these trials were required to be double-blind. We included trials examining any dose of antidepressants, with a study duration of at least 2 weeks and minimum of 10 participants per arm. We excluded non-randomised studies, case reports, experimental studies, clinical observations and prevention studies.

#### *Types of participants*

We included adults (aged 18 years or older) reporting primary or secondary pain in any part of their body (except headache) as their primary complaint, that matched the IASP definition of chronic pain (i.e. at least 3 months' duration).<sup>1</sup> We included all trials regardless of the severity of participants' chronic pain, although we extracted whether severity was part of the inclusion criteria of the individual studies. We excluded studies where the participants' primary complaint was headache or migraine, as had been performed in previous Cochrane Reviews.<sup>77</sup> Although this condition does fit within the IASP criteria, the diagnosis, classification and treatment of primary and secondary headache are often different from those of other pain conditions, and clinical trials are primarily aimed at prevention of further headaches or migraines rather than symptomatic treatment. We included participants with multiple health conditions as long as the chronic pain condition was the focus of the trial.

#### *Types of interventions*

##### **Decision set**

We included any antidepressant at any dose, for any indication, but used primarily for treatment of people with chronic pain and compared to placebo or active intervention. We included antidepressants grouped into the following classes.

- TCAs: amitriptyline, clomipramine, imipramine, trimipramine, doxepin, desipramine, protriptyline, nortriptyline, dothiepin, lofepramine and others
- SSRIs: fluvoxamine, fluoxetine, paroxetine, sertraline, citalopram, escitalopram, zimelidine and others
- SNRIs: venlafaxine, milnacipram, duloxetine and others

- Monoamine oxidase inhibitors (MAOIs):
  - irreversible: phenelzine, tranylcipromine, izocarboxazid and others
  - reversible: brofaramine, moclobemide, Tyrima and others
- Other antidepressants:
  - noradrenaline reuptake inhibitors (NARIs): reboxetine, atomoxetine and others
  - noradrenaline and dopamine reuptake inhibitors (NDRIs): amineptine, bupropion and others
  - noradrenergic and specific serotonergic antidepressants (NaSSAs) including tetracyclic antidepressants (TeCAs) such as mirtazapine, mianserin, maprotiline and others
  - serotonin antagonist and reuptake inhibitors (SARIs): trazodone and others
  - unclassified: agomelatine, vilazodone and others.

We categorised doses of included antidepressants into low, standard and high doses. These are displayed in [Table 1](#). As the majority of antidepressants are not licensed for pain, judgements were made based on the recommendations of daily doses for clinical depression in the British National Formulary.<sup>78</sup> The judgements were made by clinical authors of the review; initially by the clinical pharmacist and then approved by discussion with a psychiatrist and anaesthetist. Standard doses were the recommended doses for depression in adults. Low doses were those listed as initial doses (where a standard range is specified), the dose for elderly patients or any dose below the standard dose (where no range was specified). High doses were those listed at the upper range of standard dose ranges, or above the standard dose where no range is specified. Where trials included flexible dosing across multiple categories and did not report mean dose, these were labelled as 'unable to be categorised'.

### **Supplementary sets**

We included studies with any active comparator. We included studies where the antidepressant was combined with another intervention, as long as there was an arm solely for the other intervention so we were able to isolate the effects of the antidepressant (e.g. antidepressant + drug vs. drug). We did not include combination trials where there was no way to isolate the effects of an antidepressant (e.g. antidepressant A + drug vs. antidepressant B). For this review we assumed that any participant who met the inclusion criteria was, in principle, equally likely to be randomised to any of the eligible antidepressants; however, we acknowledge there may have been differences in patients' expectations of treatment and outcomes depending upon which antidepressant was studied.

### **Types of outcome measures**

We anticipated that there would be a variety of outcome measures used throughout the literature. Due to the distinction between distress and depression discussed above, this review used the term 'mood' as an outcome, to include depression that is diagnosed, mood that is measured via self-report, and distress.

For pain and mood, where applicable we also dichotomised outcomes into pain relief or improvement of 50% or greater, in line with the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) guidance, to indicate substantial improvement.<sup>79</sup> Where possible, we planned separate NMAs to compare antidepressants to the comparators immediately post intervention, at short-term follow-up ( $\leq 12$  weeks post trial) and at long-term follow-up ( $> 12$  weeks post trial). Where studies included multiple follow-up time points, we took the most recent time point within each period. If multiple measures were used for the same outcome (e.g. for continuous pain intensity, both a 0–10 numerical rating scale and the McGill Pain Questionnaire were reported), then we extracted from the most valid, reliable, and widely used measure in the field.

### **Primary outcomes**

- Substantial pain relief: the proportion of participants (number and percentage of total and per arm) reporting at least 50% reduction in pain intensity from baseline, irrespective of pain measurement method [e.g. visual analogue scale (VAS), numerical rating scale]

TABLE 1 Antidepressant dose categorisation

Antidepressant	Total daily dosage		
	Low	Standard	High
Amitriptyline	< 25 mg	25–75 mg	> 75 mg
Bupropion	n/a <sup>a</sup>	150–300 mg	> 300 mg
Citalopram	< 20 mg	20 mg	40 mg
Clomipramine	< 30 mg	30–150 mg	> 150 mg
Desipramine	< 100 mg	100–200 mg	> 200 mg
Desvenlafaxine	n/a <sup>b</sup>	50 mg	> 50 mg
Dothiepin (dosulepin)	< 75 mg	75–150 mg	> 150 mg
Doxepin	< 75 mg	75–150 mg	> 150 mg
Duloxetine	< 60 mg	60 mg	> 60 mg
Escitalopram	< 10 mg	10 mg	20 mg
Esreboxetine	n/a <sup>c</sup>	4–8 mg	> 8 mg
Fluoxetine	< 20 mg	20–40 mg	> 40 mg
Imipramine	< 75 mg	75–150 mg	> 150 mg
Nortriptyline	< 75 mg	75–100 mg	> 100 mg
Maprotiline	150 mg	300 mg	> 300 mg
Mianserin	< 30 mg	30–40 mg	> 40 mg
Milnacipran	< 100 mg	100 mg	> 100 mg
Mirtazapine	< 30 mg	30 mg	> 30 mg
Moclobemide	150 mg	300 mg	600 mg
Paroxetine	< 20 mg	20 mg	50 mg
Pirlindole	< 225 mg	225–300 mg	> 300 mg
Reboxetine	< 8 mg	8 mg	> 8 mg
Sertraline	n/a <sup>d</sup>	50 mg	> 50 mg
Trazodone	< 150 mg	150–300 mg	> 300 mg
Trimipramine	< 75 mg	75–150 mg	> 150 mg
Venlafaxine	< 75 mg	75–150 mg	> 150 mg
Zimelidine	< 300 mg	300 mg	> 300 mg

a Lowest dose form is 150 mg.

b Desvenlafaxine is not available in the UK; lowest dose form is 50 mg.

c Esreboxetine is not available in the UK, and no doses lower than 4 mg have been used in trials.

d 50 mg is both the initial and standard dose; no recommendations of lower doses in the British National Formulary.

- Pain intensity: continuous data from any measures of pain intensity or severity (e.g. VAS or validated measures such as Brief Pain Inventory)
- Mood: continuous data from any measures of mood (e.g. VAS, Hospital Anxiety and Depression Scale)
- Adverse events: the proportion of participants (number of percentage of total and per arm) reporting adverse events.

### Secondary outcomes

- Moderate pain relief: the proportion of participants (number and percentage of total and per arm) reporting at least 30% reduction in pain intensity from baseline, irrespective of pain measurement method (e.g. VAS, numerical rating scale)
- Physical function: continuous data from any measures of physical movement and disability (e.g. numerical rating scale, Short Form questionnaire-36 items (SF-36) Physical Component Score)
- Sleep: continuous data from any measures of quality of sleep, including insomnia, restfulness, and so on (e.g. Brief Pain Inventory, Jenkins Sleep Scale)
- Quality of life: continuous data from any measure of quality of life (e.g. numerical rating scale, EuroQol-5 Dimensions)
- PGIC: the proportion of participants (number and percentage of total and per arm) reporting 'much' and 'very much' improved on the PGIC scale, and continuous data from the PGIC scale
- Serious adverse events: the proportion of participants (number and percentage of total and per arm) reporting serious adverse events
- Withdrawal: the proportion of participants (number and percentage of total and per arm) withdrawing for any reason.

### Search methods for identification of studies

This search was last run on 4 January 2022.

#### *Electronic searches*

We searched the following databases, without language restrictions.

- The Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Library – Issue 12 of 12 2021
- MEDLINE and MEDLINE In-Process (via OVID) – 1946 to 4 January 2022
- EMBASE (via OVID) – 1974 to 4 January 2022
- CINAHL (via EBSCO) – 1981 to December 2021
- LILACS (via Birme – December 2021)
- PsycINFO (via EBSCO) – 1872 to 4 January 2022
- AMED (via OVID) – 1985 to December 2021.

We tailored searches to individual databases. The search strategies used can be found in [Appendix 1](#). The search strategy was developed by the Cochrane Pain, Palliative and Supportive Care (PaPaS) Review Group's information specialist and was independently peer reviewed. The PaPaS information specialist performed the searches.

#### *Searching other resources*

We searched ClinicalTrials.gov ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)) and the World Health Organization's International Clinical Trials Registry Platform ([apps.who.int/trialsearch/](http://apps.who.int/trialsearch/)) for unpublished and ongoing trials. In addition, we searched grey literature, checked reference lists of reviews and retrieved articles for additional studies, and performed citation searches on key articles. We contacted study authors for additional information where necessary.

### Data collection and analysis

#### *Selection of studies*

Two review authors (HB and CF) independently determined eligibility of each study identified by the search. Independent review authors eliminated studies that clearly did not satisfy inclusion criteria, and obtained full copies of the remaining studies. HB and CF read these studies independently to select

relevant studies, and in the event of a disagreement, third and fourth authors adjudicated (TP and CE). We did not anonymise the studies in any way before assessment. We have included a PRISMA flow chart which shows the status of identified studies,<sup>80</sup> as recommended in the *Cochrane Handbook*.<sup>76</sup> We included studies in the review irrespective of whether measured outcome data were reported in a 'useable' way. We recorded reasons for exclusion of any ineligible studies at the full-text stage.

### **Data extraction and management**

Two review authors (HB and CF) independently extracted data using a standard piloted form and checked for agreement before entry into Review Manager 5.4.<sup>81</sup> In the event of disagreement, third and fourth authors (TP and CE) adjudicated. We collated multiple reports of the same study, so that each study rather than each report was the unit of interest in the review. We collected characteristics of the included studies in sufficient detail to populate the table of 'Characteristics of included studies'. We extracted the following information:

- Study design: authors, publication year and journal, duration, sponsorship, conflicts of interest, aim (pain or emotional functioning), trial design, number of treatment arms, setting, missing data methods, power calculation used, definition of chronic pain, minimum level of pain for entry, inclusion and exclusion criteria
- Participant characteristics: overall number, number in each arm, withdrawal (total, per arm and by sex), type of participant, chronic pain conditions, sex, age, baseline differences
- Intervention: type of antidepressant, class, dose (freeform and dichotomised), route of administration, duration
- Comparator(s): type (e.g. placebo, psychological therapy), description (if placebo medication: active or inert, appearance, taste, smell, titration, number of tablets), type and class (if other antidepressant), doses, route of administration, length, intensity (if physical or psychological comparator)
- Outcomes (data from all time points reported in the study): domain (e.g. pain, physical functioning), measure, measure validation, baseline data, results for each time point, effect sizes
- Adverse events and withdrawals (proportion overall and per arm): any, serious, withdrawal due to adverse event, withdrawal due to lack of efficacy.

### **Assessment of risk of bias in included studies**

Two review authors (HB and CF) independently assessed risk of bias for each study, using the criteria outlined in the *Cochrane Handbook*,<sup>81</sup> with any disagreements resolved by discussion. We completed a 'risk of bias' table for each included study using the Cochrane 'risk of bias' tool version 1.0 in Review Manager 5.4.<sup>81</sup>

We assessed the following for each study:

- Random sequence generation (checking for possible selection bias):
  - We assessed the method used to generate the allocation sequence as being at low risk of bias (any truly random process, e.g. random number table, computer random number generator) or unclear risk of bias (method used to generate sequence not clearly stated).
  - We excluded studies using a non-random process (e.g. odd or even date of birth, hospital or clinic record number).
- Allocation concealment (checking for possible selection bias). The method used to conceal allocation to interventions prior to assignment determines whether intervention allocation could have been foreseen in advance of or during recruitment, or changed after assignment.
  - We assessed the methods as being at low risk of bias (e.g. telephone or central randomisation; consecutively numbered, sealed, opaque envelopes) or unclear risk of bias (method not clearly stated).
  - We excluded studies that did not conceal allocation (e.g. open list).
- Blinding of participants and personnel (checking for possible performance bias). Due to the inclusion of trials using any comparator, our review contained both double-blinded RCTs and those studies

in which double-blinding was not possible (i.e. RCTs of psychological therapy or acupuncture). In the RCTs that were double-blinded, we assessed the methods used to blind study participants and personnel from knowledge of which intervention a participant received.

- We assessed methods as being at low risk of bias (the study states that it was blinded and describes the method used to achieve blinding, such as identical tablets matched in appearance or smell, or a double-dummy technique) or unclear risk of bias (the study states that it was blinded but does not provide an adequate description of how this was achieved).
- Studies in which double-blinding was not possible due to the comparator were considered to have high risk of bias.
- Blinding of outcome assessment (checking for possible detection bias). We assessed the methods used to blind study participants and outcome assessors from knowledge of which intervention a participant received. We assessed the methods as being at:
  - low risk of bias (the study has a clear statement that outcome assessors were unaware of treatment allocation, and ideally describes how this was achieved)
  - unclear risk of bias (the study states that outcome assessors were blind to treatment allocation but it lacks a clear statement on how this was achieved)
  - high risk of bias (the outcome assessment was not blinded)
- Selective reporting (checking for reporting bias). We assessed whether primary and secondary outcome measures were prespecified and whether these were consistent with those reported. We assessed the methods as being at:
  - low risk of bias (study protocol is available with prespecified measures)
  - unclear risk of bias (insufficient information available to permit a judgement of high or low risk of bias)
  - or high risk of bias [not all of the study's prespecified primary outcomes have been reported; one or more primary outcomes have been reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not prespecified; one or more reported primary outcomes were not prespecified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); one or more outcomes of interest in the review have been reported incompletely so that they cannot be entered in a meta-analysis; the study report failed to include results for a key outcome that would be expected to have been reported for such a study]
- Incomplete outcome data (checking for possible attrition bias due to the amount, nature and handling of incomplete outcome data). We assessed the methods used to deal with incomplete data as being at:
  - low risk of bias (no missing outcome data; reasons for missing outcome data are unlikely to be related to the true outcome; missing outcome data are balanced in numbers across intervention groups, with similar reasons for missing data across groups; missing data have been imputed using BOCF analysis)
  - unclear risk of bias [insufficient reporting of attrition/exclusions to permit a judgement of low or high risk of bias (e.g. number randomised not stated; no reasons for missing data provided; or the study did not address this outcome)]
  - high risk of bias (the reason for missing outcome data is likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; 'as-treated' analysis was done with substantial departure of the intervention received from that assigned at randomisation; potentially inappropriate application of simple imputation; use of 'last observation carried forward' (LOCF) without the addition of any other low risk of bias methods)
- Other bias. We assessed any other potential sources of bias that were not included in the other domains.

We considered studies to be at high risk of bias overall if they met the criteria for high risk of bias in any of the above domains.

### ***Measures of treatment effect***

For the outcomes measuring continuous data (pain intensity, mood, physical function, sleep, quality of life and PGIC continuous), data were reported as either post-intervention scores (the mean scores at the end

of the intervention period) or change scores (mean change from baseline score). We conducted separate analyses for these. As is common in pain management studies, for all outcomes (apart from PGIC) a broad range of scales were used to measure the outcomes. Therefore, once data were extracted, they were converted into standardised mean difference (SMD) with 95% confidence intervals (CIs). We interpreted SMD as small (0.2), moderate (0.5) and large (0.8), in line with Cohen<sup>82</sup> and the *Cochrane Handbook*.<sup>76</sup> For outcomes with dichotomous data (substantial pain relief, adverse events, moderate pain relief, PGIC much/very much improved, serious adverse events and withdrawal), we used odds ratios (ORs) with 95% CIs.

### Unit of analysis issues

For most RCTs, we did not encounter any unit of analysis complexities as trial participants were randomised to different study arms, allowing direct analysis. For crossover RCTs, if the results for the first period (prior to crossover) were reported, we extracted these in an attempt to avoid crossover effects. If the results from the first period were not reported then we extracted the final trial results, provided there was a sufficient washout period of at least five times the length of the antidepressant half-life (minimum washout period length calculated separately for each antidepressant). The majority of crossover studies reported the combined effects of both periods (only one study reported first-period and second-period effects separately); therefore, we analysed crossover trials using these combined effects. Our search did not return any cluster RCTs that met our inclusion criteria.

### Dealing with missing data

For all missing study-level statistical data relevant to our outcomes, we first tried to contact the authors of the study. If we could not get the data from the authors, then we followed the guidance from the *Cochrane Handbook*.<sup>76</sup> If standard deviations were missing, then we used the Review Manager calculator to calculate these from other data reported in the study. We did not impute any data, but assessed each study's risk of bias due to missing data.

### Assessment of heterogeneity

We assessed heterogeneity within the NMAs using the Tau statistic, in line with the guidance in the *Cochrane Handbook*.<sup>76</sup> We assessed heterogeneity using Confidence in Network Meta-Analysis (CINeMA) software, which calculated the chi-squared test and the  $I^2$  statistic for each pairwise comparison on each outcome. As outlined in the *Cochrane Handbook*, we interpreted the  $I^2$  statistic as follows:<sup>76</sup>

- 0–40%: might not be important
- 30–50%: may represent moderate heterogeneity
- 50–90%: may represent substantial heterogeneity
- 75–100%: considerable heterogeneity.

We took into account the magnitude and strength of effects when assessing heterogeneity.

### Assessment of the transitivity assumption

We carefully scrutinised transitivity, which is the key underlying assumption of NMA. Transitivity requires studies to be similar on average across all factors that might alter treatment effects other than the intervention comparison being made.<sup>76</sup> To address this, we only included studies with similar clinical populations (i.e. participants reporting pain lasting at least 3 months).<sup>83</sup> Previous research, combined with review authors' clinical experience and knowledge, identified variables that could potentially influence our primary outcome:

- pain condition
- age
- pain intensity at baseline
- depressive severity at baseline
- treatment duration
- dosing schedule.

We explored the impact of these factors by assessing the indirectness of the network. The inclusion of placebo and concerns about its potential to violate the transitivity assumption have been highlighted in general,<sup>84</sup> and particularly in depression studies.<sup>85</sup> Therefore, we compared placebo-controlled studies with those that provided head-to-head evidence as a form of validation of the network.

### **Assessment of reporting biases**

We assessed reporting biases using the Cochrane 'risk of bias' tool version 1.0 in Review Manager 5.4<sup>81</sup> by checking for study protocols and prespecified outcomes (as detailed in *Assessment of risk of bias* section). We also used funnel plots for pairwise analyses for antidepressants where more than 10 studies were available, as advised in the *Cochrane Handbook*.<sup>76</sup> Funnel plots were drawn using the 'Risk Of Bias due to Missing Evidence in Network meta-analysis' (ROB-MEN) tool, which is part of CINeMA, and used to assess the significant small study effects via funnel plot asymmetry.

### **Data synthesis**

We undertook separate NMAs for each outcome. NMAs combine information (evidence) from both direct comparisons of interventions within RCTs and indirect comparisons across trials based on a common placebo comparator.<sup>86,87</sup> Direct comparisons (direct evidence) occur when two or more interventions are compared head to head in a trial; in the absence of head-to-head comparisons, interventions can be indirectly compared (indirect evidence).

We analysed the data for all primary and secondary outcomes using Bayesian random-effects NMAs implemented using the R ([r-project.org](http://r-project.org)) package *multinma* (The R Foundation for Statistical Computing, Vienna, Austria) [Phillippo DM. *multinma: Bayesian Network Meta-Analysis of Individual and Aggregate Data*. Version R package version 0.4.2. 2022. <https://doi.org/10.5281/zenodo.3904454>]. Where dose was included in the network, doses were categorised (low, standard, high) and incorporated as separate nodes. Where a study had multiple arms investigating different doses of the same antidepressant that fell into the same category (e.g. two different low doses), we did not combine these; by using the *multinma* package, we were able to keep these as separate arms in the analysis.

We fitted random-effects models using broad normal prior distributions for the treatment effects and study-specific intercepts and a half-normal prior for the heterogeneity standard deviation. We used four chains each with 2000 iterations and 1000 post-warmup draws per chain. Convergence was assessed using potential scale reduction factors and effective sample size.

We explored network connectivity via network plots. In the network plot, for treatment-only models, the nodes represent each intervention. In treatment-dose models, the antidepressant nodes represent the antidepressant and dose (low, standard, high). The colour of the node represents the antidepressant class, and the 'nonad' label refers to all interventions that were not an antidepressant. The size of each node represents the combined sample size of participants from all studies investigating that intervention, and the thickness of the lines represents the number of studies for that comparison. The forest plots present the estimates and credible intervals (CrIs) for each intervention in the network, with reference to placebo.

We assessed convergence using the potential scale reduction factor for each parameter, ensured that effective sample sizes were sufficiently large,<sup>88</sup> and verified that there were no divergent transitions.<sup>89</sup> We explored heterogeneity by fitting connected networks for treatment, treatment-dose, class, risk of bias, and condition where network geometry allowed sufficient connectivity.<sup>90</sup>

We assessed model fit using mean residual deviance, and explored inconsistency through unrelated mean effects (UME) models and node-splitting where network geometry allowed.<sup>91</sup> We reported effect

estimates and cumulative posterior ranks of effect alongside strength of evidence assessment using Grading of Recommendations, Assessment, Development and Evaluations (GRADE).

To rank the treatments for each outcome by probability of best treatment, we used the surface under the cumulative ranking curve (SUCRA) and the mean ranks. We reported relative effects and mean rank of treatments and plotted cumulative rankograms showing the range of rankings of different treatments for each outcome.

We used the deviance information criterion (DIC) to compare the different models for reporting (treatment only, treatment–dose, class, and change score and post-intervention studies for contrast-based models) to assess their parsimony. Substantive differences in DIC ( $> 5$ ) or models with marginally lower DIC but lower Tau and fewer studies with residual deviance  $> 3$  in combination were deemed superior. We selected models to report on the basis of parsimony, minimisation of inconsistency (identified via UME and node-splitting models), residual deviance and heterogeneity (measured as Tau). This approach balanced clinical exploration of results and the risk of overfitting.<sup>90</sup>

NMA, UME and node-splitting models were implemented in multinma in R (version 4.1.3). Further details of the modelling framework are described by Phillippo.<sup>77,92</sup>

### **Subgroup analysis and investigation of heterogeneity**

Where data allowed, we performed subgroup analyses for the class of antidepressant and the type of pain condition. We used a Bayesian random-effects NMA to account for expected heterogeneity and variation in the data. These methods allowed the uncertainty inherent in the between-study variance component to be reflected in effect estimate precision. We performed these subgroup analyses by building separate models; however, this was dependent on the geometry and connectedness of the networks.

Due to sparsity of data, we were unable to perform subgroup analyses on the aim of the trial (whether the trial targeted pain or mood) or on baseline levels of mood. Upon examination, the average scores for the five most commonly used scales for mood (Beck Depression Inventory, Brief Pain Inventory Mood Item, SF-36 Mental Component Score, SF-36 Mental Health Subscale, and Hamilton Depression Rating Scale) were all in the none/minimal ranges for depression.

### **Sensitivity analysis**

Analysis by risk of bias judgement (high and not high) was only possible for substantial pain relief. We were unable to perform sensitivity analyses for any outcome comparing active placebo to inert placebo, as in total only nine studies used an active placebo.

## **Summary of findings and assessment of the certainty of the evidence**

To assess the certainty of the NMA, we primarily used the CINeMA framework.<sup>93</sup> In contrast to the NMAs in this review, which were conducted within a Bayesian framework, CINeMA operates within a frequentist framework using the netmeta package in R [Rücker G, Schwarzer G, Krahn U, König J. netmeta: Network Meta-Analysis Using Frequentist Methods. Version R package version 0.9-5. 2017. URL: <https://cran.r-project.org/package=netmeta>]. The CINeMA framework considers the impact of certain issues within NMAs on clinical decision-making utilising the results. This framework is based on GRADE, and considers the following six domains specific to NMA:<sup>93</sup>

1. *Within-study bias (impact of risk of bias in the included studies)* CINeMA assesses the impact of risk of bias by combining the study's risk of bias (as judged by the reviewers using a risk of bias tool) with its contribution to the NMA.

2. *Reporting bias (publication and other reporting biases)* Reporting bias in CINeMA is categorised as either 'suspected' or 'undetected'. Suspected reporting bias is when the review methods do not take into account unpublished data, the meta-analysis is based on a small number of positive early findings, or treatments are exclusively studied in industry-funded trials. Undetected reporting bias is when data from unpublished studies have been identified and the findings agree, when prospective trial registration has been completed and there are no deviations from protocols, and comparisons of estimates between small and large studies agree.
3. *Indirectness (relevance to the research question, addressing transitivity)* Each study in the NMA is evaluated according to its relevance to the research question. Study-level judgements are combined with the percentage contribution of the study to the network. This approach assesses potential transitivity issues in the NMA.
4. *Imprecision (the precision of the NMA, by combining direct with indirect evidence)* Relevant treatment effects that represent a minimal clinically important difference (MCID) are defined and the range of clinical equivalence is produced (the value of the MCID either side of the line of no effect). CINeMA then compares the treatment effects included in the 95% CI to the range of clinical equivalence. If the 95% CI of a treatment effect crosses the range of clinical equivalence, then it is considered to have major concerns of imprecision. If the 95% CI of a treatment effect only crosses one side of the range of equivalence, then there are no concerns of imprecision.
5. *Heterogeneity (variability in the results of studies)* CINeMA accounts for heterogeneity between studies by comparing the confidence and prediction intervals of a treatment effect. When confidence and prediction intervals indicate the same effect, then there is no evidence of heterogeneity; conversely, if a prediction interval leads to a different conclusion than the CIs, then there is evidence of heterogeneity.
6. *Incoherence (agreement between the results of direct and indirect evidence)* This is the variation between direct and indirect evidence in the network and also an assessment of transitivity. CINeMA compares the 95% CIs of the direct and indirect estimates. If both of these estimates lie on the same side of the range of clinical equivalence, then there are no concerns about incoherence.

The CINeMA framework results in the reviewers summarising the judgements across the domains into the four domains of GRADE (high certainty, moderate certainty, low certainty, very low certainty).

For outcomes where we were unable to use CINeMA due to the complexity of the network (adverse events, serious adverse events, and withdrawal), we used GRADE. The GRADE system considers the following five considerations to assess the certainty of the body of evidence for each outcome:

- serious or very serious study limitations (risk of bias)
- important or serious inconsistency of results
- some or major indirectness of evidence
- serious or very serious imprecision
- probability of publication bias.

The GRADE system results in the assignment of one of the following grades to the evidence:

- High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
- Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
- Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.

Two authors (HB and GS) independently interpreted the findings, and collaboratively made the final judgements across all outcomes. To present our findings, we have produced separate 'Summary of findings' tables for all outcomes. We have used the template 'Summary of findings' tables designed for NMA.<sup>94</sup> Due to the scale of the analyses, we only included studies of antidepressants which had  $\geq 200$  participants in total receiving the antidepressant in the write-ups and Summary of findings tables. This decision was made to ensure quality and certainty of the final results and conclusions. We based this decision on reference to the tiers of evidence for pain research; tier 2 uses data from at least 200 participants.<sup>95</sup>



## Chapter 4 Results

Please note that this section duplicates the information published in the open-access full Cochrane Review (<https://doi.org/10.1002/14651858.CD014682.pub2>).

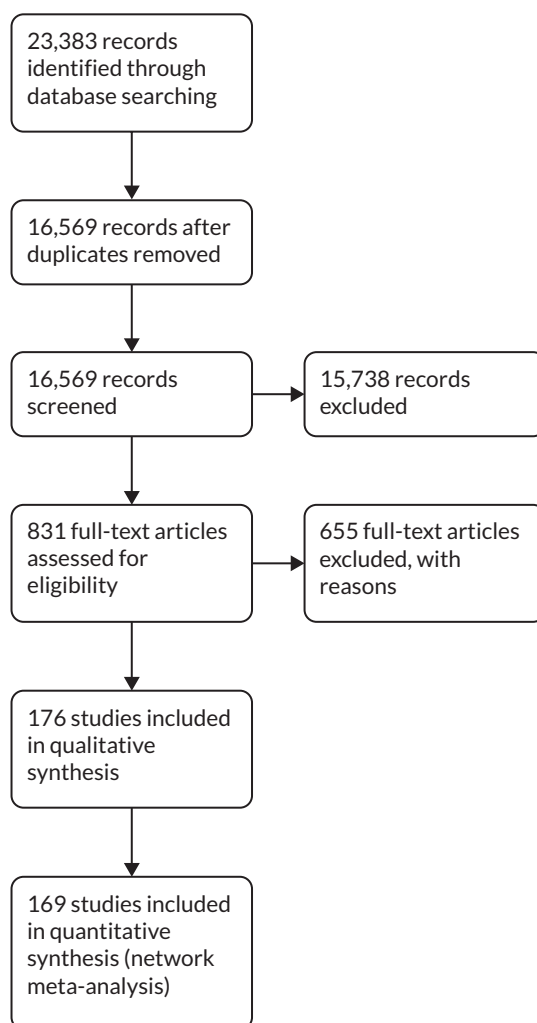
### Description of studies

#### Results of the search

We ran the original search on 6 May 2020, and the top-up search on 4 January 2022. Both searches searched six databases and [www.clinicaltrials.gov](http://www.clinicaltrials.gov). The original search returned 21,569 records, and the top-up search returned 1814 records for a total of 23,383. After removing duplicates, we screened 16,569 records at title and abstract. From this, we excluded 15,738 records, leaving 831 records at full text. After full-text screening, we included 176 studies. The study flow diagram is presented in [Figure 1](#).

#### Included studies

In total, we included 176 studies in the review, with a total of 28,664 adult participants with a mean age of 50.6 years. A table of included studies is provided in [Appendix 4](#).



**FIGURE 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

## RESULTS

There were a variety of study designs across trials:

- antidepressant versus placebo: 83 (e.g. Hudson)<sup>96</sup>
- antidepressant versus active comparator: 22 (e.g. Enomoto)<sup>97</sup>
- antidepressant versus the same antidepressant at different doses versus placebo: 17 (e.g. Arnold)<sup>98</sup>
- antidepressant versus active comparator versus combined antidepressant + active comparator: 13 (e.g. Ang)<sup>99</sup>
- antidepressant versus active comparator versus placebo: 9 (e.g. Rowbotham)<sup>100</sup>
- antidepressant versus different antidepressant: 9 (e.g. Kaur)<sup>101</sup>
- antidepressant versus active comparator versus combined antidepressant + active comparator versus placebo: 8 (e.g. Gilron)<sup>102</sup>
- antidepressant versus different antidepressant versus placebo: 7 (e.g. Heymann)<sup>103</sup>
- antidepressant versus different antidepressant versus active comparator: 4 (e.g. Boyle)<sup>104</sup>
- antidepressant versus the same antidepressant at different doses: 2 (e.g. Chappell)<sup>105</sup>
- antidepressant versus same antidepressant at different doses versus different antidepressant at different doses versus placebo: 1<sup>106</sup>
- antidepressant versus different antidepressant versus combined antidepressants versus placebo: 1<sup>107</sup>

Most studies had a parallel-arm design (141 studies) compared to a crossover design (35 studies).

Studies mainly included participants with only one type of chronic pain:

- 59 studies included fibromyalgia.
- 49 studies included neuropathic pain.
- 40 studies included musculoskeletal pain.
- 9 studies included primary pain syndromes (not including fibromyalgia) for example described only as 'somatoform' or 'idiopathic' pain.
- 6 studies included gastrointestinal pain.
- 4 studies included non-cardiac chest pain.
- 2 studies included burning mouth syndrome.
- 2 studies included visceral pain.
- 1 study included atypical facial pain.
- 1 study included phantom limb pain.
- 1 study included pelvic pain.

Most studies were funded by pharmaceutical companies:

- 72 studies were fully funded by pharmaceutical companies.
- 5 were partially funded by pharmaceutical companies.
- 67 studies were funded through non-pharmaceutical means, mainly government, charity or institutional funding.
- 32 studies did not report the source of funding.

Most studies had a primary aim of reducing pain:

- 144 studies had a primary aim of reducing pain.
- 2 studies had a primary aim of treating depression.
- 6 studies had a primary aim of treating both depression and pain.
- 14 studies had other primary aims (e.g. sleep, other symptoms).

Studies ranged in length from 2 weeks to 9 months, with an average length of 10 weeks. Only six studies followed up with participants after the trial finished.<sup>108-113</sup> The follow-up time points ranged from 4 weeks post trial to 1 year post trial. Seven studies, with a total of 156 participants, provided no useable data and were therefore omitted from the NMAs.<sup>106,114-119</sup>

Of the 176 studies and 28,664 participants, the numbers of participants receiving each antidepressant (not including combined interventions) were as follows:

- amitriptyline: 1843 (43 studies)
- bupropion: 54 (1 study)
- citalopram: 97 (5 studies)
- clomipramine: 124 (2 studies)
- desipramine: 336 (7 studies)
- desvenlafaxine: 884 (2 studies)
- dothiepin: 55 (3 studies)
- doxepin: 30 (2 studies)
- duloxetine: 6362 (43 studies)
- escitalopram: 93 (3 studies)
- esreboxetine: 978 (2 studies)
- fluoxetine: 277 (11 studies)
- imipramine: 300 (7 studies)
- maprotiline: 135 (4 studies)
- mianserin: 107 (2 studies)
- milnacipran: 3110 (18 studies)
- mirtazapine: 255 (2 studies)
- moclobemide: 42 (1 study)
- nortriptyline: 374 (7 studies)
- paroxetine: 422 (9 studies)
- pirlindole: 50 (1 study)
- reboxetine: 18 (1 study)
- sertraline: 91 (3 studies)
- trazodone: 63 (3 studies)
- trimipramine: 18 (1 study)
- venlafaxine: 489 (8 studies)
- zimeldine: 10 (1 study)

In total, 9854 participants received a placebo across 130 studies.

## Excluded studies

We excluded a total of 655 references with reasons throughout the course of this review. The main reasons for exclusion were as follows:

- duplicate records (including trial registrations): 144 records
- not chronic pain condition: 71 records
- not accessible (primarily conference abstracts): 92 records
- pooled analysis: 50 records
- open-label: 42 records
- fewer than 10 participants per arm: 22 records
- single-blind: 15 records
- washout period not at least five times the antidepressant half-life: 11.

Reasons for exclusion other than these are reported in the 'Characteristics of excluded studies' section of the full Cochrane Review.

We categorised 15 studies as 'awaiting classification' due to uncertainties regarding blinding or pain duration, and there are 26 studies identified as ongoing; these are reported in the full Cochrane Review.

### Risk of bias in included studies

Risk of bias findings from the included studies by domain are shown in [Figure 2](#). To see risk of bias findings by study, please see the full Cochrane Review ([www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD014682.pub2/full#CD014682-fig-0003](http://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD014682.pub2/full#CD014682-fig-0003)). Overall, we rated 116 of 176 studies as 'high risk' and 60 as 'not high risk'. However, of the 60 studies not rated as high risk, 29 had three or more domains rated as 'unclear'.

### Allocation

We did not assess any studies as at high risk of bias for sequence generation or allocation concealment. For sequence generation, we judged 95 studies to be of low risk, and 81 studies were judged as unclear. For allocation concealment, we judged 75 studies to have satisfactory procedures and rated them as low risk, and the other 101 studies were rated as unclear. We rated only 64 studies as at low risk of bias for both sequence generation and allocation concealment.

### Blinding

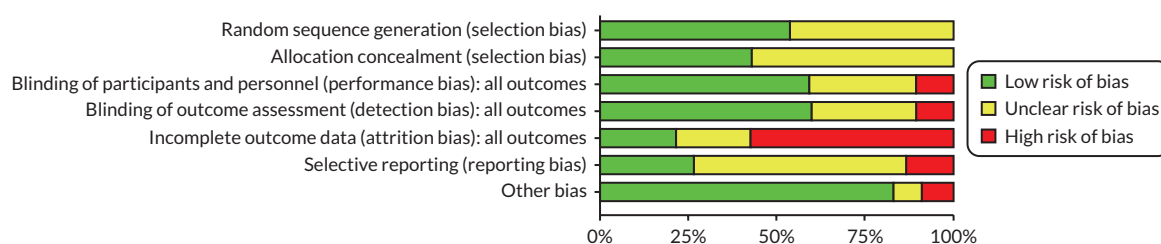
For this review, we required studies comparing antidepressants with other antidepressants, different doses of the same antidepressant, or other pharmacological interventions to be double-blind. We accepted that some interventions could not be blinded by their nature (e.g. psychological therapy, physiotherapy). These studies were included but judged to be at high risk of bias for both blinding of participants and blinding of outcomes assessors. Seventeen studies were of non-pharmacological interventions and therefore rated as at high risk of bias for both domains. As this review is focused on pain, all outcomes were self-reported by participants, and therefore judgements were often the same for both domains. In total, we rated 106 studies as at low risk for both domains and 49 studies as unclear for both domains. Low risk of bias was achieved in studies by study drugs appearing identical, having matched/sham dosing schedules across all arms, and using active placebos that mimic the side effects of antidepressants.

### Incomplete outcome data

We rated the majority of studies as at high risk of bias for incomplete outcome data; 102 studies were high risk. Studies were high risk primarily due to only using the LOCF imputation method, reporting data only on participants who completed the trial, or having significantly unequal attrition across arms. We rated 37 studies as at low risk of bias; these studies either had no/very little attrition, or used appropriate imputation methods such as BOCF or multiple imputation. We rated 37 studies as unclear due to them not clearly specifying missing data methods.

### Selective reporting

We could not find protocols or trial registrations for the majority of studies. We rated 108 studies as having an unclear risk of bias, due to missing protocols or trial registrations being published retrospectively, after the study had begun. We rated 44 studies as at low risk of bias; outcomes and analyses in the published papers matched prospective protocols or registrations. We rated 24 studies as at high risk of bias. Four of these studies were never published in journal articles, and data were extracted from trial registries.<sup>112,120-122</sup> For the other studies rated as at high risk of bias, there were



**FIGURE 2** Risk of bias of included studies by domain.

discrepancies between the protocols and published papers that were judged to result in a significant risk of bias (e.g. protocol stated that outcomes would be collected that were not reported).

### **Other potential sources of bias**

We did not identify any other sources of bias for 145 studies. We rated 17 studies as having an unclear risk of bias, primarily due to data not being presented in numerical form or being reported by a different method to that in the protocol (e.g. percentage change rather than post intervention). We rated 14 studies as at high risk of bias for the following reasons:

- no published, peer-reviewed articles<sup>112,120-122</sup>
- washout periods and tapering issues<sup>102,123</sup>
- poor reporting with mistakes in article<sup>124</sup>
- insufficient power<sup>125</sup>
- significant differences at baseline<sup>126</sup>
- selection bias prior to participation<sup>36</sup>
- significant differences between published article and trial registry<sup>127,128</sup>
- using a potential intervention as a placebo.<sup>113</sup>

We found some evidence of publication bias in one analysis (duloxetine vs. placebo for substantial pain relief), as identified from funnel plots (used to assess small study effects as a proxy for publication bias).

## **Effects of interventions**

### **Overview**

The following sections detail the results of the NMAs for all outcomes included in the review. Due to the scale of the analysis, we only include studies of antidepressants with  $\geq 200$  participants in the write-ups and Summary of findings tables. Each outcome has a table listing all the interventions included in the NMA. Antidepressant studies with  $< 200$  participants, and non-antidepressant interventions, are also included in figures for completeness and context.

For all outcomes, we made decisions on which networks to report in this results section. For all outcomes, we considered treatment and treatment-dose networks. For continuous outcomes, we considered both change scores and post-intervention scores networks. For each outcome, we have reported the most robust and reliable network. The details of these decisions are reported in [Appendix 2](#). The networks that we have not reported in this manuscript are available in [Report Supplementary Material 1](#).

The sections are reported in order of primary and secondary outcomes ([Tables 2-13](#)).

Primary outcomes:

- substantial pain relief
- pain intensity
- mood
- adverse events.

Secondary outcomes:

- moderate pain relief
- physical function
- sleep
- quality of life
- PGIC: proportion of participants reporting much/very much improved, and continuous scores

## RESULTS

- serious adverse events
- withdrawal.

### Primary outcomes

#### Substantial pain relief (50% reduction)

**TABLE 2** Substantial pain relief: summary of findings

Estimates of effects, CrIs and certainty of the evidence for substantial pain relief in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: desvenlafaxine high dose ( $\geq 50$ mg); duloxetine low dose ( $< 60$ mg), standard dose (60 mg), and high dose ( $> 60$ mg); esreboxetine standard dose (4–8 mg) and high dose ( $\geq 8$ mg); milnacipran standard dose (100 mg) and high dose ( $> 100$ mg); mirtazapine standard dose (30 mg)							
Comparator (reference): placebo							
Outcome: substantial pain relief ( $\geq 50\%$ reduction in pain intensity from baseline) as measured on various scales including 0–10 VAS, 0–100 VAS, and the Brief Pain Inventory							
Direction: higher is better (i.e. more people reporting substantial pain relief)							
Total RCTs: 42 Total participants: 14,626	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (CINeMA)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Duloxetine standard dose RCTs: 16 Participants: 4490	1.91 (1.69 to 2.17)	592/2061 287 per 1000	1058/2429 435 per 1000	148 more per 1000	Moderate <sup>c</sup>	8 (5 to 12)	Equivalent to NNTB of 7.1
Duloxetine high dose RCTs: 14 Participants: 3692	1.91 (1.66 to 2.21)	431/1855 232 per 1000	674/1837 366 per 1000	134 more per 1000	Moderate <sup>c</sup>	8 (5 to 12)	Equivalent to NNTB of 7.4
Milnacipran high dose RCTs: 1 Participants: 384	1.64 (1.04 to 2.58)	38/145 262 per 1000	88/239 368 per 1000	106 more per 1000	Very low <sup>c,d</sup>	11 (4 to 19)	Equivalent to NNTB of 9.4
Esreboxetine standard dose RCTs: 1 Participants: 828	1.72 (1.13 to 2.62)	33/275 120 per 1000	105/553 190 per 1000	70 more per 1000	Low <sup>c</sup>	11 (4 to 19)	Equivalent to NNTB of 14
Milnacipran standard dose RCTs: 2 Participants: 1298	1.65 (1.28 to 2.13)	130/654 199 per 1000	187/644 290 per 1000	91 more per 1000	Low <sup>c,e</sup>	12 (6 to 18)	Equivalent to NNTB of 11
Mirtazapine standard dose RCTs: 1 Participants: 211	1.30 (0.79 to 2.15)	33/211 156 per 1000	41/211 194 per 1000	39 more per 1000	Low <sup>g</sup>	15 (6 to 21)	Not significantly different from placebo
Duloxetine low dose RCTs: 6 Participants: 1116	1.71 (1.36 to 2.20)	150/523 287 per 1000	242/593 407 per 1000	120 more per 1000	Moderate <sup>c,d,e</sup>	16 (11 to 20)	Equivalent to NNTB of 8.3
Esreboxetine high dose RCTs: 1 Participants: 555	1.29 (0.79 to 2.11)	33/275 120 per 1000	42/280 150 per 1000	30 more per 1000	Very low <sup>c,d</sup>	16 (7 to 22)	Not significantly different from placebo
Desvenlafaxine high dose RCTs: 2 Participants: 870	1.19 (0.83 to 1.70)	51/215 237 per 1000	177/655 270 per 1000	33 more per 1000	Very low <sup>c,d</sup>	17 (11 to 21)	Not significantly different from placebo

**TABLE 2** Substantial pain relief: summary of findings (*continued*)

NNTB, number needed to treat for an additional beneficial outcome.  
 Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

- Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.
- Mean rank and Crls are presented.
- Downgraded due to within-study bias.
- Downgraded due to imprecision in the estimate.
- Downgraded due to heterogeneity in the estimate.
- Downgraded due to incoherence in the network.
- Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

**Notes:**  
 CINeMA grades of confidence in the evidence:  
 High: further research is unlikely to change our confidence in the estimate of effect.  
 Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
 Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
 Very low: we are very uncertain about the estimate.

## Results

We report the treatment–dose network for substantial pain relief, as it was the model with the least heterogeneity and had no evidence of inconsistency.

We included 42 RCTs with a total of 14,626 participants (range in study from 47 to 1108). There were 25 different interventions, and some comparisons were informed only by direct evidence from one trial. We could not include data from two trials due to disconnected networks.

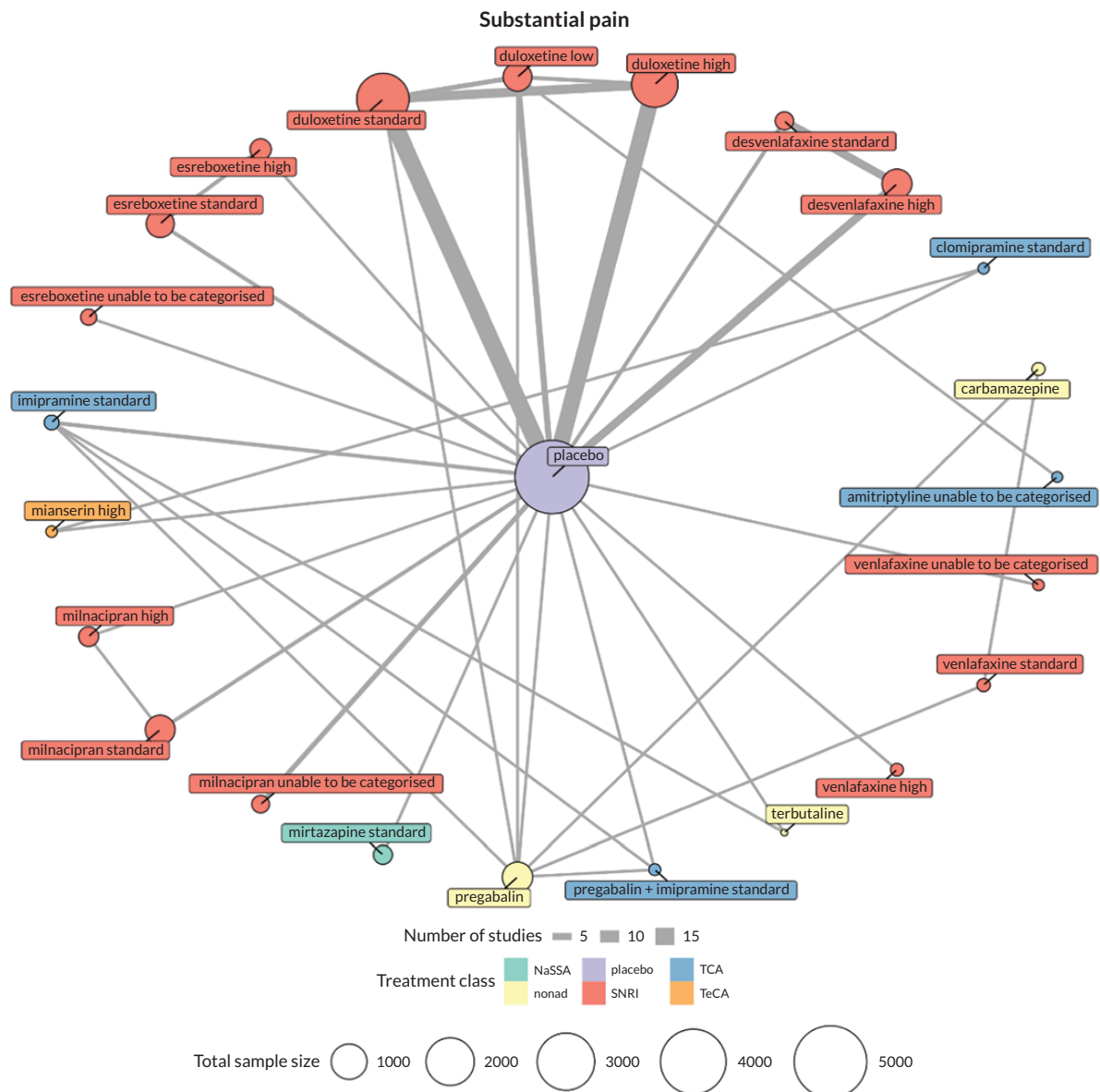
The antidepressants with  $\geq 200$  participants and therefore included in the summary are as follows:

- desvenlafaxine high dose
- duloxetine low dose
- duloxetine standard dose
- duloxetine high dose
- esreboxetine standard dose
- esreboxetine high dose
- milnacipran standard dose
- milnacipran high dose
- mirtazapine standard dose.

An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 15](#).

There were no concerns regarding model fit based on residual deviance and convergence diagnostics. The network diagram is presented in [Figure 3](#) and the forest plot in [Figure 4](#).

Duloxetine standard dose and duloxetine high dose were the highest-ranked antidepressants for substantial pain relief, and equally efficacious in comparison to placebo (OR 1.91, 95% CI 1.69 to 2.17 and OR 1.91, 95% CI 1.66 to 2.21, respectively). Milnacipran high dose (OR 1.64, 95% CI 1.04 to 2.58) and esreboxetine standard dose (OR 1.72, 95% CI 1.13 to 2.62) were also equally ranked, but less effective than duloxetine standard dose and duloxetine high dose. Mirtazapine standard dose, esreboxetine high dose and desvenlafaxine high dose showed no significant difference in comparison to placebo.



**FIGURE 3** Substantial pain relief: network plot.

A visual representation of the cumulative rankings for every treatment included in the analysis did not substantially alter the interpretation of relative effects or mean rank CrIs. The UME model had similar DICs to the dose treatment model, with no evidence of inconsistency in the dev-dev plot. We confirmed this with node-splitting models for all nine comparisons where it was possible to compare direct and indirect evidence. The comparison of pregabalin with placebo had the smallest Bayesian  $p$ -value ( $p = 0.3$ ) indicative of inconsistency where direct evidence suggests underestimation of the effect of pregabalin based on a single trial. These figures are available in [Report Supplementary Material 1](#). The availability of a consistent evidence network precluded the need for exploration of transitivity violations.

### Exploration of heterogeneity

Despite the risk of overfitting, we summarise results for multiple models because of the importance of substantial pain as an outcome for patients, clinicians and overall quality of life. The full results of all models are reported in [Report Supplementary Material 1](#).

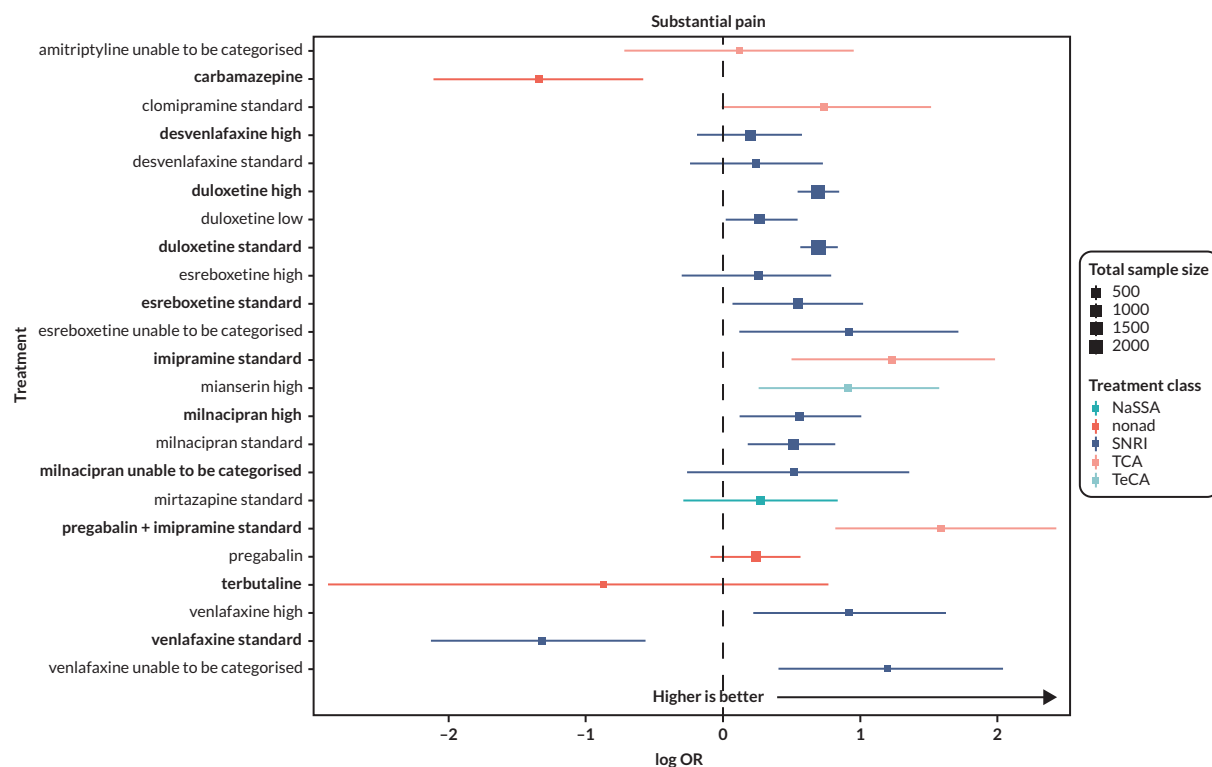


FIGURE 4 Substantial pain relief: summary forest plot (log OR with CrIs).

### Class

We generated a network by aggregating treatment into classes. The analysis included four antidepressant classes: SNRI, TCA, TeCA and NaSSA; however, we could not draw any reliable conclusions about class differences due to inconsistency and overlapping CrIs.

### Condition

Trials reporting substantial pain included neuropathic, fibromyalgia, musculoskeletal, primary and gastrointestinal pain conditions. However, only neuropathic and fibromyalgia pain conditions had connected networks. We could not derive reliable treatment rankings for neuropathic pain, as the UME models and node-splitting indicated inconsistency. For fibromyalgia, although the network geometry precluded analysis of inconsistency, esreboxetine, milnacipran and duloxetine were relatively equally ranked: esreboxetine (mean rank = 2.02, 97.5% CrI 1 to 4); milnacipran (mean rank = 2.30, 97.5% CrI 1 to 4); duloxetine (mean rank = 2.48, 97.5% CrI 1 to 4).

### Risk of bias

We conducted a sensitivity analysis to explore the effect of removing high-risk-of-bias studies. We rated 15 studies as having low risk of bias. The model of the resulting network was unstable with divergent transitions indicating problems with model convergence. UME models and the dev-dev plot did not identify inconsistency, but we could not confirm this by node-splitting due to network geometry. Results were consistent with the treatment-dose model. The two best-ranked antidepressants were esreboxetine (mean rank = 3.73, 97.5% CrI 2 to 7) and duloxetine (mean rank = 4.64, 97.5% CrI 3 to 6).

### Confidence in network meta-analysis

In addition to fitting multiple models to explore heterogeneity and utilising UME and node-splitting models to explore inconsistency, we undertook further analysis of pairwise direct evidence and

## RESULTS

network evidence (excluding multiarm trials of dose) to facilitate strength of evidence assessment using CINeMA.

The design-by-treatment test showed no inconsistency between direct and indirect evidence ( $\chi^2 = 14.069$ ,  $p = 0.296$ ), although duloxetine low dose and desvenlafaxine high dose had high  $I^2$  values (73.6% and 65.8%), indicating heterogeneity. We rated duloxetine low, standard and high doses as moderate certainty. We rated all other antidepressant doses as low, or very low certainty primarily due to major concerns regarding high-risk-of-bias studies, imprecision (estimates crossing zero), and a small number of RCTs and participants contributing to the estimates.

### Pain intensity

**TABLE 3** Pain intensity: summary of findings

Estimates of effects, CrIs and certainty of the evidence for pain intensity in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine low dose (< 60 mg), standard dose (60 mg), and high dose (> 60 mg); milnacipran standard dose (100 mg) and high dose (> 100 mg)							
Comparator (reference): placebo							
Outcome: change in pain intensity, as measured on multiple scales including 0–10 VAS, 0–100 VAS, Brief Pain Inventory, and the Short-form McGill Pain Questionnaire							
Direction: lower is better (i.e. a greater reduction in pain intensity)							
Total RCTs: 50 Total participants: 14,926	Relative effect	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Duloxetine high dose RCTs: 14 Participants: 3683	-	-	-	SMD -0.37 (-0.45 to -0.28)	Low <sup>c,d</sup>	9 (8 to 13)	Small to moderate effect
Duloxetine standard dose RCTs: 18 Participants: 4959	-	-	-	SMD -0.31 (-0.39 to -0.24)	Moderate <sup>d</sup>	11 (10 to 15)	Small to moderate effect
Milnacipran high dose RCTs: 2 Participants: 1670	-	-	-	SMD -0.22 (-0.40 to -0.05)	Low <sup>c,e</sup>	14 (12 to 19)	Small effect
Milnacipran standard dose RCTs: 4 Participants: 1866	-	-	-	SMD -0.22 (-0.39 to -0.06)	Moderate <sup>c,d</sup>	14 (12 to 20)	Small effect
Duloxetine low dose RCTs: 6 Participants: 1104	-	-	-	SMD -0.11 (-0.25 to 0.03)	Moderate <sup>c,e</sup>	17 (12 to 21)	Not significantly different from placebo

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Mean rank and CrIs are presented.

b SMD interpretation based on clinical judgement and in line with Cohen 1988<sup>82</sup> and the *Cochrane Handbook for Systematic Reviews of Interventions* as small (0.2), moderate (0.5) and large (0.8).

c Downgraded due to within-study bias.

d Downgraded due to imprecision in the estimate.

e Downgraded due to heterogeneity in the estimate.

f Downgraded due to incoherence in the network.

g Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

**TABLE 3** Pain intensity: summary of findings (*continued*)**Notes:**

CINeMA grades of confidence in the evidence:

High: further research is unlikely to change our confidence in the estimate of effect.

Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low: we are very uncertain about the estimate.

**Results**

For pain intensity, we report the change score treatment–dose network, as it was more robust than the other networks with low heterogeneity and no indications of inconsistency.

We included 49 RCTs with a total of 14,504 participants (range 26–1191). We removed one study from this analysis due to implausible results.<sup>129</sup> Of these, 28 studies compared against placebo, 9 were trials with a head-to-head comparison versus another active comparator and 12 were dose-comparison trials. There were 21 different interventions, and some comparisons were informed only by direct evidence from one trial.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are as follows:

- duloxetine low dose
- duloxetine standard dose
- duloxetine high dose
- milnacipran standard dose
- milnacipran high dose.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3, Table 16](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 5](#) and the forest plot in [Figure 6](#).

**Ranking of antidepressants**

Duloxetine high and standard dose were the highest-ranked antidepressants for pain intensity, with small to moderate effects (SMD  $-0.37$ , 95% CI  $-0.45$  to  $-0.28$  and SMD  $-0.31$ , 95% CI  $-0.39$  to  $-0.24$  respectively). Milnacipran high and standard doses had a small effect (SMD  $-0.22$ , 95% CI  $-0.40$  to  $-0.05$ ). Duloxetine low dose showed no significant difference in comparison to placebo.

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretation. The UME model had similar DICs to the dose treatment model, with no evidence of inconsistency. We confirmed this with node-splitting models for all nine comparisons where it was possible to compare direct and indirect evidence. The lowest Bayesian  $p$ -value was for the comparison of duloxetine standard dose versus duloxetine high dose ( $p = 0.08$ ). These figures are available in [Report Supplementary Material 1](#).

**Condition and risk of bias**

We were unable to undertake further NMAs of condition or risk of bias due to small sample sizes, network geometry and the risk of overfitting, but these were examined in pairwise analyses and network analysis (excluding multidose arms) in CINeMA to inform strength of evidence assessment.

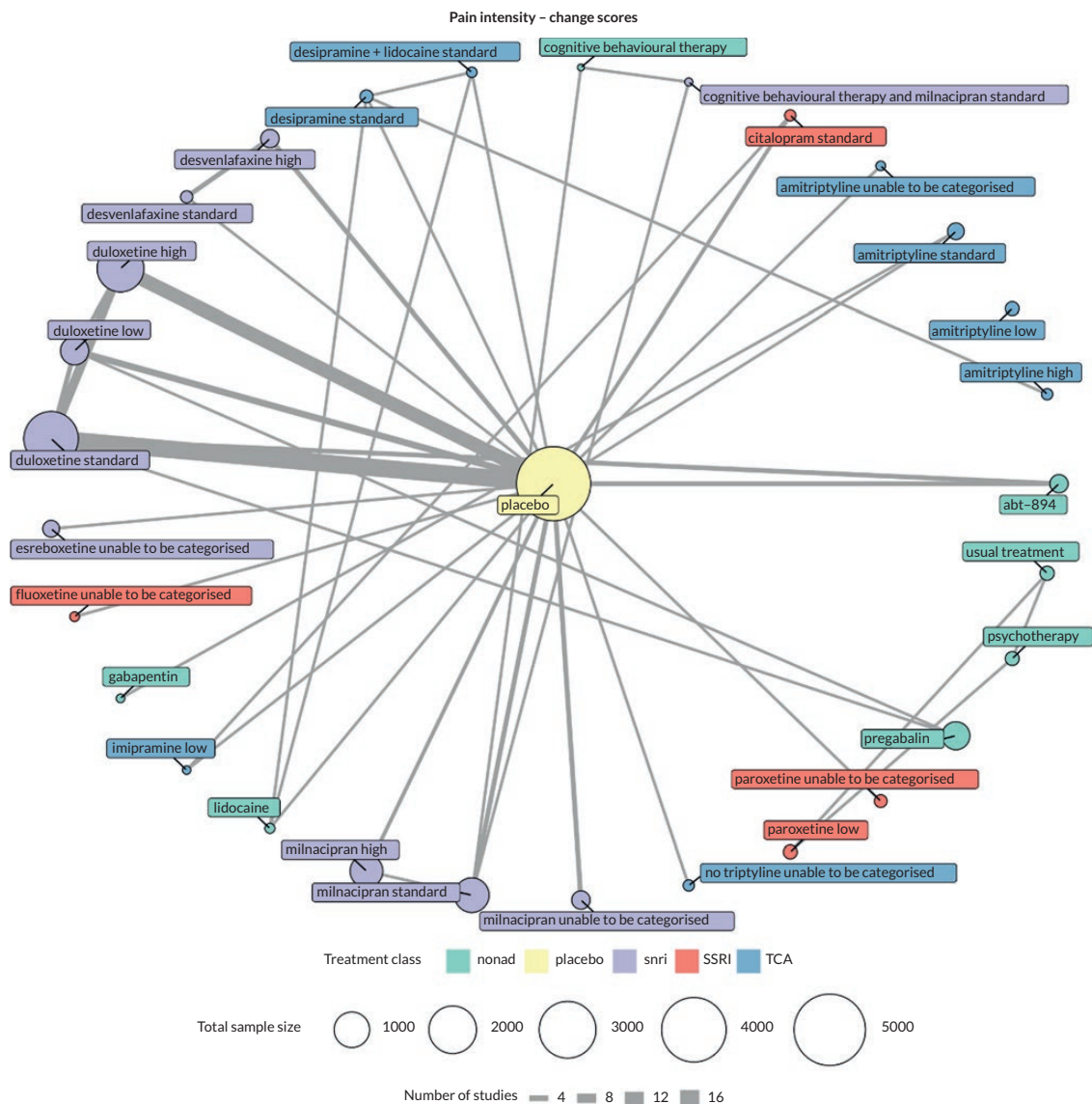


FIGURE 5 Pain intensity: network diagram.

**Confidence in network meta-analysis**

The design-by-treatment test showed no inconsistency between direct and indirect evidence ( $\chi^2 = 8.34, p = 0.82$ ), although duloxetine standard dose and milnacipran standard dose had high  $I^2$  values (65.3% and 67.7%), indicating heterogeneity. We had moderate certainty in the estimates for duloxetine low and standard and milnacipran standard doses. We rated all other antidepressant doses as low certainty due to major concerns regarding high-risk-of-bias studies and imprecision (estimates crossing zero).

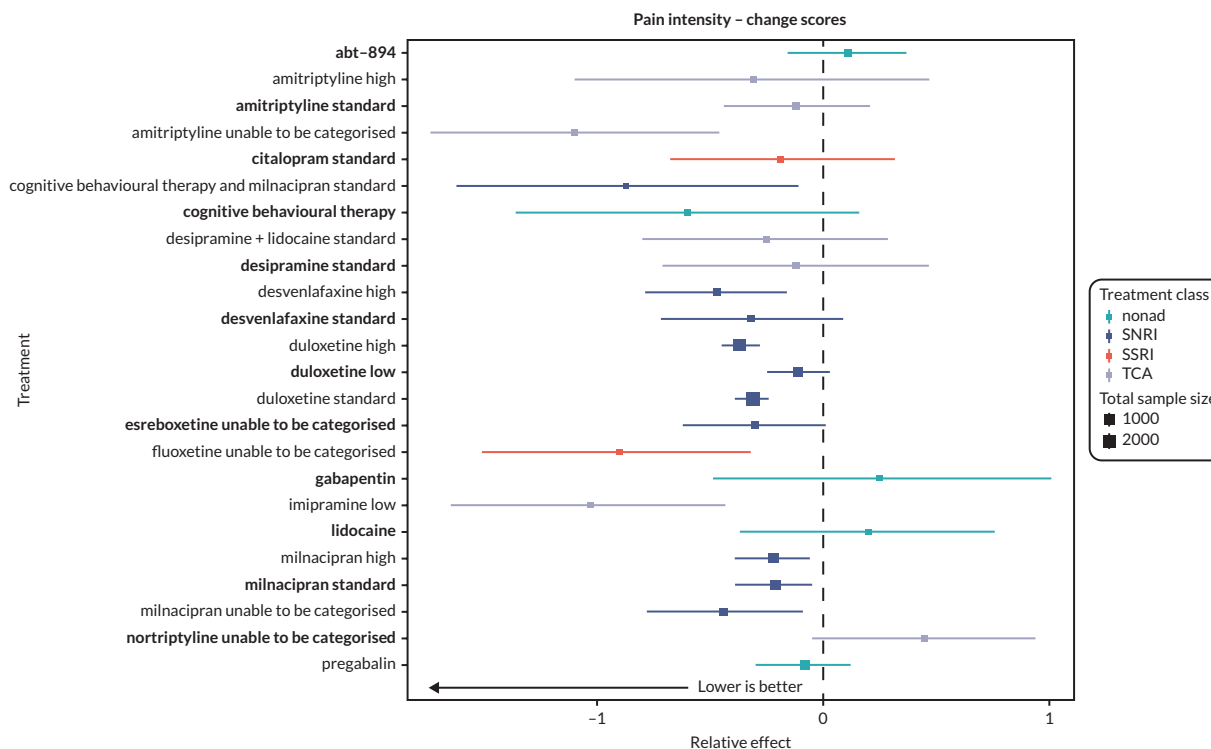


FIGURE 6 Pain intensity: summary forest plot (SMD with CrIs).

### Mood

TABLE 4 Mood: summary of findings

Estimates of effects, CrIs and certainty of the evidence for mood in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine (all doses combined), milnacipran (all doses combined), mirtazapine (all doses combined)							
Comparator (reference): placebo							
Outcome: change in mood (depression, anxiety, distress) scores as measured on various scales including the Beck Depression Inventory, Backache Index, SF-36 Mental Component Score and the SF-36 Mental Health Subscale							
Direction: lower is better (i.e. a greater reduction of distress, depression or anxiety)							
Total RCTs: 38 Total participants: 12,985	Relative effect	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Mirtazapine RCTs: 1 Participants: 406	-	-	-	SMD -0.5 (-0.78 to -0.22)	Low <sup>c</sup>	4 (2 to 7)	Moderate effect
Duloxetine RCTs: 26 Participants: 7952	-	-	-	SMD -0.16 (-0.22 to -0.1)	Moderate <sup>d</sup>	8 (5 to 11)	Small effect
Milnacipran RCTs: 5 Participants: 3109	-	-	-	SMD -0.13 (-0.26 to 0.01)	Moderate <sup>d,e</sup>	9 (5 to 13)	Not significantly different from placebo

**TABLE 4** Mood: summary of findings (*continued*)

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Mean rank and CrIs are presented.

b SMD interpretation based on clinical judgement and in line with Cohen 1988<sup>82</sup> and the *Cochrane Handbook for Systematic Reviews of Interventions* as small (0.2), moderate (0.5) and large (0.8).

c Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

d Downgraded due to within-study bias.

e Downgraded due to heterogeneity in the estimate.

f Downgraded due to imprecision in the estimate.

g Downgraded due to incoherence in the network.

**Notes:**

CINeMA grades of confidence in the evidence:

High: further research is unlikely to change our confidence in the estimate of effect.

Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low: we are very uncertain about the estimate.

## Results

For mood, we report the change score treatment network as this was the most robust and reliable network, with low heterogeneity and no indications of inconsistency.

We included 38 RCTs with a total of 12,985 participants (range 42–1191). Of these, 22 trials compared against placebo only, 6 were multiarm trials with another active comparator, 9 were comparing the same antidepressant in different doses, and 1 compared two antidepressants together. There were 16 different interventions, and some comparisons were informed only by direct evidence from one trial. We rated 23 trials as having a high risk of bias. We could not include data from one trial due to disconnected networks.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

- duloxetine
- milnacipran
- mirtazapine.

An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 17](#).

At baseline, the average scores for the five most commonly used scales (Beck Depression Inventory, Brief Pain Inventory Mood Item, SF-36 Mental Component Score, SF-36 Mental Health Subscale, and Hamilton Depression Rating Scale) were all in the none/minimal ranges.

There were no concerns regarding model fit. The network diagram is presented in [Figure 7](#) and the forest plot is presented in [Figure 8](#).

### Ranking of antidepressants

Mirtazapine was the highest-ranked antidepressant for mood with a moderate effect (SMD  $-0.5$ , 95% CI  $-0.78$  to  $-0.22$ ), based on one RCT. Duloxetine and milnacipran were equally ranked. Duloxetine showed very small effects (SMD  $-0.16$ , 95% CI  $-0.22$  to  $-0.1$ ), and milnacipran showed no difference in comparison to placebo.

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretation of the results. This figure is available in [Report Supplementary Material 1](#). The UME model had similar DICs to the dose treatment model, with no evidence of inconsistency.

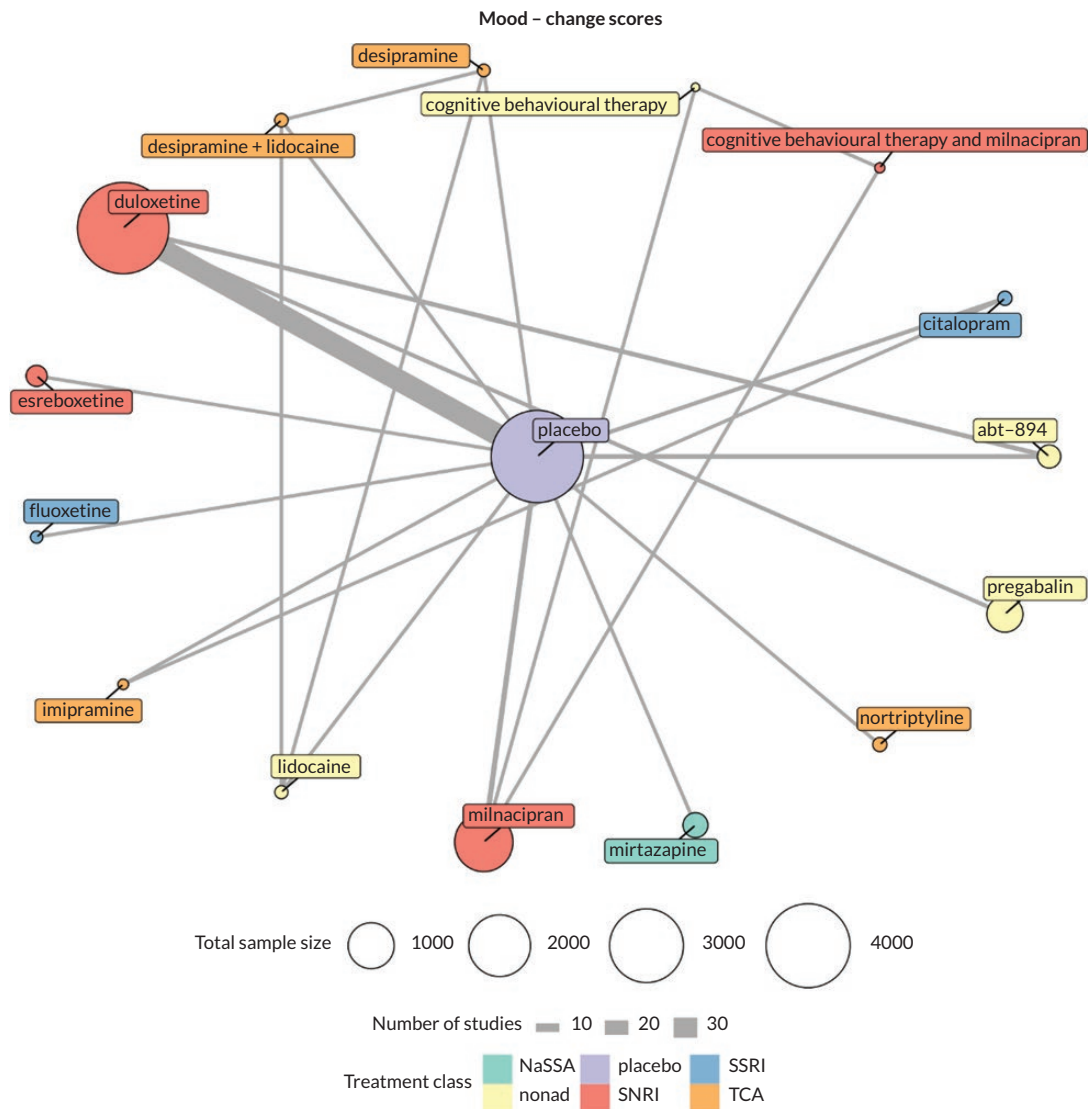


FIGURE 7 Mood: network plot.

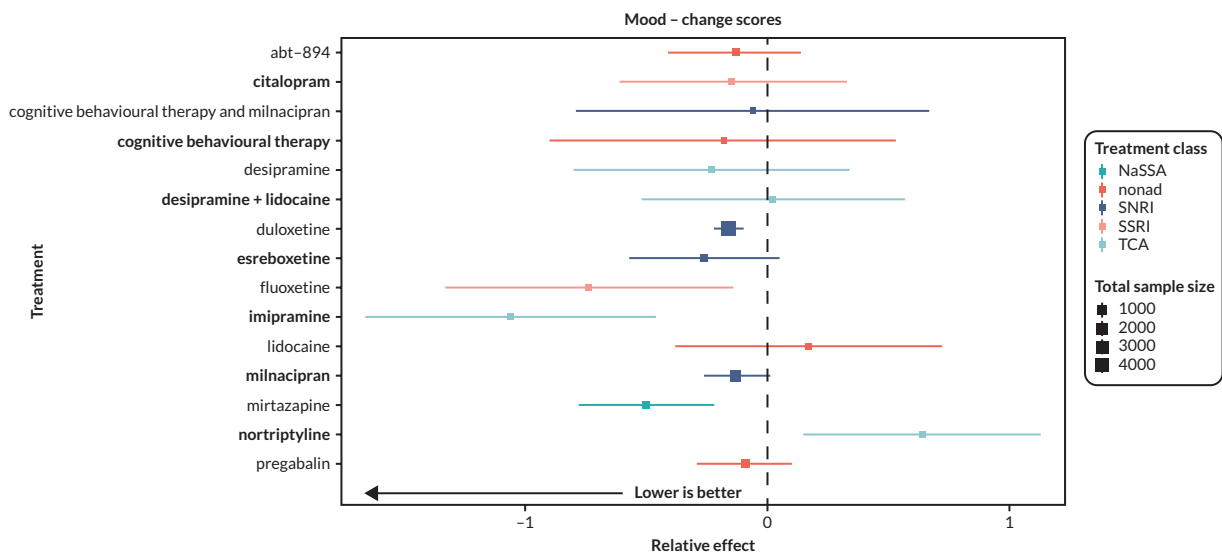


FIGURE 8 Mood: summary forest plot (SMD with CrIs).

**Class, condition and risk of bias**

We did not undertake further analyses because of small sample sizes, network geometry and the risk of overfitting, but pairwise and NMA (excluding multidose trials) were performed in CINeMA to inform strength of evidence assessment.

**Confidence in network meta-analysis**

The design-by-treatment test showed no evidence of inconsistency ( $\chi^2 = 1.83$ ,  $p = 0.4$ ), and all  $I^2$  values were below 40%, despite the analysis being unable to run node-splitting. Both duloxetine and milnacipran were rated as having moderate-certainty evidence; there were no domains indicating major concern. We rated mirtazapine as having low-certainty evidence as the estimates were formed from only one trial.

**Adverse events****TABLE 5** Adverse events: summary of findings

Estimates of effects, CrIs and certainty of the evidence for adverse events in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: amitriptyline standard dose (25–75 mg); desvenlafaxine high dose (> 50 mg); duloxetine low dose (< 60 mg), standard dose (60 mg), and high dose (> 60 mg); milnacipran standard dose (100 mg) and high dose (> 100 mg); mirtazapine standard dose (30 mg).							
Comparator (reference): placebo							
Outcome: adverse events (as reported per study)							
Direction: lower is better (i.e. fewer people reporting adverse events)							
Total RCTs: 93 Total participants: 22,558	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (GRADE)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Desvenlafaxine high dose RCTs: 2 Participants: 905	1.67 (0.92 to 2.41)	174/220 791 per 1000	590/685 863 per 1000	72 more per 1000	Very low <sup>c,d,e</sup>	30 (16 to 48)	Not significantly different from placebo
Mirtazapine standard dose RCTs: 2 Participants: 457	1.70 (0.48 to 2.91)	135/228 592 per 1000	162/229 712 per 1000	120 more per 1000	Very low <sup>d,e</sup>	31 (11 to 52)	Not significantly different from placebo
Duloxetine standard dose RCTs: 20 Participants: 4998	1.88 (1.58 to 2.17)	1259/2164 582 per 1000	1883/2834 723 per 1000	142 more per 1000	Very low <sup>c,d</sup>	33 (24 to 42)	Equivalent NNT <sub>H</sub> is 7.0
Milnacipran standard dose RCTs: 8 Participants: 2491	1.92 (1.37 to 2.46)	930/1235 753 per 1000	1039/1256 854 per 1000	101 more per 1000	Very low <sup>c,d,e</sup>	33 (20 to 45)	Equivalent NNT <sub>H</sub> is 10
Duloxetine high dose RCTs: 10 Participants: 4000	1.93 (1.64 to 2.23)	1199/1912 627 per 1000	1587/2088 764 per 1000	137 more per 1000	Very low <sup>c,d</sup>	34 (24 to 43)	Equivalent NNT <sub>H</sub> is 7.03
Duloxetine low dose RCTs: 6 Participants: 1031	2.03 (1.45 to 2.62)	271/437 620 per 1000	325/594 768 per 1000	148 more per 1000	Very low <sup>c,d</sup>	35 (21 to 47)	Equivalent NNT <sub>H</sub> is 7.0
Milnacipran high dose RCTs: 7 Participants: 2837	2.44 (1.89 to 2.98)	930/1264 736 per 1000	1294/1573 872 per 1000	136 more per 1000	Very low <sup>c,d</sup>	39 (25 to 50)	Equivalent NNT <sub>H</sub> is 6.8

TABLE 5 Adverse events: summary of findings (continued)

Total RCTs: 93 Total participants: 22,558	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (GRADE)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Amitriptyline standard dose RCTs: 10 Participants: 997	2.66 (2.14 to 3.19)	250/479 522 per 1000	351/518 744 per 1000	222 more per 1000	Very low <sup>c,d,f</sup>	41 (28 to 51)	Equivalent NNT <sub>H</sub> is 4.5
Esreboxetine standard dose RCTs: 1 Participants: 783	2.92 (1.90 to 3.93)	85/227 374 per 1000	315/556 636 per 1000	262 more per 1000	Very low <sup>c,d,e,f</sup>	42 (21 to 56)	Equivalent NNT <sub>H</sub> is 3.8

NNT<sub>H</sub>, number needed to treat for an additional harmful outcome.

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.

b Mean rank and CrIs are presented.

c Downgraded due to within-study bias.

d Downgraded due to imprecision in the estimate.

e Downgraded due to heterogeneity in the estimate.

f Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

g Downgraded due to incoherence in the network.

#### Notes:

GRADE Working Group grades of evidence:

High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

## Results

For adverse events, we report the treatment–dose network. There were similar levels of heterogeneity and inconsistency across networks but we were able to run node-splitting models for treatment–dose.

We included 93 RCTs with a total of 22,558 participants. Of all the studies in the network, 47 trials compared antidepressants only against placebo, 27 were multiarm trials with another active comparator, 15 were dose-comparison trials, and 4 compared two antidepressants to each other. We rated 62 trials as having a high risk of bias. There were 60 different interventions, and some comparisons were informed only by direct evidence from one trial. We could not include data from one trial due to disconnected networks.

The antidepressants with ≥ 200 participants and therefore included in the summary are as follows:

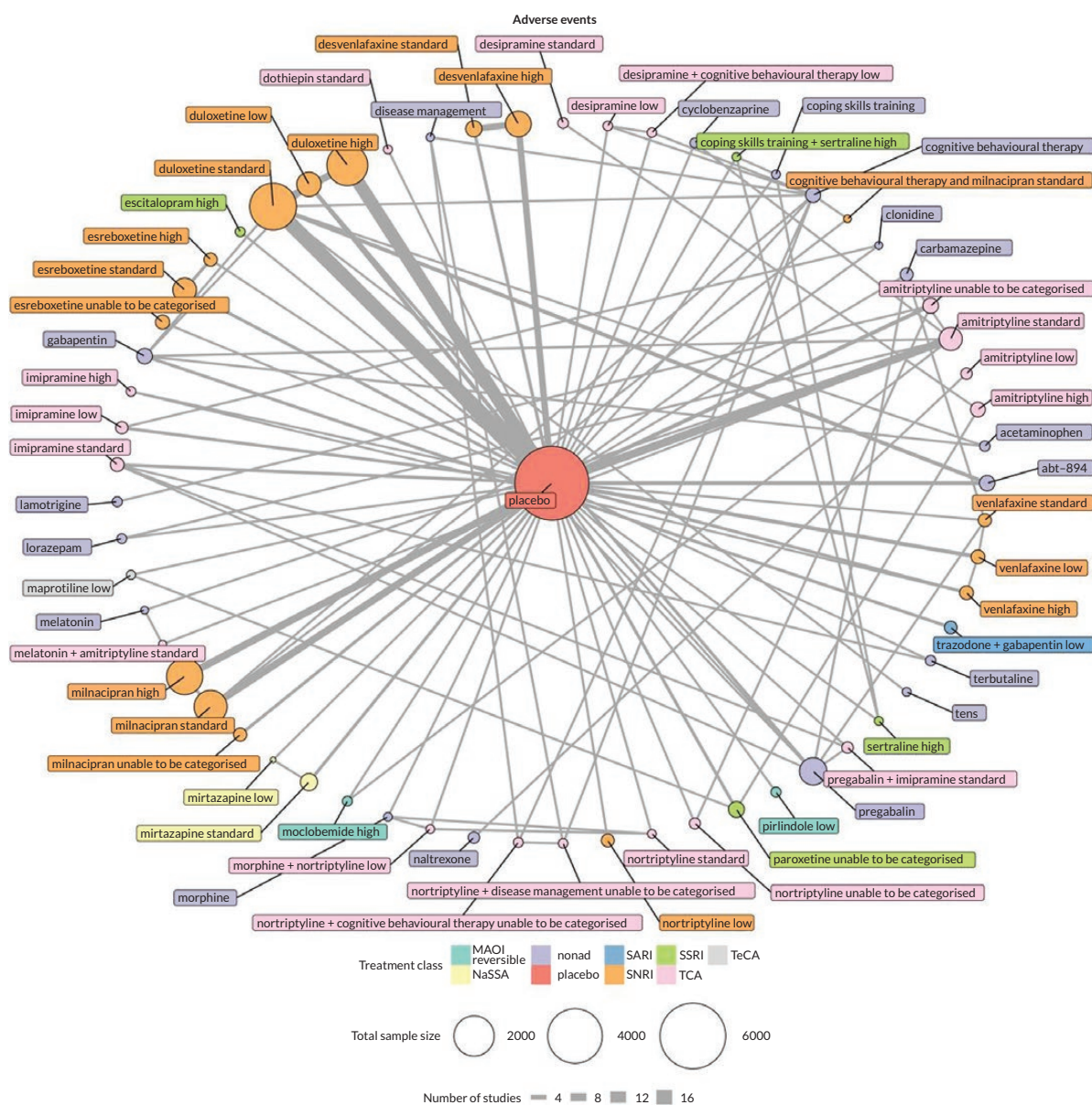
- amitriptyline standard dose
- desvenlafaxine high dose
- duloxetine high dose
- duloxetine low dose
- duloxetine standard dose
- esreboxetine standard dose
- milnacipran high dose

## RESULTS

- milnacipran standard dose
- mirtazapine standard dose.

An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 18](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 9](#), the forest plot of all interventions is presented in [Figure 10](#) and the forest plot of antidepressants only is presented in [Figure 11](#).



**FIGURE 9** Adverse events: network plot.

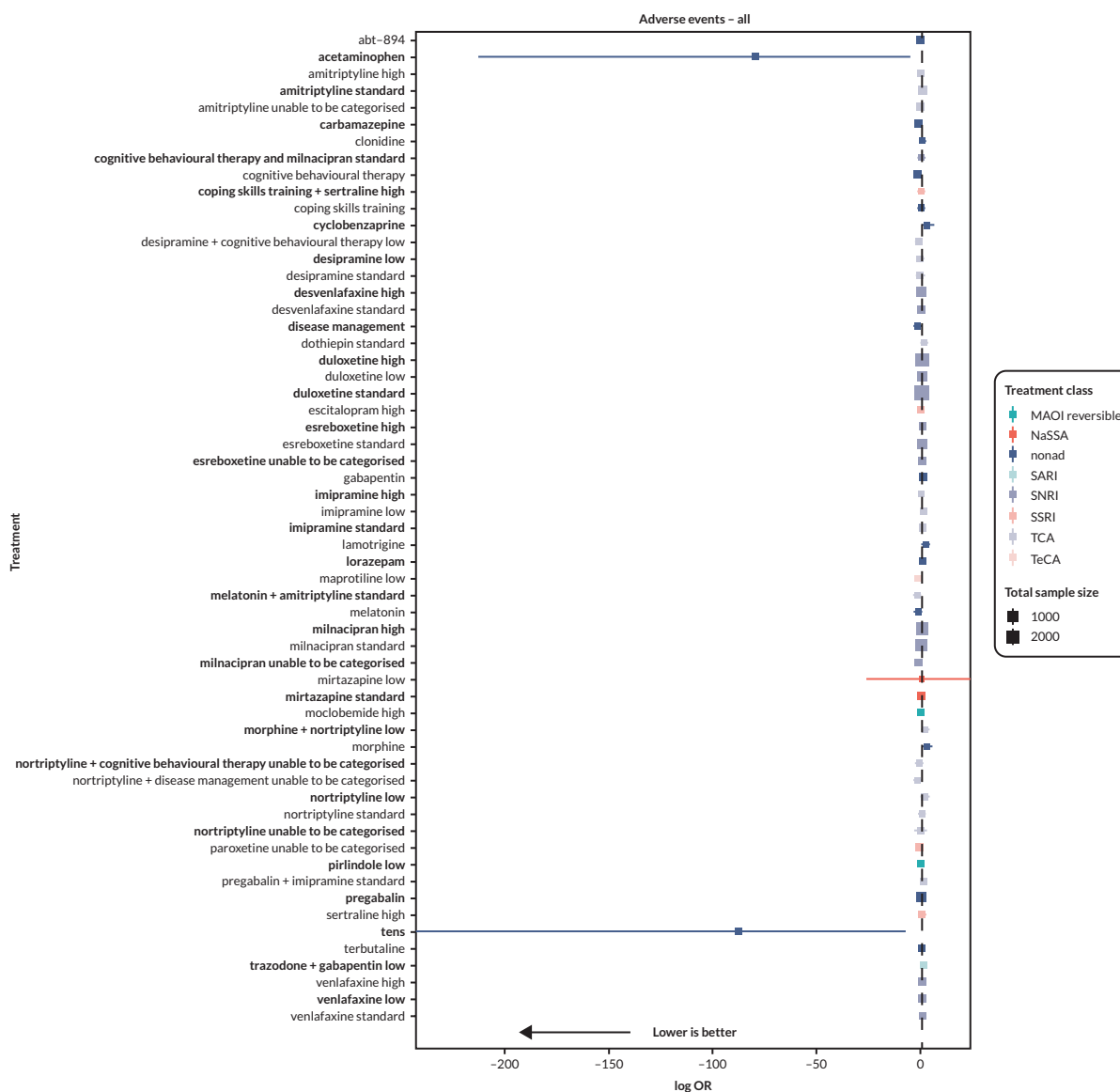


FIGURE 10 Adverse events: summary forest plot (log OR with Crls).

### Ranking of antidepressants

Data for adverse events were sparse, and studies were underpowered. All antidepressants with over 200 participants in the antidepressant arm were closely ranked. Desvenlafaxine and mirtazapine were the highest-ranked antidepressants, with no significant difference compared to placebo (OR 1.67, 95% CI 0.92 to 2.41 and OR 1.70, 95% CI 0.48 to 2.91, respectively). The evidence for both of these antidepressant doses was based on only two studies each. Duloxetine standard dose, milnacipran standard dose and duloxetine high dose were equally ranked. Duloxetine low dose, milnacipran high dose, amitriptyline standard dose and esreboxetine standard dose were the lowest-ranked antidepressants, with all ORs > 2.

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretation. We further investigated inconsistency through UME models and node-splitting

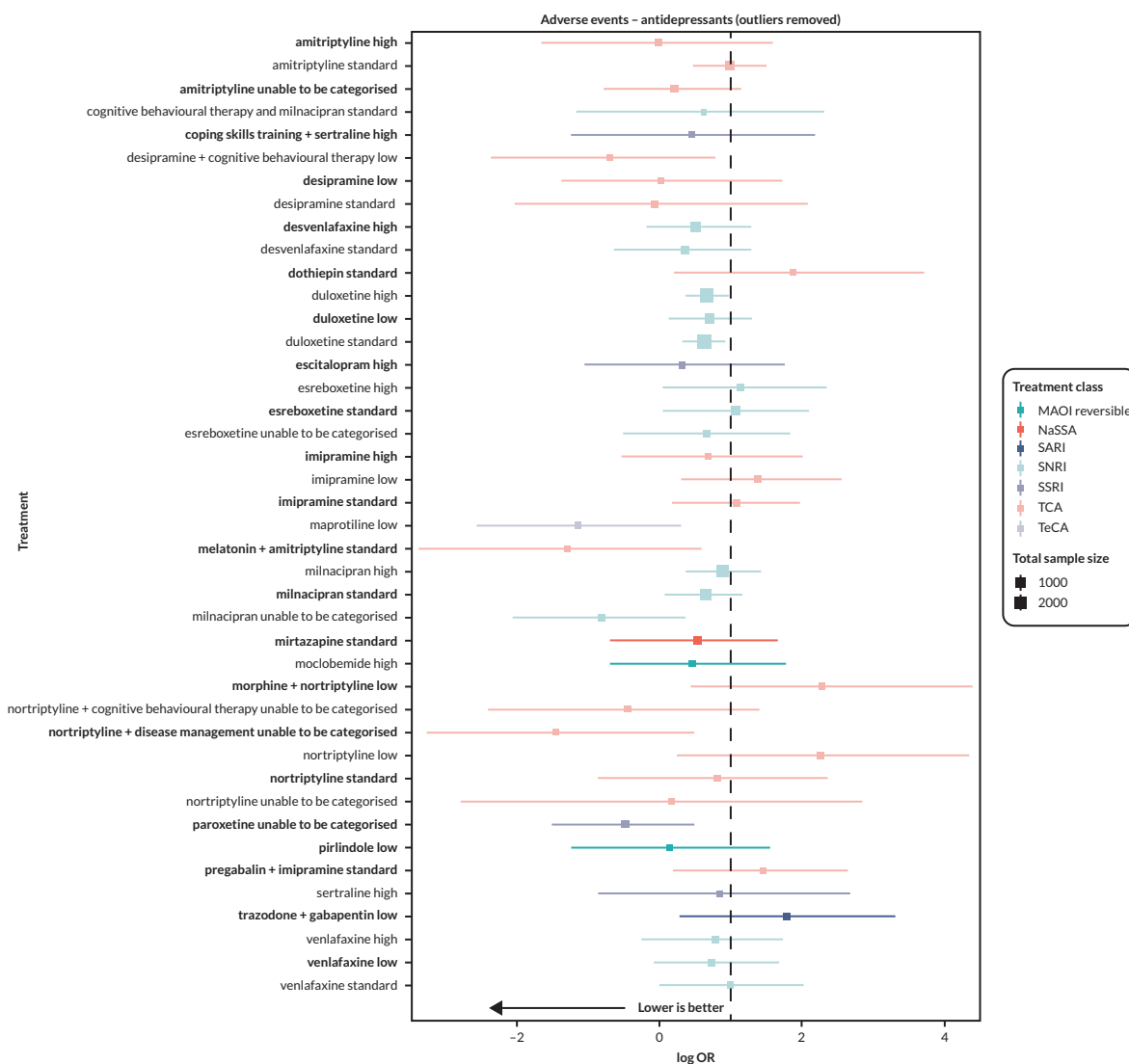


FIGURE 11 Adverse events (antidepressants only): summary forest plot.

models for all 30 comparisons where it was possible to compare direct and indirect evidence. There was evidence of inconsistency in UME but not node-splitting models. These figures are available in [Report Supplementary Material 1](#). However, multiple divergent transition warnings indicate the potential for inconsistency to be poorly estimated in the latter models.

**Class, condition and risk of bias**

Our overall model of adverse events is problematic due to divergent transitions, low effective sample sizes and inconsistency in the UME model. We were unable to undertake further exploration of class, condition and risk of bias given the high uncertainty in overall effects.

**Confidence in network meta-analysis**

We were unable to use CINeMA for this outcome due to the complexity of the network. Therefore, two authors (HB and GS) made the judgements based on GRADE and CINeMA domains and the available results. We judged all antidepressants and doses as very low confidence, primarily due to concerns with within-study bias, and imprecision in the network.

## Secondary outcomes

### Moderate pain relief

TABLE 6 Moderate pain relief: summary of findings

Estimates of effects, Crls and certainty of the evidence for moderate pain relief in people with chronic pain							
Bayesian NMA summary of findings table							
Total RCTs: 40 Total participants: 14,208	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (CINeMA)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Mirtazapine RCTs: 2 Participants: 462	1.92 (1.45 to 2.39)	70/224 313 per 1000	112/238 466 per 1000	154 more per 1000	Low <sup>c</sup>	7 (3 to 13)	Equivalent NNTB is 6.5
Duloxetine RCTs: 24 Participants: 7833	1.79 (1.67 to 1.91)	1324/3271 405 per 1000	2469/4562 549 per 1000	144 more per 1000	Moderate <sup>d</sup>	7 (4 to 11)	Equivalent NNTH is 6.9
Milnacipran RCTs: 7 Participants: 3056	1.7 (1.48 to 1.92)	347/1128 308 per 1000	825/1928 430 per 1000	123 more per 1000	Moderate <sup>d</sup>	8 (4 to 12)	Equivalent NNTH is 8.1
Esreboxetine RCTs: 2 Participants: 1374	1.65 (1.32 to 1.98)	107/409 262 per 1000	356/965 369 per 1000	107 more per 1000	Low <sup>c,d</sup>	9 (4 to 13)	Equivalent NNTH is 9.3

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.

b Mean rank and Crls are presented.

c Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

d Downgraded due to within-study bias.

e Downgraded due to imprecision in the estimate.

f Downgraded due to heterogeneity in the estimate.

g Downgraded due to incoherence in the network.

**Notes:**  
CINeMA grades of confidence in the evidence:  
High: further research is unlikely to change our confidence in the estimate of effect.  
Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
Very low: we are very uncertain about the estimate.

## Results

For moderate pain relief, we report the treatment network as this model had low heterogeneity and no evidence of inconsistency.

We included 40 RCTs with a total of 14,208 participants (range 37–1025). Of these, 20 trials compared against placebo, 8 were multiarm trials with another active comparator, 11 were dose-comparison trials,

## RESULTS

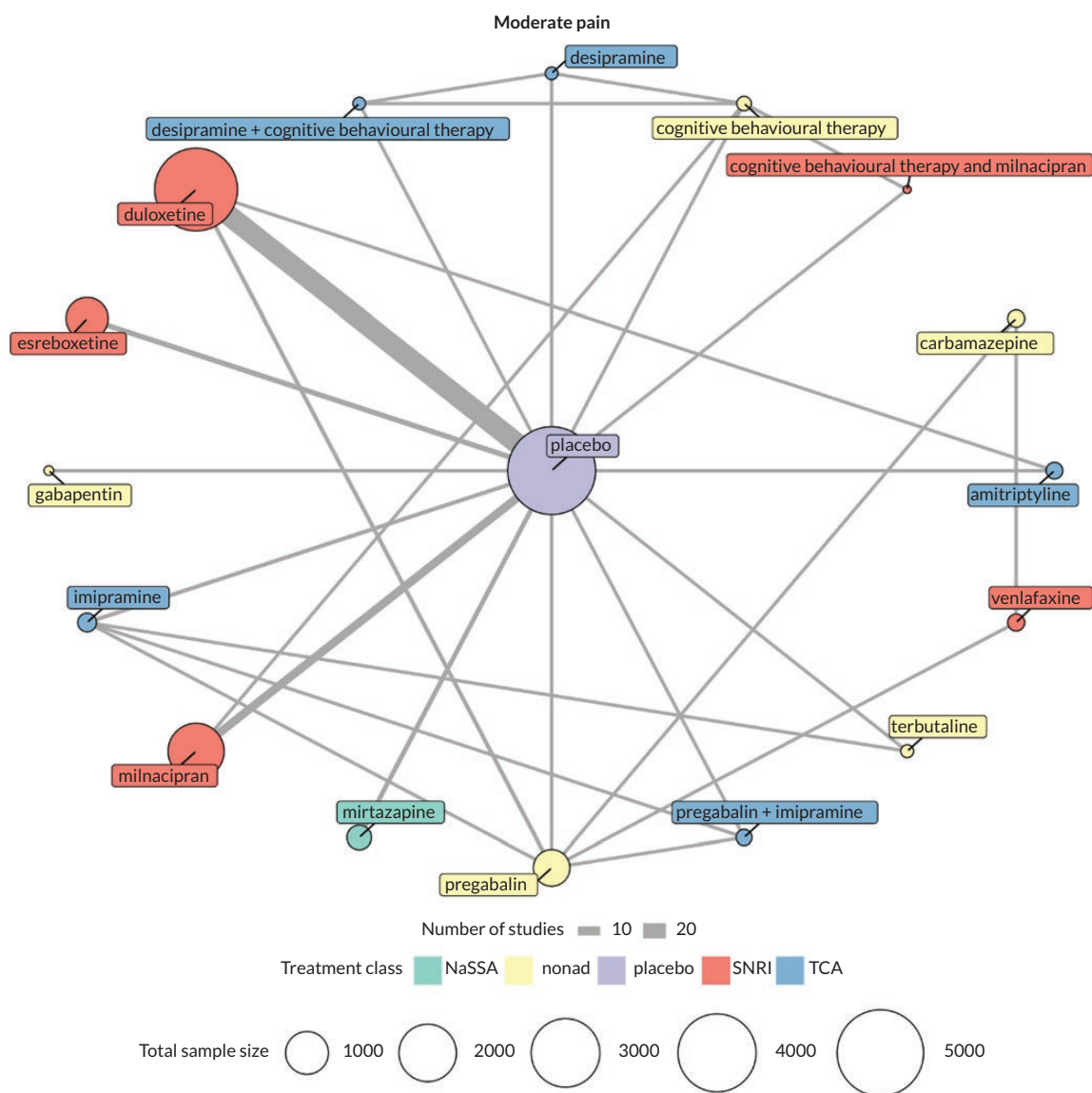
and 1 study compared two antidepressants head to head. There were 17 different interventions, and some comparisons were informed only by direct evidence from one trial. We rated 25 studies as at high risk of bias.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

- duloxetine
- esreboxetine
- milnacipran
- mirtazapine.

An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 19](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 12](#) and the forest plot is presented in [Figure 13](#).



**FIGURE 12** Moderate pain relief: network plot.

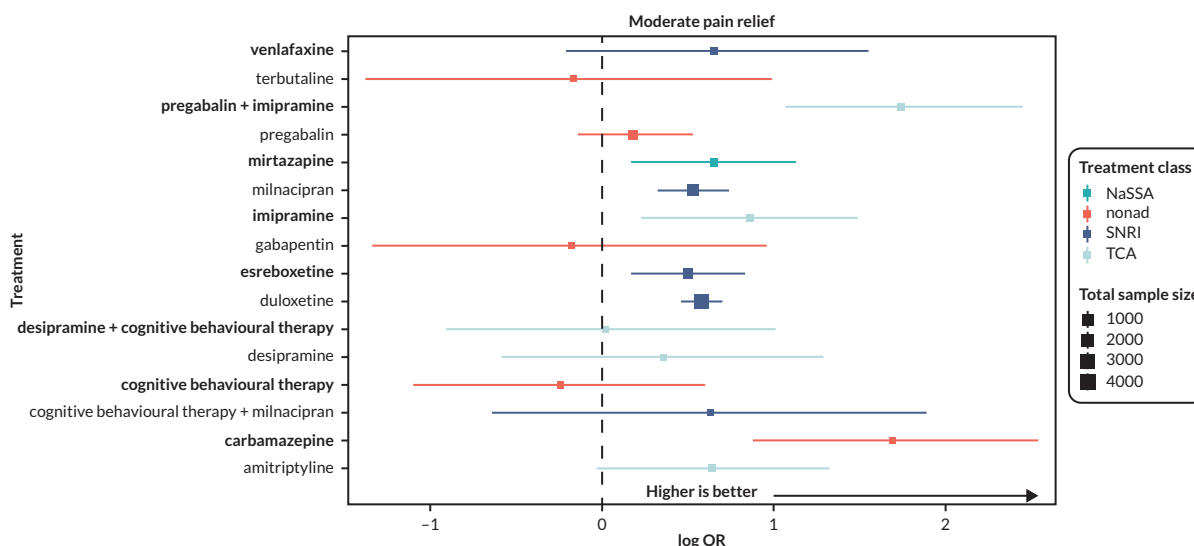


FIGURE 13 Moderate pain relief: summary forest plot (log OR with CrIs).

### Ranking of antidepressants

All antidepressants with over 200 participants in the antidepressant arm showed an effect for moderate pain relief and were very closely ranked. Mirtazapine was the highest-ranked antidepressant (OR 1.92, 95% CI 1.45 to 2.39), followed by duloxetine (OR 1.79, 95% CI 1.67 to 1.91), milnacipran (OR 1.7, 95% CI 1.48 to 1.92) and esreboxetine (OR 1.65, 95% CI 1.32 to 1.98).

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretation. The UME model showed no evidence of inconsistency. We confirmed this with node-splitting models for all nine comparisons where it was possible to compare direct and indirect evidence. The comparison of duloxetine and placebo had the lowest Bayesian  $p$ -value ( $p = 0.18$ ) with indirect evidence indicative of a larger effect than direct evidence. These figures are available in [Report Supplementary Material 1](#).

### Exploration of heterogeneity

We also explored the impact of including dose in the model. There was low heterogeneity (Tau = 0.11), and while there was no evidence of inconsistency in UME and node-splitting models, there were several divergent transitions. The analysis showed similar rankings of antidepressants to the treatment-only model, with mirtazapine, duloxetine and milnacipran remaining the highest-ranked drugs across doses. The full results of all the analyses are reported in [Report Supplementary Material 1](#).

### Class

There were three classes included in the treatment-only analysis: NaSSA, SNRI and TCA. Only the NaSSA and SNRI classes had over 200 participants in the analyses. SNRI was the highest-ranked class (logOR 0.56, CrI 0.45 to 0.60), followed by NaSSA (logOR 0.67, CrI 0.11 to 1.23).

### Condition and risk of bias

We were unable to undertake further NMAs due to small sample size, network geometry and risk of overfitting, but pairwise and NMA excluding multidose trials were undertaken to inform strength of evidence assessment using CINeMA.

### Confidence in network meta-analysis

The design-by-treatment test showed no evidence of inconsistency between the direct and indirect evidence in the network ( $\chi^2 = 2.65$ ,  $p = 0.62$ ), and only esreboxetine had an  $I^2$  value of above 40%

(44.6%). We rated duloxetine and milnacipran as having moderate certainty in the results, while we downgraded mirtazapine and esreboxetine due to low numbers of studies and participants.

## Physical function

**TABLE 7** Physical function: summary of findings

Estimates of effects, CrIs and certainty of the evidence for physical function in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine standard dose (60 mg) and high dose (> 60 mg); milnacipran standard dose (100 mg) and high dose (> 100 mg); mirtazapine standard dose (30 mg)							
Comparator (reference): placebo							
Outcome: change in physical function (lower scores are better) from a range of measures, including Fibromyalgia Impact Questionnaire and the SF-36							
Direction: lower is better (i.e. a greater improvement in physical function and disability)							
Total studies: 32 Total participants: 11,760	Relative effect	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Duloxetine standard dose RCTs: 15 Participants: 3887	-	-	-	SMD -0.24 (-0.32 to -0.18)	High	6 (3 to 8)	Small effect
Duloxetine high dose RCTs: 13 Participants: 3503	-	-	-	SMD -0.23 (-0.30 to -0.16)	Moderate <sup>c</sup>	6 (2 to 9)	Small effect
Milnacipran standard dose RCTs: 3 Participants: 1840	-	-	-	SMD -0.18 (-0.30 to -0.07)	Moderate <sup>c</sup>	7 (4 to 11)	Small effect
Milnacipran high dose RCTs: 2 Participants: 1670	-	-	-	SMD -0.1 (-0.22 to 0.07)	Very low <sup>c,d</sup>	9 (6 to 13)	Not significantly different from placebo
Mirtazapine standard dose RCTs: 1 Participants: 204	-	-	-	SMD 0.62 (0.11 to 0.69)	Very low <sup>e</sup>	16 (15 to 16)	Moderate to large effect

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Mean rank and CrIs are presented.

b SMD interpretation based on clinical judgement and in line with Cohen 1988<sup>82</sup> and the *Cochrane Handbook for Systematic Reviews of Interventions* as small (0.2), moderate (0.5) and large (0.8).

c Downgraded due to within-study bias.

d Downgraded due to heterogeneity in the estimate.

e Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

f Downgraded due to imprecision in the estimate.

g Downgraded due to incoherence in the network.

**Notes:**

CINeMA grades of confidence in the evidence:

High: further research is unlikely to change our confidence in the estimate of effect.

Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low: we are very uncertain about the estimate.

## Results

For physical function, we report the change score treatment–dose network as it had lower heterogeneity than other models and no inconsistency.

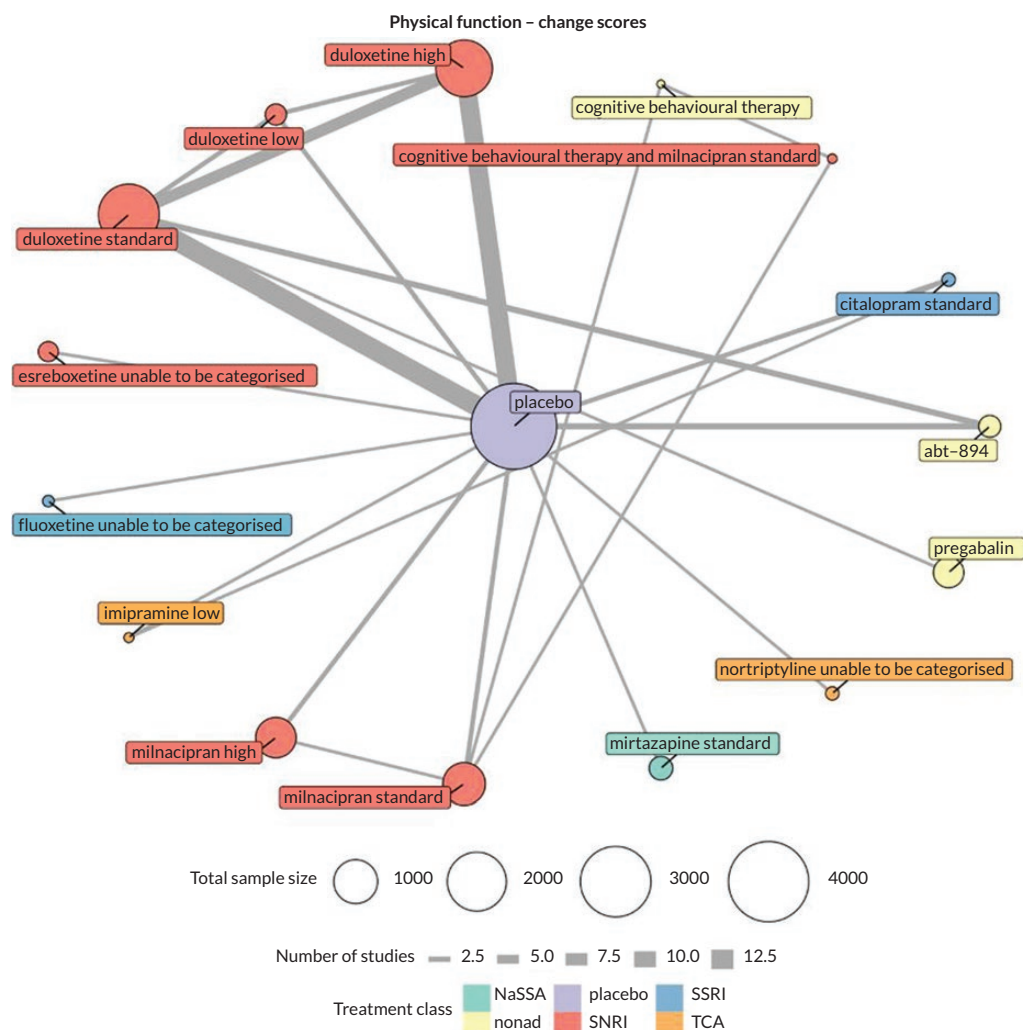
We included 32 RCTs with a total of 11,760 participants (range 42–1025). Of these, 20 trials compared against placebo, 4 were head-to-head trials with another active comparator, 7 were dose-comparison trials, and 1 was a direct head-to-head comparison between two different antidepressants. There were 18 different interventions, and some comparisons were informed only by direct evidence from one trial. We rated 21 studies as at high risk of bias.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

- duloxetine high dose
- duloxetine standard dose
- milnacipran high dose
- milnacipran standard dose
- mirtazapine standard dose.

An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 20](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 14](#) and the forest plot of placebo comparisons in [Figure 15](#).



**FIGURE 14** Physical function: network plot.

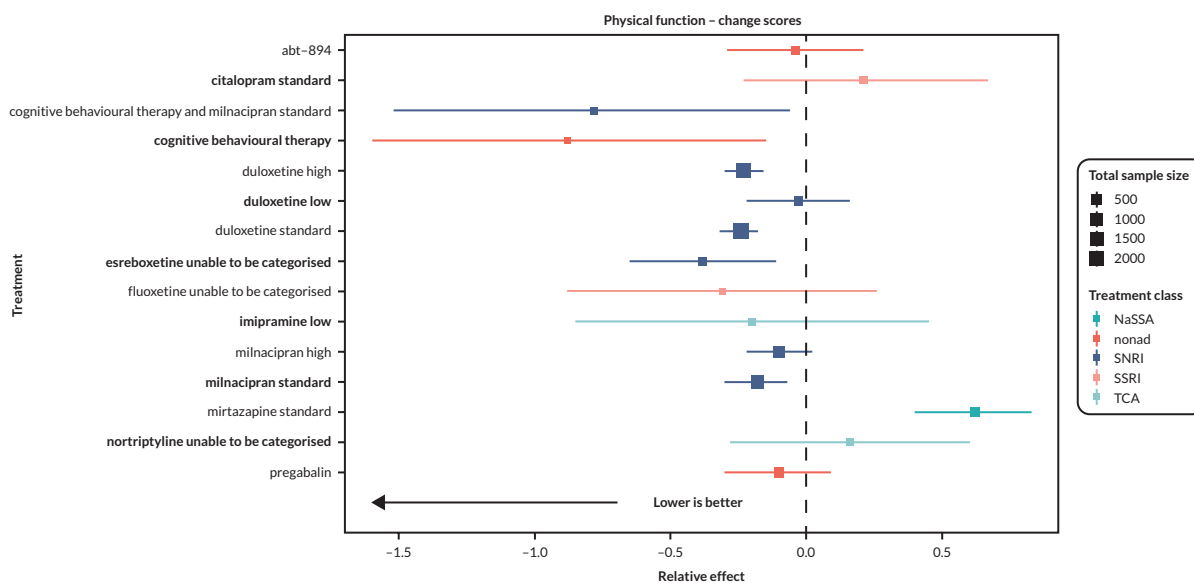


FIGURE 15 Physical function: summary forest plot (SMD with CrIs).

### Ranking of antidepressants

Duloxetine standard dose (SMD  $-0.24$ , 95% CI  $-0.32$  to  $-0.18$ ), duloxetine high dose (SMD  $-0.23$ , 95% CI  $0.30$  to  $0.16$ ) and milnacipran standard dose (SMD  $-0.18$ , 95% CI  $-0.30$  to  $-0.07$ ) were the highest-ranked antidepressants, with small effects. Duloxetine standard dose and duloxetine high doses were equally effective. Milnacipran high dose showed no significant difference compared to placebo (SMD  $-0.10$ , 95% CI  $-0.22$  to  $0.07$ ). Mirtazapine standard dose was the lowest-ranked antidepressant (SMD  $0.62$ , 95% CI  $0.11$  to  $0.69$ ).

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretation. We performed node-splitting models for all four comparisons where it was possible to compare direct and indirect evidence (see [Figure 26](#)). The lowest Bayesian  $p$ -value was for the comparison of duloxetine high dose versus placebo, where direct evidence showed a larger effect than indirect evidence ( $p = 0.07$ ). These figures are available in [Report Supplementary Material 1](#).

### Class

Four classes of antidepressants were included in the analysis (SNRI, SSRI, TCA and NaSSA); however, due to interventions including combinations of drugs, models including class could not be analysed.

### Condition and risk of bias

We were unable to undertake further NMAs due to small sample sizes, network geometry and the risk of overfitting.

### Confidence in network meta-analysis

The design-by-treatment test showed no evidence of inconsistency between the direct and indirect evidence ( $\chi^2 = 6.45$ ,  $p = 0.69$ ), and no antidepressant had an  $I^2$  value of over 40%, although values could not be generated for mirtazapine. We rated duloxetine and milnacipran as moderate certainty, downgraded only due to some concerns with within-study bias. We downgraded esreboxetine and mirtazapine further to low certainty due to the small number of studies and participants included in the analyses.

## Sleep

TABLE 8 Sleep: summary of findings

Estimates of effects, CrIs and certainty of the evidence for sleep in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine standard dose (60 mg) and high dose (> 60 mg); milnacipran standard dose (100 mg) and high dose (> 100 mg)							
Comparator (reference): placebo							
Outcome: change in sleep as measured on various scales, primarily Brief Pain Inventory Sleep Item							
Direction: lower is better (i.e. greater improvement in sleep compared to baseline)							
Total RCTs: 18 Total participants: 6301	Relative effect	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Duloxetine standard RCTs: 11 Participants: 2615	-	-	-	SMD -0.21 (-0.30 to -0.12)	Moderate <sup>c,d</sup>	3 (1 to 6)	Small effect
Duloxetine high RCTs: 6 Participants: 1494	-	-	-	SMD -0.14 (-0.27 to -0.01)	Very low <sup>c,d,e</sup>	4 (2 to 7)	Small effect
Milnacipran standard RCTs: 1 Participants: 799	-	-	-	SMD -0.06 (-0.30 to 0.17)	Very low <sup>c,d,e,f</sup>	6 (2 to 9)	Not significantly different from placebo
Milnacipran high RCTs: 1 Participants: 797	-	-	-	SMD -0.03 (-0.29 to 0.20)	Very low <sup>c,d,e,f</sup>	7 (2 to 9)	Not significantly different from placebo
Network meta-analysis summary of findings table definitions (applied across <a href="#">Tables 2-13</a> ):							
a Mean rank and CrIs are presented.							
b SMD interpretation based on clinical judgement and in line with Cohen 1988 <sup>82</sup> and the <i>Cochrane Handbook for Systematic Reviews of Interventions</i> as small (0.2), moderate (0.5) and large (0.8).							
c Downgraded due to within-study bias.							
d Downgraded due to incoherence in the network.							
e Downgraded due to heterogeneity in the estimate.							
f Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.							
g Downgraded due to imprecision in the estimate.							
<b>Notes:</b>							
CINeMA grades of confidence in the evidence:							
High: further research is unlikely to change our confidence in the estimate of effect.							
Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.							
Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.							
Very low: we are very uncertain about the estimate.							

**Results**

For sleep, we report the change score treatment–dose network as this was the most robust and reliable model.

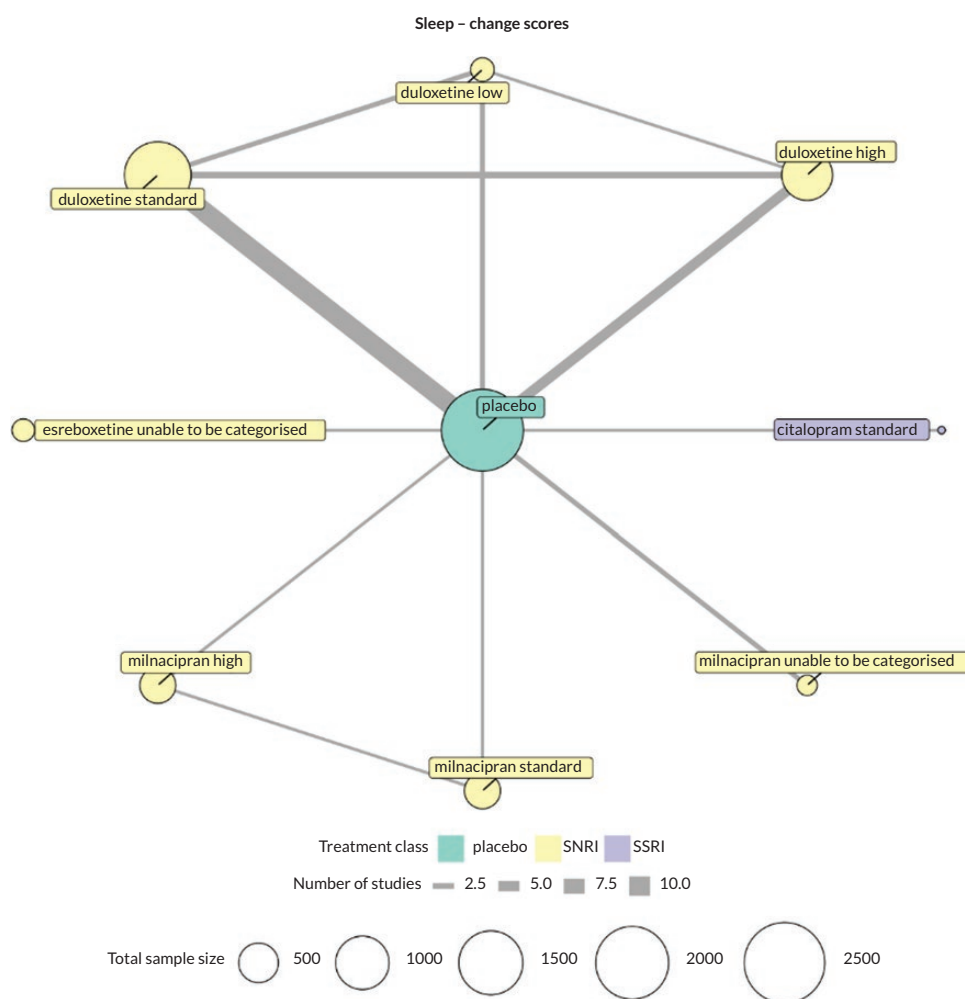
We included 18 RCTs with a total of 6301 participants (range 42–1195). Of these, 12 studies compared against placebo and 6 were dose-comparison studies. There were eight different interventions, and some comparisons were informed only by direct evidence from one trial. We rated nine trials as at high risk of bias overall.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are as follows:

- duloxetine standard dose
- duloxetine high dose
- milnacipran standard dose
- milnacipran high dose.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3, Table 21](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 16](#) and the forest plot for placebo comparison is presented in [Figure 17](#).



**FIGURE 16** Sleep: network plot.

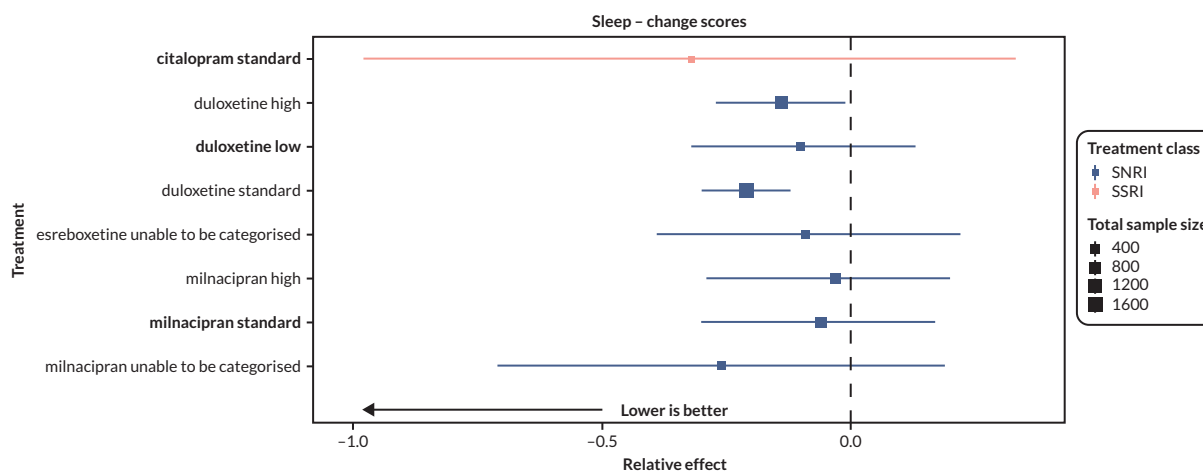


FIGURE 17 Sleep: summary forest plot (SMD with CrIs).

### Ranking of antidepressants

Duloxetine standard and high doses were the highest-ranked antidepressants and the only antidepressants to show a significant effect when compared to placebo, although the effects were small (standard dose: SMD  $-0.21$ , 95% CI  $-0.30$  to  $-0.12$ ; high dose: SMD  $-0.14$ , 95% CI  $-0.27$  to  $-0.01$ ). Milnacipran standard dose (SMD  $-0.06$ , 95% CI  $-0.30$  to  $0.17$ ) and high dose (SMD  $-0.03$ , 95% CI  $-0.29$  to  $0.20$ ) showed no significant difference in comparison to placebo.

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretations. Node-splitting models had divergent transitions and indicated inconsistency for the comparison of high- and standard-dose duloxetine ( $p = 0.02$ ). We therefore downgraded the strength of evidence for the high-dose duloxetine estimate. These figures are available in [Report Supplementary Material 1](#).

### Class, condition and risk of bias

Although there were two different classes in the network (SNRI and SSRI), SSRI was only represented by one study using citalopram with 21 participants; therefore, only SNRI crossed the threshold of 200 participants. We did not explore condition and risk of bias further using NMA because of concerns about sample size, network geometry and the risk of overfitting.

### Confidence in network meta-analysis

The design-by-treatment test showed no evidence of inconsistency between the direct and indirect evidence in the network ( $\chi^2 = 7.39$ ,  $p = 0.4$ ), despite the concerns identified in node-splitting models. No antidepressants had  $I^2$  values of above 40%, although values could not be calculated for milnacipran high or standard doses. We rated only duloxetine as moderate certainty, downgraded from high due to some concerns about within-study bias and inconsistency from the NMA. We rated duloxetine high dose, milnacipran high dose and milnacipran standard dose as having a very low certainty of evidence. We downgraded duloxetine high dose due to major concerns regarding within-study bias and incoherence. We downgraded milnacipran standard and high doses due to major concerns regarding within-study bias, and some concerns regarding imprecision, heterogeneity and inconsistency. Of note, both milnacipran dose analyses were informed by the same study.

## Quality of life

TABLE 9 Quality of life: summary of findings

Estimates of effects, CrIs and certainty of the evidence for quality of life in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine, esreboxetine. All doses were combined for each antidepressant.							
Comparator (reference): placebo							
Outcome: quality of life (post-intervention scores) as reported on various scales including the EQ-5D and the Fibromyalgia Impact Questionnaire							
Direction: higher is better (i.e. a greater improvement in quality of life compared to baseline)							
Total RCTs: 19 Total participants: 3103	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Esreboxetine RCTs: 1 Participants: 998	-	-	-	SMD -0.30 (-1.24 to 0.64)	Very low <sup>c</sup>	8 (1 to 21)	Not significantly different from placebo
Duloxetine RCTs: 6 Participants: 867	-	-	-	SMD 0.02 (-0.56 to 0.58)	Low <sup>c,d</sup>	12 (4 to 20)	Not significantly different from placebo

EQ-5D, EuroQol-5 Dimensions.  
Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):  
a Mean rank and CrIs are presented.  
b SMD interpretation based on clinical judgement and in line with Cohen 1988<sup>82</sup> and the *Cochrane Handbook for Systematic Reviews of Interventions* as small (0.2), moderate (0.5) and large (0.8).  
c Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.  
d Downgraded due to within-study bias.  
e Downgraded due to imprecision in the estimate.  
f Downgraded due to heterogeneity in the estimate.  
g Downgraded due to incoherence in the network.

**Notes:**  
CINeMA grades of confidence in the evidence:  
High: further research is unlikely to change our confidence in the estimate of effect.  
Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
Very low: we are very uncertain about the estimate.

## Results

For quality of life, we report the post-intervention treatment network, as this was the network with the lowest heterogeneity.

We included 19 RCTs with a total of 3103 participants (range 30–998). Of these, 5 studies compared against placebo, 11 were multiarm trials with another active comparator, 2 were direct head-to-head comparisons of different antidepressants and 1 was a dose-comparison trial. There were 23 different interventions, and some comparisons were informed only by direct evidence from one trial. Data from one study could not be included due to disconnected networks. We rated 13 studies as at high risk of bias overall.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are as follows:

- duloxetine
- esreboxetine.

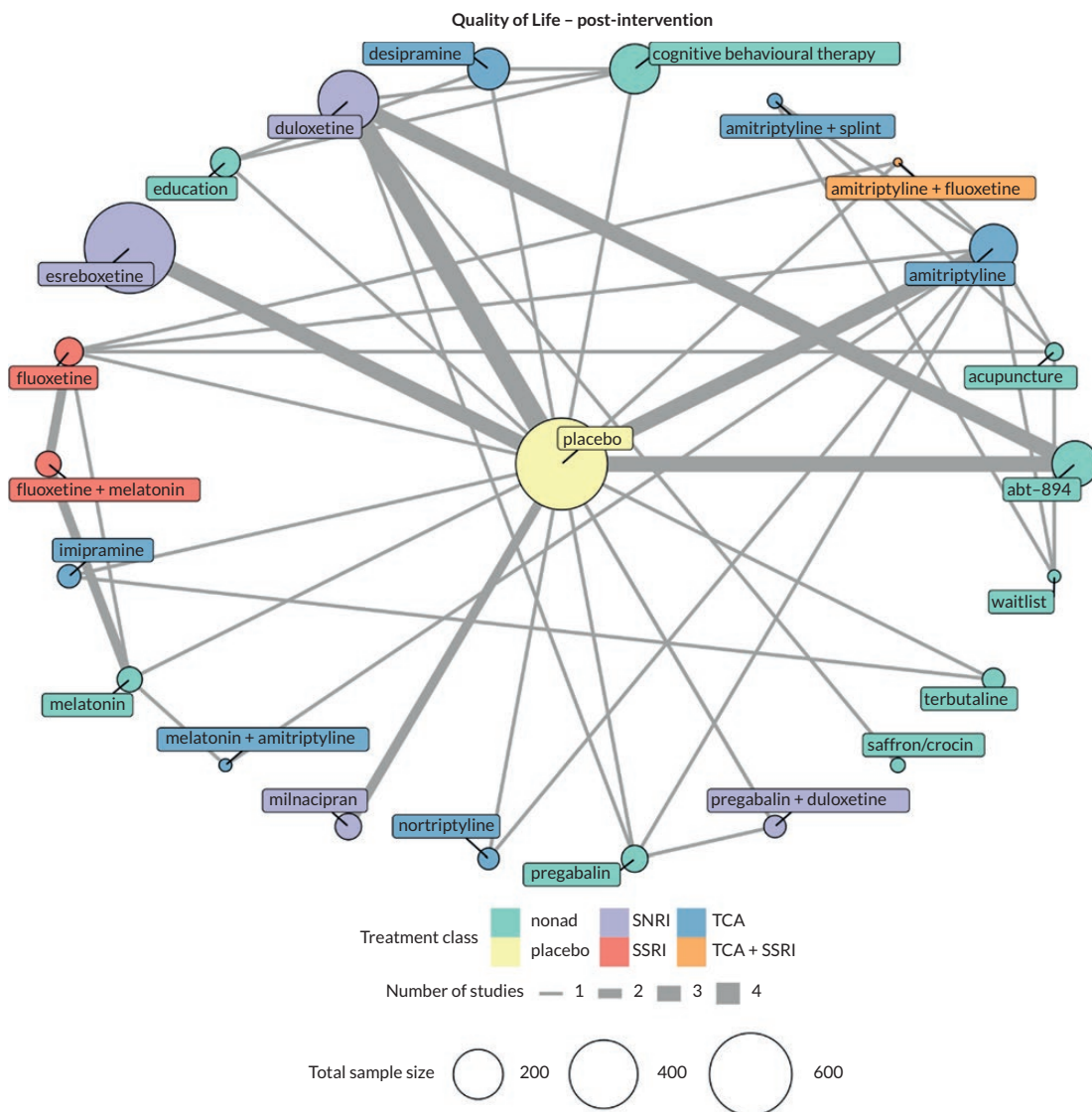
An overview of all interventions included in the analysis, the number of RCTs, and the number of participants is given in [Appendix 3, Table 22](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 18](#) and the forest plot is presented in [Figure 19](#).

### Ranking of antidepressants

Neither esreboxetine or duloxetine showed a significant difference compared to placebo for quality of life (SMD  $-0.30$ , 95% CI  $-1.24$  to  $0.64$  and SMD  $0.02$ , 95% CI  $-0.56$  to  $0.58$ , respectively).

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretations. Node-splitting models were undertaken for all 13 comparisons where



**FIGURE 18** Quality of life: network plot.

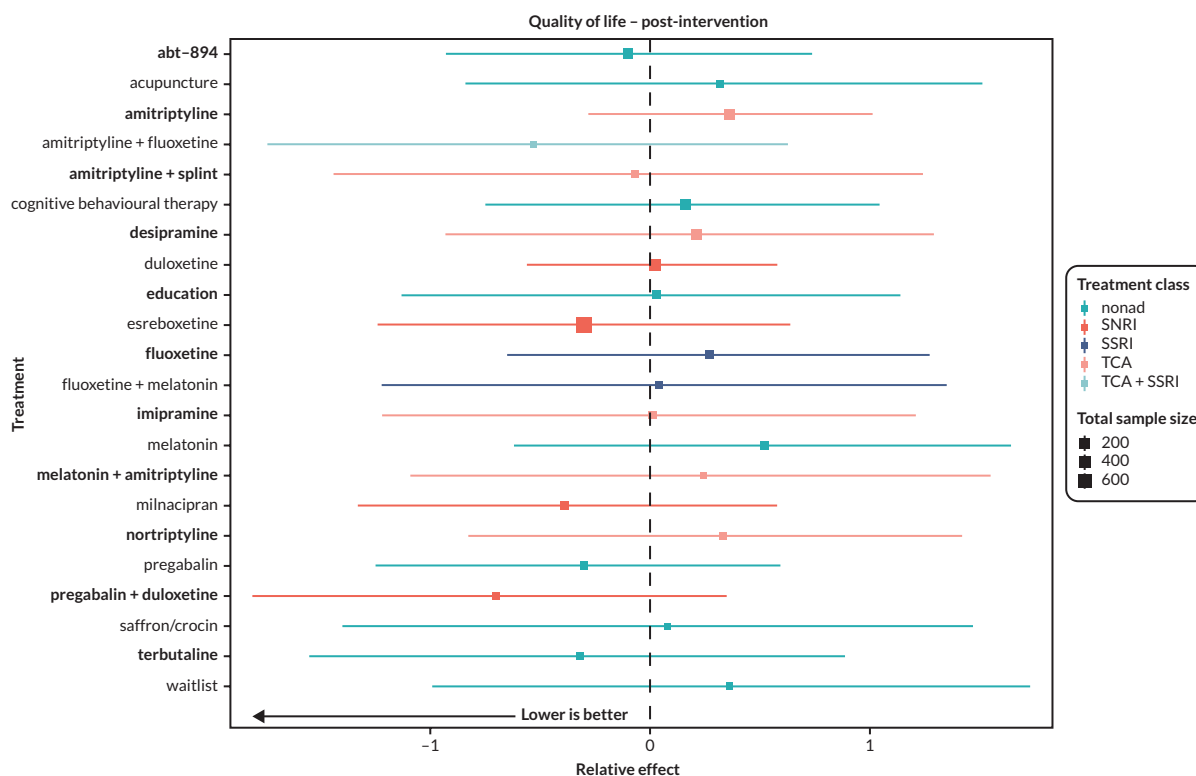


FIGURE 19 Quality of life: summary forest plot (SMD with CIs).

it was possible to compare direct and indirect evidence. The comparison with the lowest Bayesian  $p$ -value ( $p = 0.16$ ) was for fluoxetine compared to amitriptyline. These figures are available in [Report Supplementary Material 1](#). UME models also failed to identify inconsistency.

### Exploration of heterogeneity

We explored models including both treatment and dose; this model had higher heterogeneity (Tau = 0.67) and similar residual deviance to that of the treatment-only model.

### Class, condition and risk of bias

We were unable to generate meaningful networks including class, condition and risk of bias. Only one class had antidepressants with over 200 participants (SNRI). Small sample sizes, network geometry and the risk of overfitting precluded analyses of condition and risk and bias.

### Confidence in network meta-analysis

The design-by-treatment test showed evidence of significant inconsistency between the direct and indirect evidence in the network ( $\chi^2 = 80.27$ ,  $p = 0.00$ ) despite node-splitting and UME models indicating no concern. The  $I^2$  value for duloxetine showed evidence of heterogeneity ( $I^2 = 67.2\%$ ) and could not be calculated for esreboxetine. Therefore, we rated duloxetine as having low certainty of evidence (downgraded due to within-study bias, heterogeneity and inconsistency) and esreboxetine as having very low certainty of evidence (downgraded due to within-study bias, inconsistency and low numbers of studies).

## Patient Global Impression of Change (responders)

TABLE 10 Patient Global Impression of Change (responders): summary of findings

Estimates of effects, CrIs and certainty of the evidence for PGIC in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: desvenlafaxine high dose (> 50 mg); duloxetine standard dose (60 mg) and high dose (> 60 mg); esreboxetine standard dose (4–8 mg) and high dose (> 8 mg); milnacipran standard dose (100 mg) and high dose (> 100 mg)							
Comparator (reference): placebo							
Outcome: PGIC – people reporting much or very much improved (i.e. 1 or 2 on the 7-point PGIC scale)							
Direction: higher is better (i.e. more people reporting much or very much improved from baseline)							
Total RCTs: 12 Total participants: 6995	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (CINeMA)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Duloxetine standard dose RCTs: 3 Participants: 974	2.29 (1.98 to 2.60)	215 per 1000 106/493	382 per 1000 184/481	170 more per 1000	Moderate	2 (1 to 6)	Equivalent to NNTB of 5.9
Duloxetine high dose RCTs: 2 Participants: 567	2.03 (1.62 to 2.44)	250 per 1000 70/280	404 per 1000 113/287	154 more per 1000	Very low <sup>c</sup>	4 (1 to 7)	Equivalent to NNTB of 6.5
Milnacipran high dose RCTs: 3 Participants: 2057	1.99 (1.77 to 2.21)	282 per 1000 280/992	439 per 1000 480/1065	157 more per 1000	Low <sup>c</sup>	4 (1 to 7)	Equivalent to NNTB of 6.4
Milnacipran standard dose RCTs: 3 Participants: 2098	1.95 (1.73 to 2.17)	303 per 1000 320/1055	459 per 1000 462/1043	156 more per 1000	Moderate <sup>c</sup>	4 (1 to 7)	Equivalent to NNTB of 6.4
Esreboxetine standard dose RCTs: 1 Participants: 811	1.79 (1.44 to 2.14)	291 per 1000 80/275	423 per 1000 226/536	133 more per 1000	Very low <sup>c,d</sup>	5 (1 to 7)	Equivalent to NNTB of 7.5
Esreboxetine high dose RCTs: 1 Participants: 550	1.63 (1.24 to 2.02)	291 per 1000 80/275	401 per 1000 110/275	110 more per 1000	Very low <sup>c,d</sup>	6 (2 to 8)	Equivalent to NNTB of 9.1
Desvenlafaxine high dose RCTs: 1 Participants: 528	1.01 (0.58 to 1.44)	429 per 1000 54/126	431 per 1000 173/402	2 more per 1000	Very low <sup>c,d,e</sup>	8 (6 to 9)	Not significantly different from placebo
Network meta-analysis summary of findings table definitions (applied across <a href="#">Tables 2–13</a> ):							
a Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.							
b Mean rank and CrIs are presented.							
c Downgraded due to within-study bias.							
d Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.							
e Downgraded due to imprecision in the estimate.							
f Downgraded due to heterogeneity in the estimate.							
g Downgraded due to incoherence in the network.							

**TABLE 10** Patient Global Impression of Change (responders): summary of findings (*continued*)**Notes:**

CINeMA grades of confidence in the evidence:

High: further research is unlikely to change our confidence in the estimate of effect.

Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low: we are very uncertain about the estimate.

**Results**

For PGIC much/very much improved, we report the treatment–dose network as this had low heterogeneity with no inconsistency.

We included 12 RCTs with a total of 6995 participants (range 43–1025). Of these, eight studies compared against placebo and four were dose-comparison trials. There were nine different interventions, and some comparisons were informed only by direct evidence from one trial. We judged seven studies to be at high risk of bias.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

- desvenlafaxine high dose
- duloxetine high dose
- duloxetine standard dose
- esreboxetine high dose
- esreboxetine standard dose
- milnacipran high dose
- milnacipran standard dose.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3, Table 24](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 20](#) and the forest plot is presented in [Figure 21](#).

**Ranking of antidepressants**

Duloxetine standard dose was the highest-ranked antidepressant for PGIC much and very much improved, with a large effect (OR 2.29, 95% CI 1.98 to 2.60). Duloxetine high dose (OR 2.03, 95% CI 1.62 to 2.44), milnacipran high dose (OR 1.99, 95% CI 1.77 to 2.21) and milnacipran standard dose (OR 1.95, 95% CI 1.73 to 2.17) were the next highest-ranked antidepressants. Both esreboxetine doses showed a smaller effect (standard: OR 1.79, 95% CI 1.44 to 2.14; high: OR 1.63, 95% CI 1.24 to 2.02) but were among the lowest-ranked antidepressants. Desvenlafaxine high dose showed no significant effects when compared to placebo (OR 1.01, 95% CI 0.58 to 1.44).

A visual representation of the SUCRA rankings for every intervention included in the analysis did not alter the interpretation. The UME model had no evidence of inconsistency. We were only able to compare direct and indirect evidence for milnacipran standard versus milnacipran high dose, with a Bayesian  $p$ -value of 0.66 indicative of no inconsistency. These figures are available in [Report Supplementary Material 1](#).

**Class, condition and risk of bias**

We were unable to include class, condition and risk of bias in the models. For class, all the antidepressants included in the model were SNRIs. For condition and risk of bias, the sparse network geometry created disconnected networks with small sample sizes and high risk of overfitting.

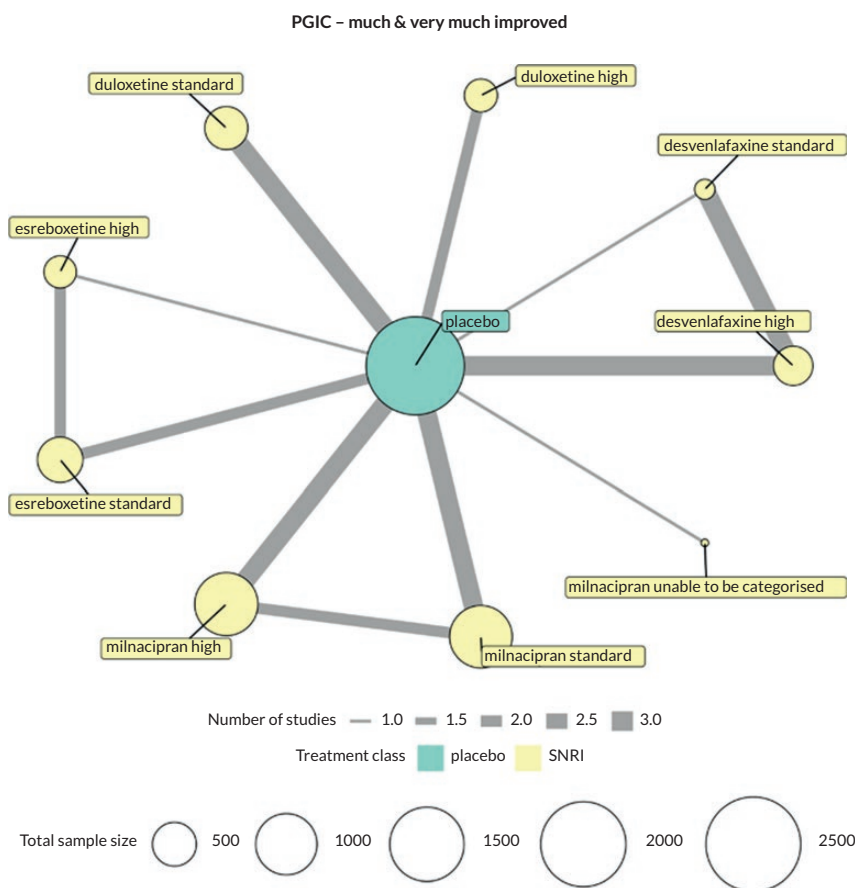


FIGURE 20 Patient Global Impression of Change much/very much improved: network plot.

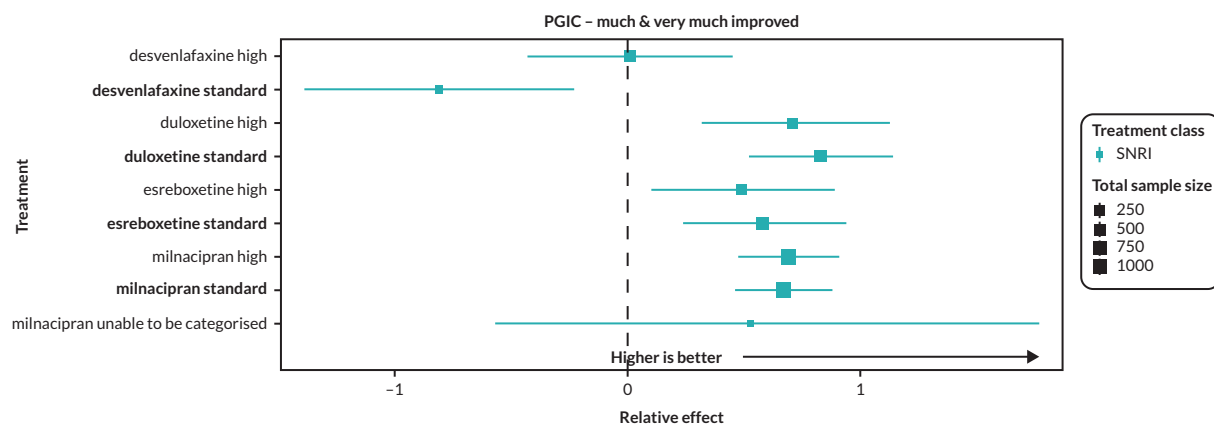


FIGURE 21 Patient Global Impression of Change much/very much improved: summary forest plot (log OR with CrIs).

**Confidence in network meta-analysis**

The design-by-treatment test showed no evidence of inconsistency ( $\chi^2 = 0.35, p = 0.84$ ), and no antidepressants had  $I^2$  values of over 40%. We rated the majority of the evidence to be very low certainty, due to within-study bias and low study and participant numbers. We rated milnacipran high dose as low certainty, downgraded due to major concerns of within-study bias. We rated milnacipran and duloxetine standard dose as moderate certainty, only downgraded due to concerns about within-study bias.

## Patient Global Impression of Change (continuous)

**TABLE 11** Patient Global Impression of Change (continuous): summary of findings

Estimates of effects, CrIs and certainty of the evidence for PGIC (continuous) in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: duloxetine low dose (< 60 mg), standard dose (60 mg) and high dose (> 60 mg)							
Comparator (reference): placebo							
Outcome: PGIC measured continuously on the PGIC 1–7 scale							
Direction: lower is better (1 on the scale represents 'very much improved', 7 represents 'very much worse')							
Total studies: 24 Total participants: 8415	Relative effect	Anticipated absolute effect (event rate)			Certainty of the evidence (CINeMA)	Ranking <sup>a</sup> (2.5% to 97.5% CrI)	Interpretation of findings <sup>b</sup>
		With placebo	With intervention	Difference (95% CI)			
Duloxetine standard dose RCTs: 14 Participants: 3847	-	-	-	SMD -0.36 (-0.44 to -0.29)	Moderate <sup>c</sup>	3 (1 to 4)	Small to moderate effect
Duloxetine high dose RCTs: 14 Participants: 3520	-	-	-	SMD -0.33 (-0.40 to -0.26)	Moderate <sup>c</sup>	3 (2 to 5)	Small to moderate effect
Duloxetine low dose RCTs: 5 Participants: 1097	-	-	-	SMD -0.23 (-0.35 to -0.11)	Moderate <sup>c,d</sup>	5 (3 to 6)	Small effect

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

a Mean rank and CrIs are presented.

b SMD interpretation based on clinical judgement and in line with Cohen 1988<sup>82</sup> and the *Cochrane Handbook for Systematic Reviews of Interventions* as small (0.2), moderate (0.5) and large (0.8).

c Downgraded due to incoherence in the network.

d Downgraded due to within-study bias.

e Downgraded due to imprecision in the estimate.

f Downgraded due to heterogeneity in the estimate.

g Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

**Notes:**  
CINeMA grades of confidence in the evidence:  
High: further research is unlikely to change our confidence in the estimate of effect.  
Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
Very low: we are very uncertain about the estimate.

### Results

For PGIC (continuous), we report the treatment–dose network as it had the lowest heterogeneity and the most clinical utility.

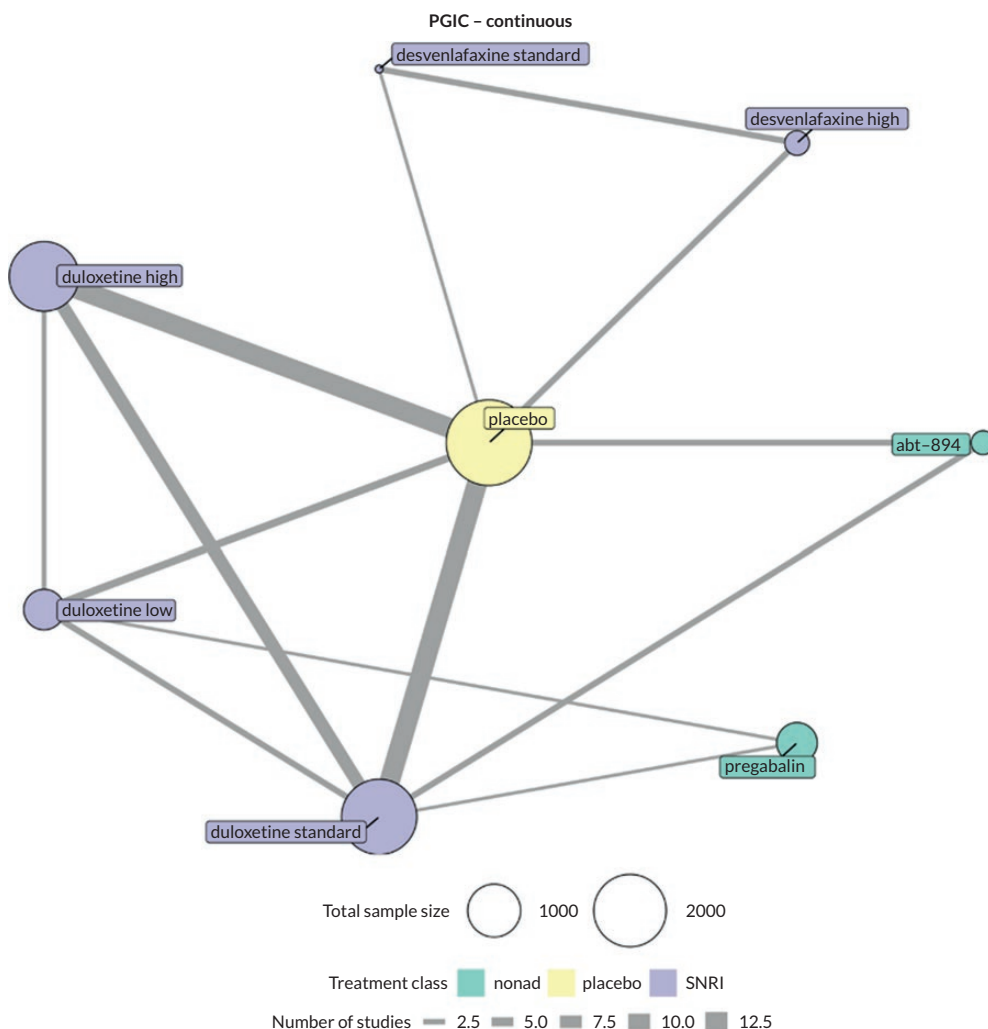
We included 24 RCTs with a total of 8415 participants (range 194–804). Of these, 12 studies compared against only placebo, 3 were multiarm trials with another active comparator, and 9 were dose-comparison trials. There were seven different interventions, and some comparisons were informed only by direct evidence from one study. We judged 15 studies as at high risk of bias overall.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

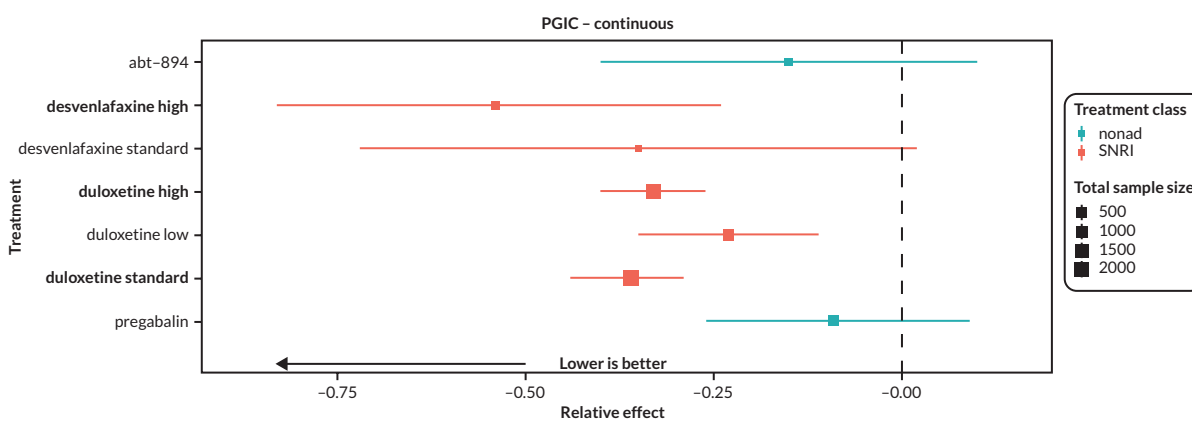
- duloxetine low dose
- duloxetine standard dose
- duloxetine high dose.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3, Table 25](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 22](#) and the forest plot of placebo comparisons is presented in [Figure 23](#).



**FIGURE 22** Patient Global Impression of Change (continuous): network plot.



**FIGURE 23** Patient Global Impression of Change (continuous): summary forest plot (SMD with CrIs).

### Ranking of antidepressants

Duloxetine standard and high doses were the highest-ranked antidepressants, with a small to moderate effect (SMD  $-0.36$ , 95% CI  $-0.44$  to  $-0.29$  and SMD  $-0.33$ , 95% CI  $-0.40$  to  $-0.26$ , respectively). Duloxetine low dose was the lowest-ranked antidepressant with a small effect (SMD  $-0.23$ , 95% CI  $-0.35$  to  $-0.11$ ).

A visual representation of the cumulative rankings for every intervention included in the analysis did not alter interpretations. Both UME and node-splitting models showed evidence of inconsistency. The highest Bayesian  $p$ -value ( $p = 0.03$ ) suggested that direct evidence overestimated the effectiveness of high-dose duloxetine versus placebo compared to indirect evidence, resulting in the strength of evidence being downgraded. These figures are available in [Report Supplementary Material 1](#).

### Class, condition and risk of bias

We were unable to run models including class, condition and risk of bias. We were unable to analyse class as there was only one class present in the network (SNRI). We were unable to analyse condition and risk of bias due to the high risk of overfitting.

### Confidence in network meta-analysis

The design-by-treatment test showed no evidence of inconsistency between the direct and indirect evidence in the network ( $\chi^2 = 14.98$ ,  $p = 0.13$ ), and no antidepressant had an  $I^2$  value higher than 40%. We rated duloxetine standard and high doses as having moderate-certainty evidence as a result of incoherence. We downgraded duloxetine low dose to moderate certainty due to some concerns regarding within-study bias in addition to network inconsistency.

### Serious adverse events

**TABLE 12** Serious adverse events: summary of findings

Estimates of effects, Crls and certainty of the evidence for serious adverse events in people with chronic pain							
Bayesian NMA summary of findings table							
Total RCTs: 70 Total participants: 19,304	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (GRADE)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Desvenlafaxine high dose RCTs: 2 Participants: 912	0.51 ( $-0.27$ to 1.29)	12/221 54 per 1000	20/691 28 per 1000	26 fewer per 1000	Very low <sup>c,d,e</sup>	11 (4 to 24)	Not significantly different from placebo
Milnacipran dose unable to be categorised RCTs: 3 Participants: 272	0.66 ( $-0.95$ to 2.27)	3/69 43 per 1000	5/203 29 per 1000	14 fewer per 1000	Very low <sup>c,d,e</sup>	15 (2 to 36)	Not significantly different from placebo
Duloxetine low dose RCTs: 4 Participants: 935	0.89 ( $-0.05$ to 1.83)	11/462 24 per 1000	9/473 21 per 1000	3 fewer per 1000	Very low <sup>c,d,e</sup>	19 (6 to 32)	Not significantly different from placebo

TABLE 12 Serious adverse events: summary of findings (continued)

Total RCTs: 70 Total participants: 19,304	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (GRADE)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Duloxetine high dose RCTs: 12 Participants: 3404	0.92 (0.43 to 1.41)	33/1601 21 per 1000	40/1803 19 per 1000	2 fewer per 1000	Very low <sup>c,d,e</sup>	19 (9 to 29)	Not significantly different from placebo
Milnacipran standard dose RCTs: 7 Participants: 2474	0.94 (0.31 to 1.57)	22/1234 18 per 1000	21/1240 17 per 1000	1 fewer per 1000	Very low <sup>c,d,e</sup>	19 (9 to 31)	Not significantly different from placebo
Mirtazapine standard dose RCTs: 3 Participants: 484	0.99 (-0.83 to 2.81)	3/241 12 per 1000	3/243 12 per 1000	0 fewer per 1000	Very low <sup>d,e</sup>	10 (3 to 38)	Not significantly different from placebo
Milnacipran high dose RCTs: 7 Participants: 2826	1.08 (0.55 to 1.61)	28/1257 22 per 1000	35/1569 24 per 1000	2 more per 1000	Very low <sup>c,d,e</sup>	22 (11 to 32)	Not significantly different from placebo
Duloxetine standard dose RCTs: 15 Participants: 4589	1.16 (0.71 to 1.61)	34/1082 16 per 1000	52/2507 19 per 1000	3 more per 1000	Very low <sup>c,d,e</sup>	23 (13 to 32)	Not significantly different from placebo
Esreboxetine standard dose RCTs: 1 Participants: 833	2.25 (-0.69 to 5.19)	1/277 4 per 1000	3/556 8 per 1000	4 more per 1000	Very low <sup>c,d,e,f</sup>	27 (4 to 41)	Not significantly different from placebo
Esreboxetine high dose RCTs: 1 Participants: 558	2.75 (-0.35 to 5.85)	1/277 4 per 1000	2/281 10 per 1000	6 more per 1000	Very low <sup>c,d,e,f</sup>	28 (4 to 41)	Not significantly different from placebo

Network meta-analysis summary of findings table definitions (applied across [Tables 2-13](#)):

a Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.

b Mean rank and CrIs are presented.

c Downgraded due to within-study bias.

d Downgraded due to imprecision in the estimate.

e Downgraded due to heterogeneity in the estimate.

f Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

g Downgraded due to incoherence in the network.

#### Notes:

GRADE Working Group grades of evidence:

High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.

### Results

For serious adverse events, we report the treatment–dose model. Both treatment and treatment–dose models had studies with high levels of imprecision, but treatment–dose was selected for reporting due to its clinical utility.

We included 70 RCTs with a total of 19,304 participants (range 26–1025). Of these, 39 studies compared against placebo, 12 compared against another active comparator, 15 were dose-comparison trials, and 4 studies compared two different antidepressants against each other. There were 31 different interventions, and some comparisons were informed only by direct evidence from one trial. We judged 45 studies as at high risk of bias. We could not include data from three studies due to disconnected networks.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are as follows:

- desvenlafaxine high dose
- duloxetine high dose
- duloxetine low dose
- duloxetine standard dose
- esreboxetine high dose
- esreboxetine standard dose
- milnacipran high dose
- milnacipran standard dose
- milnacipran dose unable to be categorised
- mirtazapine standard dose.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3, Table 26](#).

There were no concerns regarding model fit. The network diagram is presented in [Figure 24](#) and the forest plot of placebo comparisons is presented in [Figure 25](#).

### Ranking of antidepressants

Data for serious adverse events were very sparse, and studies were generally underpowered to detect rare events. No antidepressant showed any significant difference when compared with placebo, and the CIs were very wide.

We undertook a visual representation of the cumulative rankings for every intervention included in the analysis. The UME model had no evidence of inconsistency. We confirmed this with node-splitting models for all 16 comparisons where it was possible to compare direct and indirect evidence. The lowest Bayesian  $p$ -value ( $p = 0.07$ ) was for the comparison of pregabalin and low-dose duloxetine. These figures are available in [Report Supplementary Material 1](#).

### Class, condition and risk of bias

We were unable to undertake further analysis of class, condition or risk of bias in networks due to small sample sizes, network geometry and the risk of overfitting.

### Confidence in network meta-analysis

We were unable to use CINeMA for this outcome due to the complexity of the network. Therefore, two review authors (HB and GS) made the judgements based on GRADE and CINeMA domains and the available results. We judged all antidepressants and doses as very low confidence, primarily due to concerns with within-study bias, heterogeneity and imprecision in the network.

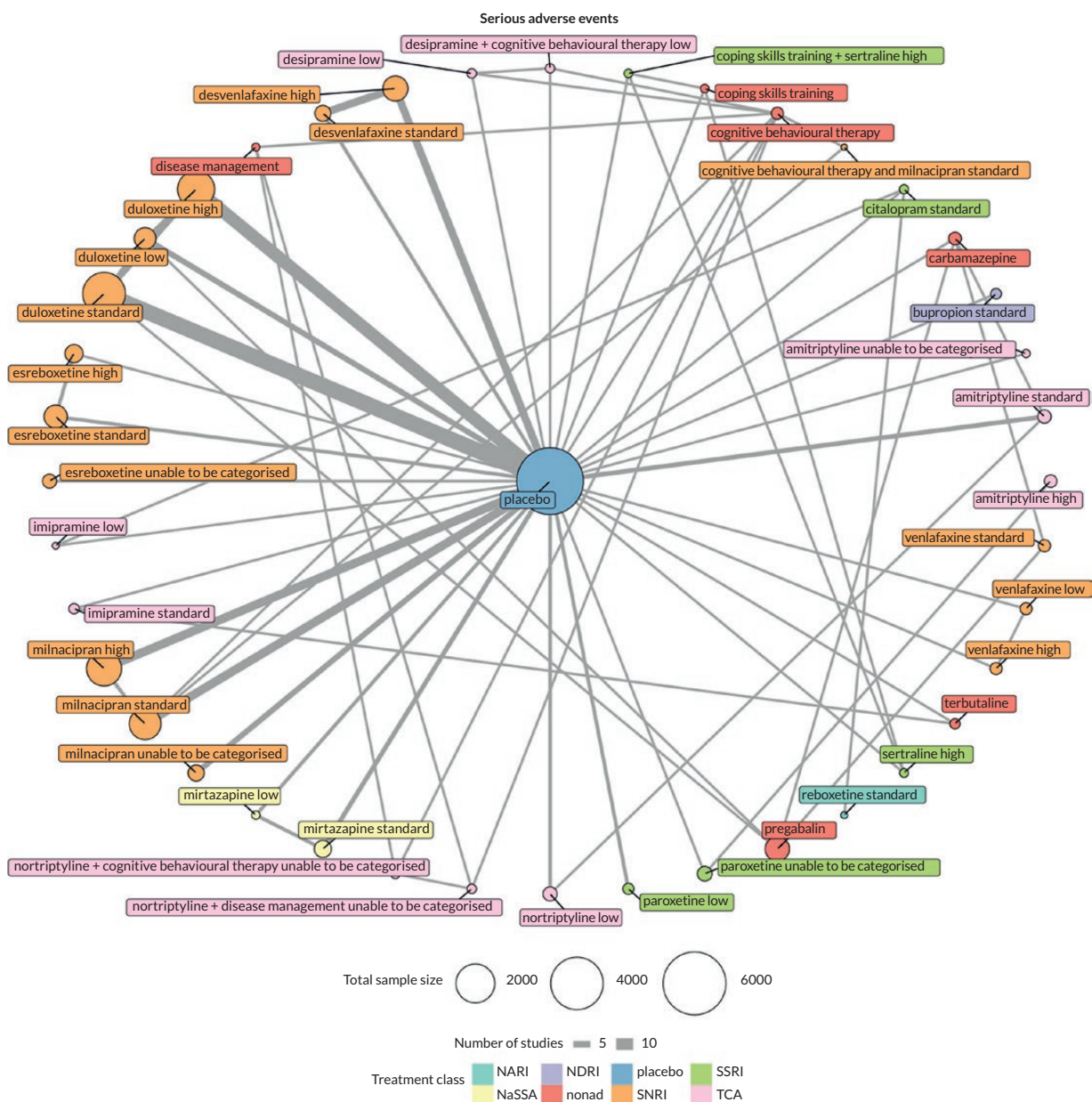


FIGURE 24 Serious adverse events: network plot.

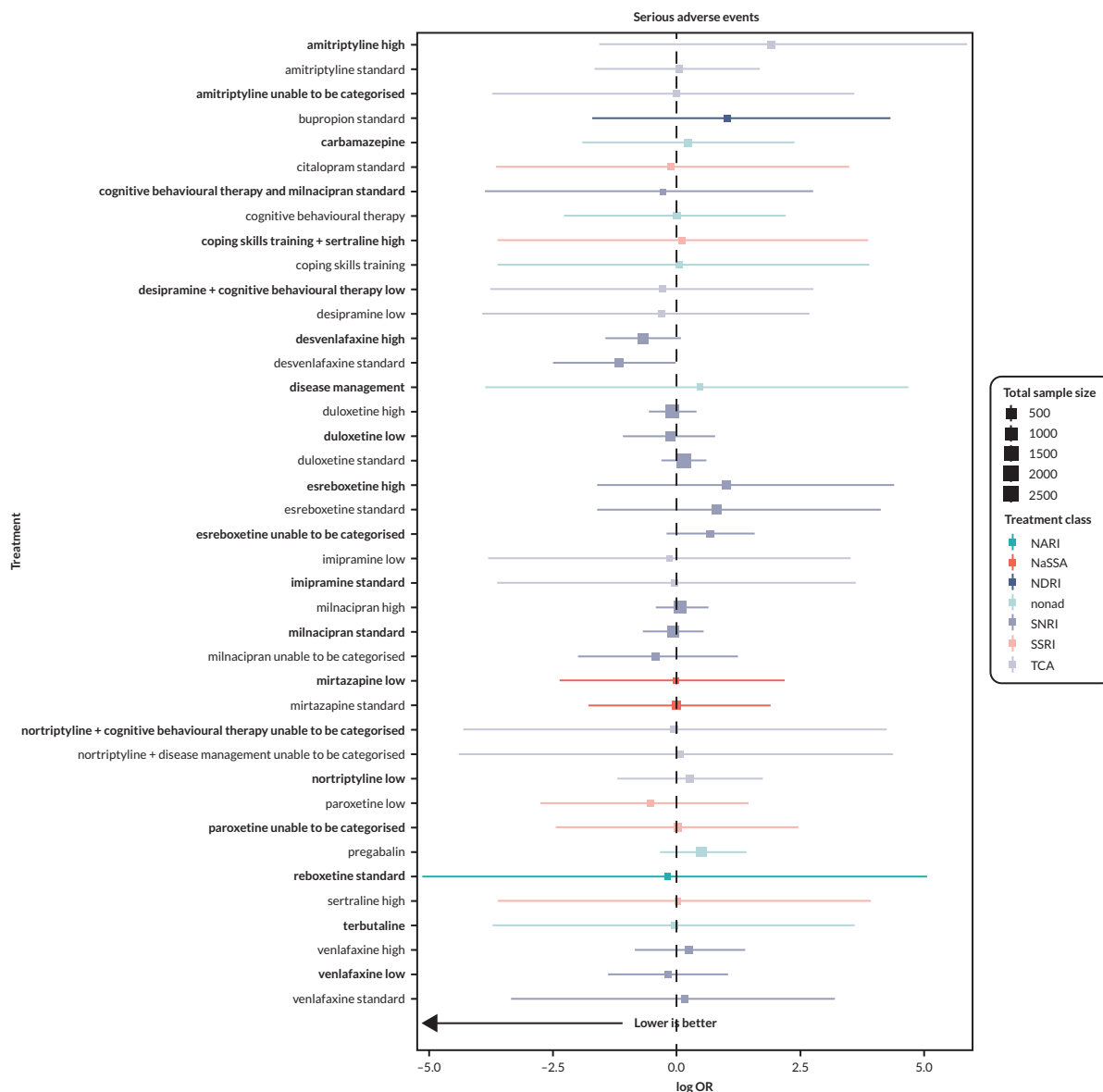


FIGURE 25 Serious adverse events: summary forest plot (log OR with CIs).

## Withdrawal

TABLE 13 Withdrawal: summary of findings

Estimates of effects, CrIs and certainty of the evidence for withdrawal from studies in people with chronic pain							
Bayesian NMA summary of findings table							
Patient or population: people with chronic pain							
Interventions: amitriptyline, desipramine, desvenlafaxine, duloxetine, esreboxetine, milnacipran, mirtazapine, nortriptyline, paroxetine, venlafaxine. All doses were combined for each antidepressant.							
Comparator (reference): placebo							
Outcome: withdrawal (for any reason)							
Direction: lower is better (i.e. fewer people withdrawing from studies)							
Total RCTs: 152 Total participants: 28,120	Relative effect (OR and 95% CI)	Anticipated absolute effect (event rate) <sup>a</sup>			Certainty of the evidence (GRADE)	Ranking <sup>b</sup> (2.5% to 97.5% CrI)	Interpretation of findings
		With placebo	With intervention	Difference			
Nortriptyline RCTs: 7 Participants: 612	0.54 (0.09 to 1.17)	101 per 1000	57 per 1000	44 fewer per 1000 (111 fewer to 15 more)	Very low <sup>c,d</sup>	13 (5 to 26)	Not significantly different from placebo
Mirtazapine RCTs: 3 Participants: 510	0.99 (0.34 to 1.64)	120 per 1000	119 per 1000	1 fewer per 1000 (76 fewer to 63 more)	Very low <sup>d,e</sup>	28 (11 to 52)	Not significantly different from placebo
Amitriptyline RCTs: 34 Participants: 2126	1.12 (0.85 to 1.39)	138 per 1000	152 per 1000	14 more per 1000 (18 fewer to 44 more)	Very low <sup>c,d,e</sup>	31 (20 to 43)	Not significantly different from placebo
Duloxetine RCTs: 45 Participants: 10,140	1.20 (1.06 to 1.34)	207 per 1000	239 per 1000	32 more per 1000 (10 more to 52 more)	Low <sup>c,d</sup>	33 (24 to 43)	Equivalent to NNT of 31
Desvenlafaxine RCTs: 2 Participants: 1105	1.25 (0.82 to 1.68)	450 per 1000	506 per 1000	56 more per 1000 (48 fewer to 129 more)	Very low <sup>c,d,e</sup>	35 (19 to 53)	Not significantly different from placebo
Milnacipran RCTs: 17 Participants: 5088	1.34 (1.12 to 1.56)	254 per 1000	314 per 1000	59 more per 1000 (22 more to 93 more)	Very low <sup>c,d</sup>	38 (27 to 49)	Equivalent to NNT of 17
Venlafaxine RCTs: 6 Participants: 624	1.40 (0.91 to 1.89)	158 per 1000	208 per 1000	50 more per 1000 (12 fewer to 104 more)	Very low <sup>c,d,e</sup>	40 (21 to 59)	Not significantly different from placebo
Esreboxetine RCTs: 2 Participants: 1389	1.42 (1.01 to 1.83)	251 per 1000	322 per 1000	71 more per 1000 (2 more to 129 more)	Very low <sup>c,d,e</sup>	41 (23 to 56)	Equivalent to NNT of 31
Desipramine RCTs: 4 Participants: 368	1.57 (1.02 to 2.12)	196 per 1000	276 per 1000	81 more per 1000 (3 more to 145 more)	Very low <sup>c,d,e</sup>	44 (24 to 61)	Equivalent to NNT of 14
Paroxetine RCTs: 9 Participants: 568	1.68 (1.23 to 2.12)	173 per 1000	260 per 1000	87 more per 1000 (32 more to 134 more)	Very low <sup>c,d</sup>	46 (28 to 60)	Equivalent to NNT of 11

**TABLE 13** Withdrawal: summary of findings (*continued*)

Network meta-analysis summary of findings table definitions (applied across [Tables 2–13](#)):

- a Anticipated absolute effect. Anticipated absolute effect compares two risks by calculating the difference between the risk of the intervention group and the risk of the control group.
- b Mean rank and CrIs are presented.
- c Downgraded due to within-study bias.
- d Downgraded due to imprecision in the estimate.
- e Downgraded due to heterogeneity in the estimate.
- f Downgraded due to incoherence in the network.
- g Downgraded due to a small number of trials and participants; we cannot draw reliable conclusions.

**Notes:**

GRADE Working Group grades of evidence:

High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

**Results**

For withdrawal, we report the treatment network. Although this model has high heterogeneity, we determined that including dose would increase the network complexity to a point where analysis would be infeasible.

We included 152 RCTs with a total of 28,120 participants (range 24–1025). Of these, 73 studies compared against placebo, 47 were multiarm trials with another active comparator, 18 were dose-comparison trials and 14 were head-to-head trials comparing two different antidepressants. There were 77 different interventions, and some comparisons were informed only by direct evidence from one trial. We rated 106 studies as at high risk of bias. We could not include data from two studies due to disconnected networks.

The antidepressants with  $\geq 200$  participants and therefore included in the summary are:

- amitriptyline
- desipramine
- desvenlafaxine
- duloxetine
- esreboxetine
- imipramine
- milnacipran
- mirtazapine
- nortriptyline
- paroxetine
- venlafaxine.

An overview of all interventions included in the analysis, the number of RCTs and the number of participants is given in [Appendix 3](#), Table 27.

There were no concerns regarding model fit. We present the network diagram in [Figure 26](#) and the forest plot of placebo comparisons in [Figure 27](#).

**Ranking of antidepressants**

Nortriptyline was the highest-ranked antidepressant. Nortriptyline, mirtazapine, amitriptyline, desvenlafaxine and venlafaxine all showed no significant difference compared to placebo for withdrawal.

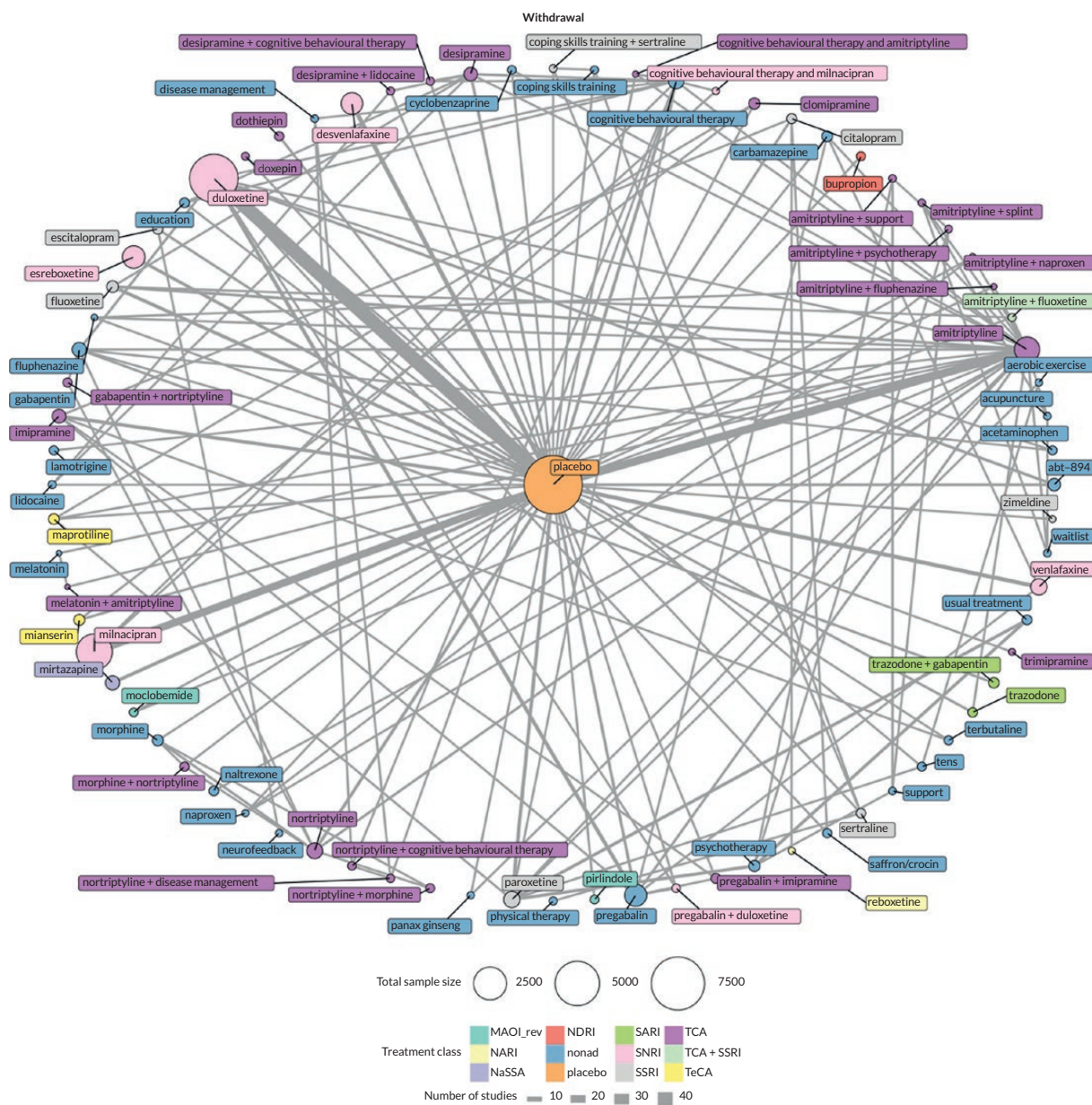


FIGURE 26 Withdrawal: network plot.

Duloxetine, milnacipran, esreboxetine, desipramine and paroxetine all showed significant effects, ranging from small to moderate.

A visual representation of the cumulative rankings for every intervention included in the analysis is available in [Report Supplementary Material 1](#). We were unable to draw any very reliable conclusions due to all antidepressants having wide, overlapping CIs.

### Exploration of heterogeneity

Due to the complexity and geometry of the network, we were only able to examine models including class, and were unable to examine condition or risk of bias.

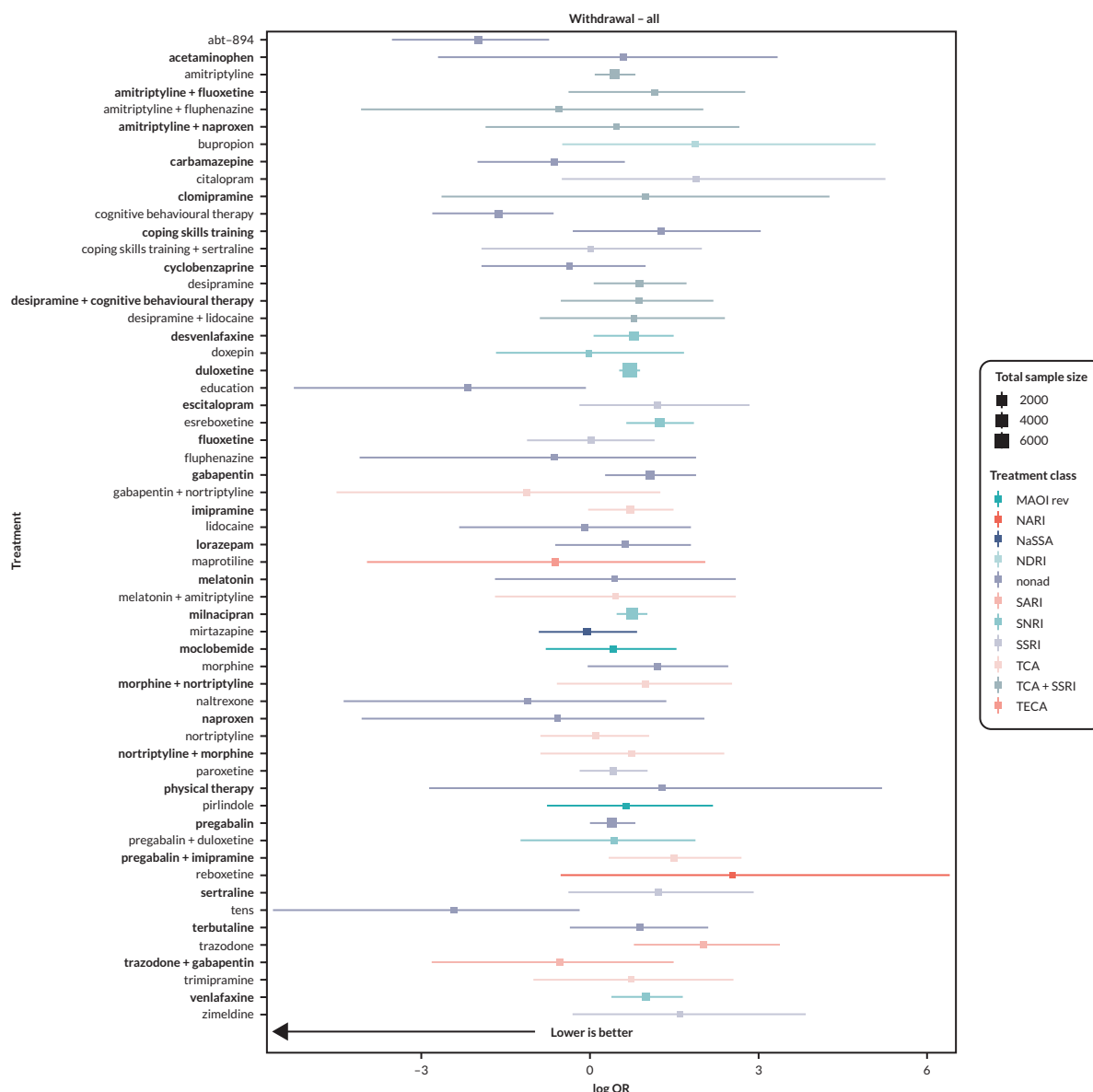


FIGURE 27 Withdrawal: summary forest plot (log OR with Crls).

**Class**

Ten classes of antidepressant were included in the analysis: SNRI, SSRI, TCA, MAOI reversible, NARI, NaSSA, NDRI, SARI, TCA+SSRI and TeCA. There was slightly higher heterogeneity than the treatment-only model (Tau = 0.33), but no evidence of inconsistency in the UME models. Half of the classes had fewer than 200 participants, leaving SNRI, SSRI, TCA, NaSSA and TeCA with reliable sample sizes. The rankings of these classes are presented in [Table 14](#).

**Confidence in network meta-analysis**

We were unable to use CINeMA for this outcome due to the complexity of the network. Therefore, two review authors (HB and GS) made the judgements based on GRADE and CINeMA domains and the available results. We judged all antidepressants except duloxetine as very low certainty, primarily due to concerns with within-study bias, heterogeneity and imprecision in the network. We rated duloxetine as low certainty, as the only antidepressant without major concerns due to imprecision.

TABLE 14 Top-ranked antidepressant classes in withdrawal analysis

Class	Antidepressant	n	Mean rank	Credible intervals	
				2.5%	97.5%
NaSSA	Mirtazapine	242	3.61	1	10
TCA	Amitriptyline Clomipramine Desipramine Dothiepin Doxepin Imipramine Nortriptyline	2593	4.33	2	7
SNRI	Duloxetine Esreboxetine Milnacipran Venlafaxine	7804	6.24	4	9
TeCA	Maprotiline Mianserin	207	6.96	2	11
SSRI	Citalopram Escitalopram Fluoxetine Paroxetine Sertraline Zimeldine	713	7.7	4	10



## Chapter 5 Differences between protocol and review

Please note that this section duplicates the information published in the open-access full Cochrane Review (<https://doi.org/10.1002/14651858.CD014682.pub2>).

There are a number of differences between the protocol and the review.

### Updating of the background section

- We removed the IASP pain categories from the background, as there is currently discourse about the clinical usefulness of primary pain, and we subsequently did not categorise pain types into these. If we had used the IASP categories, then a number of distinct pain conditions (e.g. fibromyalgia, low back pain) would have been combined, whereas there is evidence for these types of conditions being kept separate to evaluate the effects.
- We have added in some NICE guidelines to the background which were not published at the time of protocol publication.
- We have updated the literature in the 'how might the intervention work' section to reflect current understanding and theories.

### Methods

- We reported continuous pain intensity as an outcome, which was not included in the published version of the protocol. This was originally in the protocol and was removed accidentally during the protocol editing process.
- We separated adverse events and serious adverse events into separate outcomes as they are defined differently and were assessed using separate NMAs. Therefore, 'serious adverse events' was moved to a secondary outcome.
- We rated studies that imputed missing data using the LOC method as high risk, unless attrition was very low. This rule was not included in the protocol.
- We stated that we would present the primary outcomes on a scale of 0–100. As outcomes were reported on a wide variety of scales, this was not possible. Instead, we have reported the number needed to treat for an additional beneficial outcome and number needed to treat for an additional harmful outcome in the Summary of findings tables.
- We planned to use threshold analysis to analyse how much evidence needed to be added for our conclusions to change. We did not undertake threshold analysis in the review as the majority of evidence was judged to be of low or very low quality, and therefore it is already likely that new evidence will affect the conclusions.
- We have added in the criteria for antidepressant doses being categorised as 'low', 'standard' or 'high', and clarified how dose was included in the analysis in the *Data synthesis* section (moved from the subgroup analysis section).
- We omitted the 'Other bias' domain from the protocol accidentally – we did assess for this in our risk of bias assessments, and so have included this in the methods under the risk of bias section.
- We reordered parts of the methods section regarding sensitivity analyses for clarification: assessment of consistency has been moved to data synthesis, and further information regarding the sensitivity analyses has been added to the sensitivity analysis section.



## Chapter 6 Patient and public involvement

We invited people living with chronic pain themselves or caring for people living with chronic pain, to meet with the research team online, through the PPI network in Primary Care, University of Southampton. Six people took part in the meeting: three males and three females. The group lived with a variety of different conditions; it included two people below the age of 30 years and two people of black ethnicity. Specifically, the characteristics of the participants were as follows:

1. partner has fibromyalgia, older adult
2. knee pain, recent total knee arthroplasty, older adult
3. chronic hip pain post injury, older adult
4. adrenal failure and chronic fatigue syndrome, mum has fibromyalgia and arthritis, younger person
5. chronic spinal pain, previous experience of taking antidepressants for chronic pain (including amitriptyline)
6. chronic knee condition, lower back disc problem, younger person.

The main feedback included the following observations:

- PPI members were concerned about the short length of most trials, which they considered did not reflect real-life timelines of taking antidepressants.
- They were also concerned about the poor quality of adverse event reporting.

The discussion included descriptions of personal experiences of side effects from taking antidepressants for pain management, and it highlighted that addiction/dependency is a key issue, which is often not measured in trials. In addition, trials failed to measure withdrawal symptoms post intervention.

- Concern about interactions with other pain medications in real-world clinical practice that are not addressed in trials – typically anyone taking other pain medications is excluded.
- Lack of participants from ethnic minorities – the large-scale, higher-quality trials were majority Western with white participants.
- Concern about why trials are always so focused on pain intensity instead of participants' better quality of life; a feeling that quality of life is what matters, because pain is not necessarily going to get better and can be lived with if quality of life is improved.

The group agreed that the findings were really important and relevant to people with pain and clinicians. They advised that dissemination must present the same message to both clinicians and patients. They also commented that it is essential that there is a written version that is accessible for people with pain, such as the plain language summary.

In reference to dissemination, they felt strongly that it was essential to engage the community. For people with pain, dissemination comes through real-life experience, from other patients' 'witness statements'. Often patients will believe another patient over a professional, as they share the same experiences and understanding, while clinicians may have another agenda: 'it's just their job'. The meeting participants wanted to see the findings on social media, in patient groups, through patients' associations and so on. They would like to see the topic and the findings taken up by mainstream media.

Finally, they wondered whether the widespread practice of prescribing antidepressants stemmed from compassion fatigue, when there was no other option.



## Chapter 7 Discussion

Please note that this section duplicates some of the information published in the open-access full Cochrane Review (<https://doi.org/10.1002/14651858.CD014682.pub2>).

### Summary of main results from review

#### Overall

We report a NMA of 176 double-blind RCTs investigating antidepressants for chronic pain. Studies included 28,664 adult participants with a mean age of 50.6 years. The majority of studies investigated antidepressants from three classes: SNRI (74 studies), TCA (72 studies) and SSRI (34 studies). There was a variety of study designs; however, the majority of trials were placebo controlled (83 studies). The remainder compared an antidepressant against an active comparator with no placebo (22 studies), or compared two or more different doses of the same antidepressant with a placebo arm (17 studies). Most studies had a parallel-arm design (141 studies) compared to a crossover design (35 studies). Studies mainly included participants with only one type of chronic pain: 59 studies included participants with fibromyalgia, 49 neuropathic pain, 40 musculoskeletal pain and 26 included participants with other conditions (e.g. gastrointestinal, primary pain conditions, non-cardiac chest pain). Finally, 72 studies were fully funded by pharmaceutical companies, and 30 studies did not report the source of funding.

Seven studies, with a total of 156 participants, provided no useable data and were therefore omitted from the NMAs. At the time of the review, the majority of antidepressants are not licenced for use in chronic pain. Only amitriptyline and duloxetine are indicated for types of chronic pain in the British National Formulary: amitriptyline for neuropathic pain, and duloxetine for diabetic neuropathy.<sup>41,130</sup>

The following results are based on NMA. One study<sup>131</sup> reported the results separately according to the type of pain condition. This study was stratified into two to include the results for both conditions.

#### Primary efficacy outcomes

For the primary efficacy outcomes (substantial pain relief, pain intensity and mood), duloxetine was consistently the highest-ranked antidepressant that had data from over 200 participants in total across trials, and the only antidepressant with robust evidence that showed an effect with moderate-certainty evidence. For substantial pain and pain intensity, standard-dose duloxetine was equally as efficacious as high dose. For pain intensity and mood, milnacipran also showed reliable effectiveness, with moderate-certainty evidence. At a class level, SNRIs were the only class to have an effect with reliable evidence. For pain intensity, we removed one study that showed improbable effects from the data extracted from the published article.<sup>129</sup> We e-mailed the study authors for clarification but received no response.

#### Secondary efficacy outcomes

Across all the secondary efficacy outcomes (moderate pain relief, physical function, sleep, quality of life and PGIC) duloxetine and milnacipran were the highest-ranked and most trustworthy antidepressants, respectively. Very few other antidepressants had been studied in over 200 participants, and those that had been were ranked as having very low-certainty evidence. For both duloxetine and milnacipran, standard doses were equally as effective as high doses, although effects for both were small.

#### Safety

We extracted adverse event, serious adverse event, and withdrawal data from the studies included in the review. The data for these outcomes were poor. Although we have reported the ranking of antidepressants in the Summary of findings tables, the quality and certainty of this evidence for all antidepressants and doses is very low, and we cannot draw any reliable conclusions from the analyses.

## Overall completeness and applicability of evidence

We were able to draw some conclusions about the effectiveness and rankings of antidepressants in the efficacy and safety of their use for treating chronic pain. The evidence is particularly lacking for long-term outcomes and safety data.

### Participants

The sample of participants in the included studies was mostly female (68.3%) and had a mean age of 50.6 years. Most studies had a minimum pain intensity inclusion criterion, with 92 studies requiring participants to score  $\geq 4$  on a 0–10 scale or equivalent at baseline, and most participants reported experiencing pain for over 1 year.

Our inclusion criteria for participants were strict, and we required the study population to have had pain for 3 months or longer. If this time frame was not explicitly reported by the study or required for a diagnosis of the pain condition, then the study was excluded. Therefore, we excluded six studies from our full-text screening with a study population described as having a 'chronic' pain condition without information regarding duration. This may mean that we excluded other relevant studies, but we believe the number of studies to be affected by this to be minimal.

### Interventions

There were 89 different interventions included in the review, 26 of which were antidepressants. We included all interventions that matched the inclusion criteria regardless of dose, formulation and route of administration. Only four antidepressants were investigated in more than 10 studies. The only antidepressant that had robust studies and evidence is duloxetine, with 43 studies and a total of 11,608 participants randomised. Participants in duloxetine trials accounted for over one-third of all the participants included in this review. Milnacipran also showed some reliable evidence across outcomes, with 11 studies and a total of 5083 participants. Amitriptyline was investigated in 43 studies, with a total of 3372 participants, although the certainty of this evidence was very low, and only 3 studies randomised more than 200 participants. Fluoxetine was the fourth antidepressant to be included in more than 10 studies, but the quality and certainty of the evidence was very low, with 11 studies including 630 participants in total. All other antidepressants were included in fewer than 10 studies.

### Study designs and comparisons

A variety of study designs were used by studies included in the review. Half the studies included in the review were two-arm, parallel designed trials comparing antidepressant to placebo (89 out of 176 studies). There were also dose-comparison trials, comparisons against active comparators, and combined antidepressant interventions (e.g. antidepressant + psychological therapy), and a number of studies included multiple types of these comparisons. Some of the combined antidepressant comparisons precluded full analysis in the NMA as we were unable to isolate the effects of the antidepressant alone. There were few head-to-head trials comparing two antidepressants with a placebo arm for reference.

The majority of studies provided useable data for the primary efficacy outcomes: 131 studies measured pain intensity, and 87 measured mood. Although these figures represent the majority of studies, it is evident that a large number of trials in chronic pain do not report these key outcomes. In the review, over half of trials did not measure mood, and almost one-third did not measure or report pain intensity. Despite the 2005 publication of the IMMPACT guidelines<sup>79</sup> for core outcomes of chronic pain trials, only 44 and 43 studies reported the proportion of participants achieving 50% and 30% pain relief, respectively. For the secondary outcomes, around one-third of studies reported physical function, less than one-quarter reported sleep and only one-quarter reported quality of life.

All outcomes aside from withdrawal used self-reported measures. There was considerable heterogeneity in the outcome measures used across all outcomes such that SMD was required for the continuous outcomes. Additionally, studies reported a mix of change scores (change in outcome from baseline to

post intervention) and post-intervention scores. As we had to use SMD, this meant that we could not build one NMA including all data for each outcome; rather, we were required to build both change score and post-intervention score models and subsequently decide which model to report for each outcome. Typically, larger, pharmaceutical-funded trials reported change scores, while smaller trials reported post-intervention scores. Future reviews would benefit from studies reporting both types of scores, so that results can be combined for a holistic evidence synthesis. We found that the data for the safety outcomes were particularly poor; adverse events were reported in various different ways across studies, and studies were often not powered adequately or had not lasted long enough to detect events.

### **Mood**

As antidepressants are primarily designed and used to manage depression, and low mood is a common comorbidity with chronic pain, we planned to explore their impact upon mood in this analysis in several ways.

First, we planned to undertake a subgroup analysis exploring whether there were any differences in outcomes between studies reporting a primary aim of targeting pain and those reporting a main aim of targeting mood. We were unable to undertake this analysis as only two studies had a main aim of targeting mood. In contrast, 144 studies had a main aim of targeting pain.

Second, we planned to undertake analyses examining differences in outcomes for studies stratified by levels of depression at baseline (none, mild, moderate and severe as defined by the diagnostic tools used). The majority of studies excluded participants with diagnoses of major depressive disorder and other mental health conditions. Because of this, baseline measures of depression and/or anxiety failed to exceed average scores of mild depression at baseline.

As we were unable to undertake these analyses, we were unable to assess the effect of depression and mood on the outcomes of the NMA, and unable to draw any meaningful conclusions regarding the mood outcome.

### **Follow-up**

We were only able to undertake analyses at the post-intervention time point because only a small number of studies had follow-up periods of any length (6/176 studies). Therefore, we were unable to draw any conclusions regarding the long-term efficacy and safety of using antidepressants for chronic pain.

## **Quality of the evidence**

### **Overall quality**

We assessed the quality of the evidence using CINeMA (and ROB-MEN and GRADE where appropriate). Across the outcomes, the only antidepressant with consistently robust evidence was duloxetine, followed by milnacipran. All other antidepressants were judged as having low or very low certainty of evidence. The most common reasons for downgrading comparisons were within-study bias, imprecision in the NMA (wide CrIs), and small numbers of studies and participants. Additionally, all evidence for safety data was graded as very low due to heterogeneity, imprecision and sparsity of data.

### **Risk of bias**

Overall, the risk of bias for included studies was relatively high. Using the Cochrane Risk of Bias Tool 1 resulted in 116 studies being defined as at high risk of bias overall. Evidence was often downgraded due to within-study bias across antidepressants and outcomes. There are several points relating to risk of bias to be discussed. The common method of deciding the overall rating of a study's risk of bias stipulates that if any one domain is high risk, then the whole study is rated as at high risk of bias. As we included studies that compared antidepressants to other active comparators, this included interventions

whose designs inherently require participants and trial staff to be unblinded (e.g. psychological therapies). To be consistent with other studies in the review, these trials have been rated as having a high risk of bias for the blinding domains, but it has been recognised previously that these domains are not appropriate for these interventions, and in previous reviews, these domains have been omitted.<sup>77</sup>

Additionally, we found that a number of studies simply do not report the information required needed to make a judgement. Of the 60 studies rated as 'not high' risk of bias, over half had three or more domains judged as 'unclear'. Therefore, this raises concerns as to the reporting quality of these trials, an ongoing problem in health research.<sup>132</sup> There are a number of clinical trial reporting guidelines available which these trials have not abided by, which suggests that some of the studies might have been rated as at high risk of bias if the correct information had been provided.

### **Heterogeneity**

We found substantial heterogeneity in direct comparisons and entire networks across outcomes when lumping doses of each treatment together in the NMAs. Where this was evident, splitting treatments by dose removed heterogeneity for most outcomes. Therefore, most of the outcomes were analysed using a split dose model. Further exploration of heterogeneity by including antidepressant class and pain condition had to be balanced against the risk of overfitting multiple models.<sup>91</sup> The decision process for this is discussed within each outcome results section.

### **Imprecision**

Imprecision was a problem across most of our NMAs. Of the 26 different antidepressants included in our review, only four were used in more than 10 studies. Although we included all treatments in each analysis, for each outcome we graded any antidepressants with fewer than 200 participants in the antidepressant arms as very low by default and excluded these from the written summaries and Summary of findings tables. The remaining networks were generally robust at a network level, but problems remained with network connectivity relying on single trials. Imprecision was a major problem for safety data, particularly adverse events and serious adverse events, meaning that we cannot be sure of the true effect for these outcomes.

### **Inconsistency**

For each outcome, we used UME and node-splitting models to assess inconsistency in treatment and split treatment-dose networks and found no evidence of inconsistency for the primary outcomes. Therefore, we concluded that transitivity across sex, age and pain condition was valid in our models.

### **Publication bias**

We used ROB-MEN<sup>133</sup> to assess publication bias in the review. For the primary outcomes, we were only able to produce funnel plots for the duloxetine-placebo comparison as it was the only comparison with over 10 studies. These funnel plots showed some evidence of publication bias, and therefore the comparisons were rated as 'some concerns'. As all other antidepressants tended to report small effects with small numbers of studies and participants, we judged all comparisons to have 'some concerns'.

### **Potential biases in the review process**

We minimised the potential for bias in the review process as much as possible. We published our protocol through the Cochrane Library and followed this for the review process. We had an extensive search strategy that included six databases and also searched clinical trial registries for unpublished and ongoing trials. The chance of a missed trial is minimal, and even more minimal is the chance of any missed trial having a substantial effect on the overall results.

Screening, data extraction and risk of bias assessments were completed in duplicate and independently by two researchers, with all disagreements resolved by discussion. Where possible, we contacted authors to request missing data, but the response rate from authors was low. Where the study had a

clinical trial registry, we collected data that were not reported in the published paper from the results section of the registry.

We used CINeMA<sup>92</sup> and ROB-MEN<sup>133</sup> to assess our confidence in the results. The final interpretation and judgements were made by two review authors in discussion.

## Agreements and disagreements with other studies or reviews

To our knowledge, this is the only NMA that has examined all antidepressants for all types of chronic pain; previous reviews in this topic area have focused solely on one pain condition or one antidepressant, or have examined efficacy by drug, dose and pain condition. There have been a number of systematic reviews and meta-analyses over the past decade examining antidepressants for different types of pain conditions, the majority of which were Cochrane Reviews.

For neuropathic pain, multiple reviews have shown there is no high-quality or high-certainty evidence for the efficacy of amitriptyline, desipramine, imipramine, milnacipran, nortriptyline or venlafaxine.<sup>52-57</sup> However, there was moderate-certainty evidence that duloxetine is efficacious for diabetic peripheral neuropathy.<sup>134</sup> For fibromyalgia, reviews show that there was no unbiased evidence that amitriptyline, desvenlafaxine, venlafaxine or SSRIs were better than placebo, but there is low-certainty evidence that duloxetine, milnacipran and mirtazapine are efficacious.<sup>61,62</sup> Finally, for musculoskeletal pain, two reviews found no clear evidence to support the use of antidepressants for low back pain,<sup>64,65</sup> though a recent systematic review and meta-analysis showed moderate-certainty evidence for SNRIs for low back pain.<sup>135</sup> The majority of studies in Ferreira and colleagues' review and meta-analysis investigated chronic low back pain, although acute low back pain studies were also included.<sup>136</sup>

Although we were unable to examine the outcomes by condition, our results are broadly in line with previous reviews. We found no high-quality or certain evidence for the efficacy of amitriptyline, desipramine, desvenlafaxine, imipramine, mirtazapine, nortriptyline or venlafaxine in any of our outcomes. Our review and NMA found that duloxetine had robust evidence and was the highest-rated antidepressant for the majority of outcomes. For most outcomes, milnacipran was the second most efficacious antidepressant, although the certainty of evidence ranged between very low and moderate. For outcomes where a treatment-dose model was used, standard and high doses of both duloxetine and milnacipran were equally effective.

## Implications

### Implications for practice

#### For people with chronic pain

Research from RCTs suggests that duloxetine is more effective than other antidepressants (including amitriptyline) for management of chronic pain. For people with chronic pain considering trying an antidepressant for pain relief, it may be worth trying duloxetine first before other antidepressants. However, it is important to acknowledge that there is no 'one size fits all' with both antidepressants and pain. Adopting a person-centred approach is critical.

#### For clinicians

Amitriptyline was not among the highest-ranked antidepressants in terms of efficacy for either substantial pain relief or reduction in pain intensity. The evidence suggests that generic duloxetine could be the first option when considering the use of antidepressants for chronic pain management. Additionally, for duloxetine there is often no benefit to using a high dose; using a standard dose (60 mg) is often as effective as using a high dose (> 60 mg). We were unable to be certain about the

adverse events and harms for any antidepressant, so this is important to consider when prescribing antidepressants for chronic pain.

### ***Implications for policy-makers and guidelines***

A full analysis of international guidelines is out of scope, but the NICE guidelines for the treatment of chronic primary pain recommend antidepressants as the only pharmacological treatment option.<sup>30</sup> In these guidelines, NICE specifically recommends amitriptyline, citalopram, duloxetine, fluoxetine, paroxetine or sertraline, with no recommendations regarding dose. Our review and analyses found moderate- to high-certainty evidence for only duloxetine in the management of chronic pain; evidence for amitriptyline, citalopram, fluoxetine, paroxetine or sertraline was of low quality and of very low certainty. Our findings challenge current guidelines on several grounds. They suggest that there is insufficient evidence for NICE guidelines to recommend amitriptyline on the same level as duloxetine. It is important to note that the issue is a lack of quality evidence for amitriptyline, not necessarily evidence showing an absence of benefit. Therefore, guideline recommendations could be improved by including an evidence-based ranked hierarchy rather than a broad recommendation of several interventions, some of which are only supported by weak evidence. This is already implemented elsewhere: the Japanese guidelines categorise their recommendations with the qualifications 'weakly recommend', 'moderately recommend' and 'strongly recommend' for amitriptyline and duloxetine for each pain condition.<sup>29</sup> A similar approach would be appropriate for NICE guidelines, particularly when considering the weak evidence for the other antidepressants (citalopram, fluoxetine, paroxetine, sertraline), all of which are currently recommended equally for chronic primary pain. Second, our analyses by pain condition demonstrate the effectiveness of duloxetine across three major groups of people with pain: fibromyalgia, musculoskeletal (low back pain and osteoarthritis), and neuropathic pain. It is therefore unclear why the recommendation of antidepressants varies across the NICE guidelines for these conditions. From our findings, duloxetine should be recommended for all of these conditions.

### ***Implications for funders of the intervention***

Currently, amitriptyline is the most common and first-line antidepressant prescribed for the management of chronic pain; however, there are no large high-quality trials to support this position. There is also a lack of head-to-head trials where multiple antidepressants are compared in the same trial. It is important to recognise that there are no long-term safety data available for any antidepressant used for chronic pain treatment, and that collection and reporting of such data during trials is essential.

### ***Implications for research***

#### **General implications**

- For all antidepressants aside from duloxetine, there is a lack of high-quality, robust trials to establish effectiveness and safety. Amitriptyline and milnacipran particularly require further research; amitriptyline as this is the most common antidepressant prescribed for chronic pain management, and milnacipran as it has consistently ranked equivalent or very close to duloxetine. Amitriptyline was the most common antidepressant that our PPI group were familiar with, and they were concerned about its effectiveness and safety.
- SNRIs as a class require further research. Duloxetine and milnacipran were consistently the highest-ranked antidepressants across outcomes. Research to identify and explore the mechanisms underpinning the effectiveness of these antidepressants is required.
- The relationship between chronic pain and depression deserves further attention. It is common in trials of analgesics to exclude participants with comorbid mental health disorders such as clinical depression, anxiety or psychosis. As a consequence, we know nothing of the effects of antidepressants on pain in these populations. Further, depression and anxiety are common consequences of chronic pain, and often co-exist. Although the dosing schedules of antidepressant medicines are different when prescribed for analgesia rather than depression (typically smaller), there is a possibility of dual effect, but it is not possible to study this in these trials.

## Design implications

- Longer trials are required: there is no evidence regarding the long-term efficacy or safety of using antidepressants for the treatment of chronic pain. This is critical as it is likely that patients will be prescribed antidepressants for long periods of time, and currently we do not know if there are likely to be any harms related to this. This was a critical point raised by our PPI group, as the majority of members had experience of taking antidepressants for their pain over multiple years.
- Head-to-head trials between antidepressants are required to accurately measure the effects of antidepressants for chronic pain.
- Larger sample sizes: there is no need for small trials; sufficient sizes are required to establish effect.
- There is a need for pragmatic trials with more complex designs to address changes in medication. Pragmatic trial designs that account for individual differences have been recommended for over a decade,<sup>136</sup> yet the majority of studies still have a two-arm placebo-controlled design.

## Measurement implications

- There is now guidance on the optimal conduct and reporting of clinical trials, and specific guidance on the reporting of pain trials: the CONSORT and IMMPACT recommendations.<sup>79,137</sup> These should be adhered to in order to reduce research waste and efficiently inform clinical decision-making.
- Where applicable, both post-intervention and change scores should be reported to enable comprehensive evidence synthesis.
- Adverse events should be reported following the CONSORT guidelines, as highlighted many times previously.<sup>138,139</sup>

## Equality, diversity and inclusion

### Research participants

As this study was a systematic review, it did not recruit participants. We reviewed the equality, diversity and inclusion aspects of the trials in the review. Most studies were undertaken in Western populations, primarily USA and Western Europe, with the majority of participants being of white ethnicity. Furthermore, nearly all studies excluded participants with mental health comorbidities, including depression and anxiety. Therefore, the studies in this review do not reflect real-world settings. It is essential that future research in this area focuses on equality, diversity and inclusion principles, particularly the inclusion of participants of non-white ethnicities and with comorbidities.

### Research team and wider involvement

The research team is a multidisciplinary team including psychologists, pain researchers, a psychiatrist, a pharmacist, an anaesthetist and statisticians, ranging from junior members (pre- and post-PhD research associates) to professors. The review was led by an early-career researcher (HB), and provided opportunities for development including training, gaining of new skills, conference attendance and presenting, and dissemination.

Equality, diversity and inclusion was championed in our PPI groups. We had representation from both men and women with a range of chronic pain conditions, a mix of younger and older adults, and members from both white and black ethnicities.

## Summary

### Strengths and limitations

This is the first and only full systematic review with NMA for all antidepressants in all chronic pain conditions. The strength of this review is in the design and methodology, which applied gold-standard

guidance at all stages of the project. We believe our search was comprehensive, that we minimised human bias by using double extraction, and that we graded each trial on the quality of the methodology, therefore providing information on the certainty we can attach to each finding. We are also able to compare direct and indirect evidence for some antidepressants in some conditions.

The main limitation of this study is that we were unable to assess the harms associated with any antidepressant. None of the ways in which adverse events were measured in individual trials produced results that were interpretable in meta-analysis or NMA. The inconsistency and unreliability of adverse event reporting has been raised previously,<sup>139</sup> and improvement is still needed. This is important to consider in relation to the recommendations for antidepressants given by guidelines; a judgement has to be made as to the balance between efficacy and potential adverse events.

A second limitation is that the majority of trials included in our analyses excluded people with diagnoses of depression, anxiety, and other mental health conditions. Across trials, the average scores at baseline on mood measures were all within the 'normal' or 'subclinical' range. These results may not therefore be representative of the large group of people who experience any mental health conditions alongside chronic pain.

Finally, as the average length of trials was only 11.5 weeks, there are no data on the longer-term efficacy of any antidepressants in the management of chronic pain. This is critical, as clinically patients with chronic pain usually use medicines for long periods of time to consistently manage their pain.

## Chapter 8 Conclusion

In conclusion, our findings indicate that duloxetine is efficacious across pain conditions for a number of outcomes. There is some evidence that milnacipran is also effective for chronic pain, although it is currently unavailable for use in the UK. For duloxetine, the effects are similar across different pain conditions, and standard dose is as effective as higher dose. There is no certain evidence for any other antidepressant. It is important to caveat these conclusions with two points. First, while the evidence for the use of amitriptyline is of low certainty, there is also no evidence we can be certain of against the use of amitriptyline. Second, the findings should not be read as an encouragement to prescribe an antidepressant where other non-pharmacological intervention could be equally effective, especially in the absence of good evidence on side effects and safety. Despite these caveats, the overwhelming evidence for efficacy of duloxetine requires guideline updates to reflect these findings.



# Additional information

## Acknowledgements

We acknowledge and thank the NIHR Health Technology Assessment programme (HTA) for funding this review.

We thank Joanne Abbott (Information Specialist for the Cochrane PaPaS Review Group) for developing and running the search strategy, and Iris Gordon (Information Specialist for the Cochrane Eyes and Vision Review Group) for peer reviewing the search strategy.

We thank John Birkbeck and Lawrence Yule for their work in reformatting the plots and graphs for this manuscript.

We thank the peer reviewers Dr Sarah Nevitt, Kevin Pacheco-Barrios MD, Dr Christina Abdel Shaheed, and Prof Amanda C de C Williams, and consumer reviewers Harrison Nelson and Stella O'Brien.

The Cochrane PaPaS Review Group supported the authors in the development of this review.

## Contributions of authors

**Hollie Birkinshaw** (<https://orcid.org/0000-0003-0853-2995>) (Senior Research Associate, with expertise in pain research and the psychology of pain) co-ordinated the review; was responsible for screening and selection of studies from the search results, data extraction, and risk of bias assessments; cleaned the data for analysis; assessed the certainty of the evidence using CINeMA; drafted the manuscript; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Claire Friedrich** (<https://orcid.org/0000-0002-5841-4324>) (Research Associate with expertise in systematic review methodology) was responsible for screening and selection of studies from the search results, data extraction, and risk of bias assessments; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Peter Cole** (<https://orcid.org/0000-0003-2762-2144>) (Consultant in Anaesthesia and Pain Medicine, with expertise in pain management) conceived, designed, and gained funding for the review; was responsible for decisions requiring clinical knowledge; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Christopher Eccleston** (<https://orcid.org/0000-0003-0698-1543>) (Professor of Medical Psychology, with expertise in systematic reviews, clinical trials, and psychological aspects of pain) conceived, designed, and gained funding for the review; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Marc Serfaty** (<https://orcid.org/0000-0001-8388-0776>) (Psychiatrist and a Professor of Psychotherapy, with expertise in mood disorders and pain) conceived, designed, and gained funding for the review; was responsible for decisions requiring clinical knowledge; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Gavin Stewart** (<https://orcid.org/0000-0001-5684-1544>) (Senior Lecturer in Evidence Synthesis, with expertise in network meta-analysis and evidence synthesis) conceived, designed, and gained funding for the review; undertook data analysis through network meta-analyses; assessed the certainty of the

evidence using CInEMA; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Simon White** (<https://orcid.org/0000-0003-0096-251X>) (Clinical Pharmacist and Reader in Pharmacy Practice, with expertise in pharmacology) conceived, designed, and gained funding for the review; was responsible for decisions requiring clinical knowledge; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Andrew Moore** (<https://orcid.org/0000-0002-7932-5136>) (retired researcher into pain, with expertise in systematic reviews and clinical trials for pain) contributed to the interpretation of findings and the writing and editing of the manuscript.

**David Phillippo** (<https://orcid.org/0000-0003-2672-7841>) (Research Fellow in Evidence Synthesis, with expertise in network meta-analysis methods) supported Gavin Stewart in data analysis through network meta-analyses; and contributed to the interpretation of findings and the writing and editing of the manuscript.

**Tamar Pincus** (<https://orcid.org/0000-0002-3172-5624>) (Professor in Health Psychology, Dean of the Faculty for Environmental and Life Sciences, with expertise in psychological factors associated with pain) conceived, designed, and gained funding for the review; contributed to the interpretation of findings and the writing and editing of the manuscript; and will be responsible for the update of this review.

### Disclosure of interests

**Full disclosure of interests:** Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/MKRT2948>.

**Primary conflicts of interest:** Tamar Pincus reports grants from the Medical Research Council and Versus Arthritis and NIHR, consulting fees from Inspira Centro and Agence EBP, conference presentation support from EFIC, and the role of an Officer for Eurospine. Christopher Eccleston reports a grant from the Medical Research Council and Versus Arthritis, funding from Orion Pharma, royalties from Oxford University Press, and patents from the USA and Finland. Peter Cole is a consultant anaesthetist in pain management for Oxford University Hospitals Trust. Marc Serfaty is a member of the HTA General Committee. All other authors report no competing interests.

### Data-sharing statement

All data requests should be submitted via e-mail to the corresponding author. Access to data sets developed for this study may be permitted following review by the author team and University of Southampton and Newcastle University's Research Governance procedures.

### Ethics statement

As this was a systematic review, no primary data collection was undertaken. Therefore, ethical approval was not required for this study.

## Information governance statement

The University of Southampton is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679.

Under the Data Protection legislation, The University of Southampton is the Data Controller, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here: [www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page](http://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

The data for this report contained no personal data.



## References

1. Bark A. *Chronic Pain has Arrived in the ICD-11*. IASP News. URL: [www.iasp-pain.org/PublicationsNews/NewsDetail.aspx?ItemNumber=8340&navItemNumber=643](http://www.iasp-pain.org/PublicationsNews/NewsDetail.aspx?ItemNumber=8340&navItemNumber=643) (accessed 1 July 2020).
2. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, *et al*. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain* 2019;**160**:19–27.
3. Moore RA, Cai N, Skljarevski V, Tölle TR. Duloxetine use in chronic painful conditions – individual patient data responder analysis. *Eur J Pain* 2014;**18**:67–75.
4. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. *BMJ Open* 2016;**6**:e010364–12.
5. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population based studies. *Br J Anaesth* 2019;**123**:273–83.
6. Watt T, Raymond A, Ratchet-Jacquet L. *Quantifying Health Inequalities in England*. The Health Foundation; 2022. URL: [www.health.org.uk/news-and-comment/charts-and-infographics/quantifying-health-inequalities](http://www.health.org.uk/news-and-comment/charts-and-infographics/quantifying-health-inequalities) (accessed 3 April 2023).
7. Smith D, Wilkie R, Uthman O, Jordan JL, McBeth J. Chronic pain and mortality: a systematic review. *PLOS ONE* 2014;**9**:e99048.
8. Smith D, Wilkie R, Croft P, Parmar S, McBeth J. Pain and mortality: mechanisms for a relationship. *Pain* 2018;**159**:1112–8.
9. Interagency Pain Research Coordinating Committee, National Institutes of Health. *National Pain Strategy Report – A Comprehensive Population Health-Level Strategy for Pain*. National Institutes of Health; 2022. URL: [www.iprcc.nih.gov/node/5/national-pain-strategy-report](http://www.iprcc.nih.gov/node/5/national-pain-strategy-report) (accessed 3 April 2023).
10. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life and treatment. *Eur J Pain* 2006;**10**:287–333.
11. Rice A, Smith B, Blyth F. Pain and the global burden of disease. *Pain* 2016;**157**:791–6.
12. European Pain Federation. *Pain Proposal: Improving the Current and Future Management of Chronic Pain*. URL: [https://europeanpainfederation.eu/wp-content/uploads/2016/06/pain\\_proposal.pdf](https://europeanpainfederation.eu/wp-content/uploads/2016/06/pain_proposal.pdf) (accessed 1 July 2020).
13. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. *BMC Musculoskelet Disord* 2010;**11**:144.
14. Morris LD, Daniels KJ, Ganguli B, Louw QA. An update on the prevalence of low back pain in Africa: a systematic review and meta-analyses. *BMC Musculoskelet Disord* 2018;**19**:1–15.
15. Bevan S. *The Impact of Back Pain on Sickness Absence in Europe*. Lancaster: The Work Foundation; 2012.
16. Office for National Statistics. *Sickness Absence in the UK*. URL: [www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/20](http://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/20) (accessed 1 July 2020).
17. Goesling J, Clauw DJ, Hassett AL. Pain and depression: an integrative review of neurobiological and psychological factors. *Curr Psychiatry Rep* 2013;**15**:421.

18. Moore RA, Straube S, Paine J, Phillips CJ, Derry S, McQuay HJ. Fibromyalgia: moderate and substantial pain intensity reduction predicts improvement in other outcomes and substantial quality of life gain. *Pain* 2010;**149**:360–4.
19. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, *et al.* Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;**391**:1357–66.
20. Harmer CJ, Duman RS, Cowen PJ. How do antidepressants work? New perspectives for refining future treatment approaches. *Lancet Psychiatry* 2017;**4**:409–18.
21. Finnerup NB, Kuner R, Jensen TS. Neuropathic pain: from mechanisms to treatment. *Physiol Rev* 2021;**101**:259–301.
22. Kremer M, Yalcin I, Goumon Y, Wurtz X, Nexon L, Daniel D, *et al.* A dual noradrenergic mechanism for the relief of neuropathic allodynia by the antidepressant drugs duloxetine and amitriptyline. *J Neurosci* 2018;**38**:9934–54.
23. Llorca-Torralba M, Borges G, Neto F, Mico JA, Berrocoso E. Noradrenergic Locus Coeruleus pathways in pain modulation. *Neuroscience* 2016;**338**:93–113.
24. Obata H. Analgesic mechanisms of antidepressants for neuropathic pain. *Int J Mol Sci* 2017;**18**:2483.
25. Alba-Delgado C, Mico JA, Sánchez-Blázquez P, Berrocoso E. Analgesic antidepressants promote the responsiveness of locus coeruleus neurons to noxious stimulation: implications for neuropathic pain. *Pain* 2012;**153**:1438–49.
26. Eccleston C, Aldington D, Moore A, de C Williams AC. Pragmatic but flawed: the NICE guideline on chronic pain. *Lancet* 2021;**397**:2029–31.
27. Dowell D, Ragan KR, Jones CM, Baldwin GT, Chou R. CDC Clinical practice guideline for prescribing opioids for pain, United States, 2022. *MMWR Recomm Rep* 2022;**71**:1–95.
28. Korownyk CS, Montgomery L, Young J, Moore S, Singer AG, MacDougall P, *et al.* PEER simplified chronic pain guideline: Management of chronic low back, osteoarthritic, and neuropathic pain in primary care. *Can Fam Physician* 2022;**68**:179–90.
29. The Committee for Clinical Practice Guideline, for the Management of Chronic Pain. *Clinical Practice Guidelines for the Management of Chronic Pain*. Japan; 2021. URL: [https://minds.jcqhc.or.jp/docs/gl\\_pdf/G0001351/4/chronic\\_pain.pdf](https://minds.jcqhc.or.jp/docs/gl_pdf/G0001351/4/chronic_pain.pdf) (accessed 12 August 2022).
30. National Institute for Health and Care Excellence. *Chronic Pain (Primary and Secondary) in Over 16s: Assessment of All Chronic Pain and Management of Chronic Primary Pain*; 2021. URL: [www.nice.org.uk/guidance/ng193/chapter/recommendations#managing-chronic-primary-pain](http://www.nice.org.uk/guidance/ng193/chapter/recommendations#managing-chronic-primary-pain) (accessed 28 November 2022).
31. National Institute for Health and Care Excellence. *Neuropathic Pain in Adults: Pharmacological Management in Non-specialist Settings*; 2020. URL: [www.nice.org.uk/guidance/cg173](http://www.nice.org.uk/guidance/cg173) (accessed 26 August 2022).
32. National Institute for Health and Care Excellence. *Low Back Pain and Sciatica in Over 16s: Assessment and Management*; 2020. URL: [www.nice.org.uk/guidance/ng59](http://www.nice.org.uk/guidance/ng59) (accessed 28 August 2022).
33. National Institute for Health and Care Excellence. *Osteoarthritis in Over 16s: Diagnosis and Management*. 2022. URL: [www.nice.org.uk/guidance/ng226](http://www.nice.org.uk/guidance/ng226) (accessed 28 November 2022).
34. Nørregaard J, Volkmann H, Danneskiold-Samstøe B. A randomized controlled trial of citalopram in the treatment of fibromyalgia. *Pain* 1995;**61**:445–9.

35. Lee RA, West RM, Wilson JD. The response to sertraline in men with chronic pelvic pain syndrome. *Sex Transm Infect* 2005;**81**:147–9.
36. Spinhoven P, Van der Does AJW, Van Dijk E, Van Rood YR. Heart-focused anxiety as a mediating variable in the treatment of noncardiac chest pain by cognitive-behavioral therapy and paroxetine. *J Psychosom Res* 2010;**69**:227–35.
37. Gouveia N, Rodrigues A, Ramiro S, Eusébio M, Machado PM, Canhao H, *et al*. The use of analgesic and other pain relief drugs to manage chronic low back pain: results from a national survey. *Pain Pract* 2017;**17**:353–65.
38. Kurita GP, Sjøgren P, Juel K, Højsted J, Ekholm O. The burden of chronic pain: a cross-sectional survey focussing on diseases, immigration, and opioid use. *Pain* 2012;**153**:2332–8.
39. Audi S, Burrage DR, Lonsdale DO, Pontefract S, Coleman JJ, Hitchings AW, Baker EH. The 'top 100' drugs and classes in England: an update 'starter formulary' for trainee prescribers. *Br J Clin Pharmacol* 2018;**84**:2562–71.
40. Bennett Institute for Applied Data Science, University of Oxford. *OpenPrescribing.net*. [www.openprescribing.net](http://www.openprescribing.net); 2022. URL: <https://openprescribing.net/> (accessed 4 August 2022).
41. Joint Formulary Committee. *British National Formulary: Amitriptyline Hydrochloride* (online). London: BMJ Group and Pharmaceutical Press. URL: <https://bnf.nice.org.uk/drugs/amitriptyline-hydrochloride/> (accessed 4 August 2022).
42. Tamblyn R, Bates DW, Buckeridge DL, Dixon W, Forster AJ, Girard N, *et al*. Multinational comparison of new antidepressant use in older adults: a cohort study. *BMJ Open* 2019;**9**:e027663.
43. Riediger C, Schuster T, Barlind K, Maier S, Weitz J, Siepmann T. Adverse effects of antidepressants for chronic pain: a systematic review and meta-analysis. *Front Neurol* 2017;**8**:307.
44. Sinyor M, Cheung CP, Abraha HY, Lanctôt KL, Saleem M, Liu CS, *et al*. Antidepressant-placebo differences for specific adverse events in major depressive disorder: a systematic review. *J Affect Disord* 2020;**267**:185–90.
45. Coupland C, Dhiman P, Barton G, Morriss R, Arthur A, Sach T, Hippisley-Cox J. A study of the safety and harms of antidepressant drugs for older people: a cohort study analysis using a large primary care database. *Health Technol Assess* 2011;**15**:1–202.
46. Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. *BMJ* 2011;**343**:d4551.
47. Coupland C, Hill T, Morriss R, Moore M, Arthur A, Hippisley-Cox J. Antidepressant use and risk of cardiovascular outcomes in people aged 20 to 64: cohort study using primary care database. *BMJ* 2016;**352**:i1350.
48. Coupland C, Hill T, Morriss R, Arthur A, Moore M, Hippisley-Cox J. Antidepressant use and risk of suicide and attempted suicide or self harm in people aged 20 to 64: cohort study using a primary care database. *BMJ* 2015;**350**:h517.
49. Hill T, Coupland C, Morriss R, Arthur A, Moore M, Hippisley-Cox J. Antidepressant use and risk of epilepsy and seizures in people aged 20 to 64 years: cohort study using a primary care database. *BMC Psychiatry* 2015;**15**:315.
50. Jiang HY, Chen HZ, Hu XJ, Yu ZH, Yang W, Deng M, *et al*. Use of selective serotonin reuptake inhibitors and risk of upper gastrointestinal bleeding: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2015;**13**:42–50.e3.

51. National Institute for Health and Care Excellence. *Chronic Pain in Over 16s: Assessment and Management*. URL: [www.nice.org.uk/guidance/gid-ng10069/documents/draft-guideline](http://www.nice.org.uk/guidance/gid-ng10069/documents/draft-guideline) (accessed 1 September 2020).
52. Derry S, Phillips T, Moore RA, Wiffen PJ. Milnacipran for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015;2015:CD011789. <https://doi.org/10.1002/14651858.CD011789>
53. Derry S, Wiffen PJ, Aldington D, Moore RA. Nortriptyline for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015;1:CD011209. <https://doi.org/10.1002/14651858.CD011209.pub2>
54. Gallagher HC, Gallagher RM, Butler M, Buggy DJ, Henman MC. Venlafaxine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015;2015:CD011091. <https://doi.org/10.1002/14651858.CD011091.pub2>
55. Hearn L, Moore RA, Derry S, Wiffen PJ, Phillips T. Desipramine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2014;2014:CD011003. <https://doi.org/10.1002/14651858.CD011003.pub2>
56. Hearn L, Derry S, Phillips T, Moore RA, Wiffen PJ. Imipramine for neuropathic pain in adults. *Cochrane Database Syst Rev* 2014;2014:CD010769. <https://doi.org/10.1002/14651858.CD010769.pub2>
57. Moore RA, Derry S, Aldington D, Cole P, Wiffen PJ. Amitriptyline for neuropathic pain in adults. *Cochrane Database Syst Rev* 2015;2015:CD008242. <https://doi.org/10.1002/14651858.CD008242.pub3>
58. Bates D, Schultheis C, Hanes MC, Jolly SM, Chakravarthy KV, Deer TR, *et al*. A comprehensive algorithm for the management of neuropathic pain. *Pain Med* 2019;20:2–12.
59. Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Gilron I, *et al*. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. *Lancet Neurol* 2015;14:162–73.
60. Moulin DE, Boulanger A, Clark AJ, Clarke H, Dai T, Finley GA, *et al*. Pharmacological management of chronic neuropathic pain: revised consensus statement from the Canadian Pain Society. *Pain Res Manag* 2014;19:328–35.
61. Walitt B, Urrútia G, Nishishinya MB, Cantrell SE, Häuser W. Selective serotonin reuptake inhibitors for fibromyalgia syndrome. *Cochrane Database Syst Rev* 2015;2015:CD011735. <https://doi.org/10.1002/14651858.CD011735>
62. Welsch P, Üçeyler N, Klose P, Walitt B, Häuser W. Serotonin and noradrenaline reuptake inhibitors (SNRIs) for fibromyalgia. *Cochrane Database Syst Rev* 2018;2:CD010292. <https://doi.org/10.1002/14651858.CD010292.pub2>
63. Welsch P, Bernardy K, Derry S, Moore RA, Häuser W. Mirtazapine for fibromyalgia in adults. *Cochrane Database Syst Rev* 2018;8:CD012708. <https://doi.org/10.1002/14651858.CD012708.pub2>
64. Urquhart DM, Hoving JL, Assendelft WJJ, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. *Cochrane Database Syst Rev* 2008;2008:CD001703. <https://doi.org/10.1002/14651858.CD001703.pub3>
65. Koes BW, Backes D, Bindels PJE. Pharmacotherapy for chronic non-specific low back pain: current and future options. *Expert Opin Pharmacother* 2018;19:537–45.
66. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA* 2010;303:1295–302.

67. Sullivan MJL, Reesor K, Mikail S, Fisher R. The treatment of depression in chronic low back pain: review and recommendations. *Pain* 1992;**50**:5–13.
68. Tunks ER, Crook J, Weir R. Epidemiology of chronic pain with psychological comorbidity: prevalence, risk, course, and prognosis. *Can J Psychiatry* 2008;**53**:224–34.
69. Mallen CD, Peat G, Thomas E, Dunn KM, Croft PR. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract* 2007;**57**:655–61.
70. Pincus T, Vogel S, Burton AK, Santos R, Field AP. Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis Rheum* 2006;**54**:3999–4010.
71. Rusu A, Pincus T. Pain-related distress and clinical depression in chronic pain: a comparison between two measures. *Scand J Pain* 2016;**12**:62–7.
72. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;**374**:609–19.
73. Nuyen J, Volkens AC, Verhaak PF, Schellevis FG, Groenewegen PP, Van den Bos GA. Accuracy of diagnosing depression in primary care: the impact of chronic somatic and psychiatric co-morbidity. *Psychol Med* 2005;**35**:1185–95.
74. Dowrick C, Frances A. Medicalising unhappiness: new classification of depression risks more patients being put on drug treatment from which they will not benefit. *BMJ* 2013;**347**:f7140–f7140.
75. Mulder RT. An epidemic of depression or the medicalization of distress? *Perspect Biol Med* 2008;**51**:238–50.
76. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch V, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.1*. London: Cochrane; 2020.
77. Williams AC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database Syst Rev* 2020;**8**:CD007407. <https://doi.org/10.1002/14651858.CD007407.pub4>
78. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: <https://bnf.nice.org.uk/> (accessed 29 June 2022).
79. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2005;**9**:105–21.
80. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Med* 2009;**6**:e1000097–9.
81. Review Manager. *Version 5.4*. Copenhagen: Cochrane; 2020.
82. Cohen J. *Statistical Power Analysis in the Behavioral Sciences*. 2nd edn. Hillsdale (NJ): Lawrence Erlbaum Associates, Inc; 1988.
83. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* 2016;**6**:e010919–10.
84. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013;**159**:130–7.
85. Rutherford BR, Sneed JR, Roose SP. Does study design influence outcome? The effects of placebo control and treatment duration in antidepressant trials. *Psychother Psychosom* 2009;**78**:172–81.

86. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;**331**:897–900.
87. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, *et al.* Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011;**14**:417–28.
88. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: an improved method for assessing convergence of MCMC. *Bayesian Anal* 2021;**16**:667–718. <https://doi.org/10.1214/20-BA1221>
89. Betancourt M, Girolami M. Hamiltonian Monte Carlo for Hierarchical Models. In *Current Trends in Bayesian Methodology with Applications*. London: Chapman and Hall; 2015. p. 24.
90. Dias S, Sutton A, Welton N, Ades A. Evidence synthesis for decision making 3: heterogeneity-subgroups, meta-regression, bias, and bias-adjustment. *Med Decis Making* 2013;**33**:618–40. <https://doi.org/10.1177/0272989X13485157>
91. Dias S, Welton N, Sutton A, Caldwell D, Lu G, Ades A. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 2013;**33**:641–56. <https://doi.org/10.1177/0272989X12455847>
92. Phillipppo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making* 2018;**38**:200–11.
93. Nikolakopoulou A, Higgins JP, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, *et al.* CINeMA: an approach for assessing confidence in the results of a network meta-analysis. *PLOS Med* 2020;**17**:1–19.
94. Yepes-Nuñez JJ, Li SA, Guyatt G, Jack SM, Brozek JL, Beyene J, *et al.* Development of the summary of findings table for network meta-analysis. *J Clin Epidemiol* 2019;**115**:1–13. <https://doi.org/10.1016/j.jclinepi.2019.04.018>
95. Wiffen P, Moore RA. Pain leads the way: the development of evidence-based medicine for pain relief. *Int J Clin Pharmacol Ther* 2016;**54**:505. <https://doi.org/10.5414/CP202546>
96. Hudson B, Williman JA, Stamp LK, Alchin JS, Hooper GJ, Mangin D, *et al.* Nortriptyline for pain in knee osteoarthritis: a double-blind randomised controlled trial in New Zealand general practice. *Br J Gen Pract* 2021;**71**:e538–46.
97. Enomoto H, Yasuda H, Nishiyori A, Fujikoshi S, Furukawa M, Ishida M, *et al.* Duloxetine in patients with diabetic peripheral neuropathic pain in Japan: a randomized, double-blind, noninferiority comparative study with pregabalin. *J Pain Res* 2018;**11**:1857–68.
98. 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* 2002;**112**:191–7.
99. Ang DC, Jensen MP, Steiner JL, Hilligoss J, Gracely RH, Saha C. Combining cognitive-behavioral therapy and milnacipran for fibromyalgia: a feasibility randomized-controlled trial. *Clin J Pain* 2013;**29**:747–54.
100. Rowbotham MC, Arslanian A, Nothaft W, Duan WR, Best AE, Pritchett Y, *et al.* Efficacy and safety of the alpha4beta2 neuronal nicotinic receptor agonist ABT-894 in patients with diabetic peripheral neuropathic pain. *Pain* 2012;**153**:862–8.
101. Kaur H, Hota D, Bhansali A, Dutta P, Bansal D, Chakrabarti A. A comparative evaluation of amitriptyline and duloxetine in painful diabetic neuropathy: a randomized, double-blind, cross-over clinical trial. *Diabetes Care* 2011;**34**:818–22.

102. Gilron I, Chaparro LE, Tu D, Holden RR, Milev R, Towheed T, *et al.* Combination of pregabalin with duloxetine for fibromyalgia: a randomized controlled trial. *Pain* 2016;**157**:1532–40.
103. 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* 2001;**19**:697–702.
104. Boyle J, Eriksson ME, Gribble L, Gouni R, Johnsen S, Coppini DV, Kerr D. Randomized, placebo-controlled comparison of amitriptyline, duloxetine, and pregabalin in patients with chronic diabetic peripheral neuropathic pain: impact on pain, polysomnographic sleep, daytime functioning, and quality of life. *Diabetes Care* 2012;**35**:2451–8.
105. Chappell AS, Littlejohn G, Kajdasz DK, Scheinberg M, D'Souza DN, Moldofsky H. A 1-year safety and efficacy study of duloxetine in patients with fibromyalgia. *Clin J Pain* 2009;**25**:365–75.
106. Atkinson JH, Slater MA, Capparelli EV, Wallace MS, Zisook S, Abramson I, *et al.* Efficacy of noradrenergic and serotonergic antidepressants in chronic back pain: a preliminary concentration-controlled trial. *J Clin Psychopharmacol* 2007;**27**:135–42.
107. 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* 1996;**39**:1852–9.
108. Creed F, Fernandes L, Guthrie E, Palmer S, Ratcliffe J, Read N, *et al.*; North of England IBS Research Group. North of England IBS Research Group. The cost-effectiveness of psychotherapy and paroxetine for severe irritable bowel syndrome. *Gastroenterology* 2003;**124**:303–17.
109. Kayiran S, Dursun E, Dursun N, Ermutlu N, Karamursel S. Neurofeedback intervention in fibromyalgia syndrome; a randomized, controlled, rater blind clinical trial. *Appl Psychophysiol Biofeedback* 2010;**35**:293–302.
110. Sencan S, Ak S, Karan A, Muslumanoglu L, Ozcan E, Berker E. A study to compare the therapeutic efficacy of aerobic exercise and paroxetine in fibromyalgia syndrome. *J Back Musculoskeletal Rehabil* 2004;**17**:57–61.
111. Tanum L, Malt UF. A new pharmacologic treatment of functional gastrointestinal disorder. A double-blind placebo controlled study with mianserin. *Scand J Gastroenterol* 1996;**31**:318–25.
112. NCT. *Comparison of Psychological and Pharmacological Treatments for Pain Due to Temporomandibular Joint Disorder (TMD)*; 2003. URL: <https://clinicaltrials.gov/show/NCT00066937> (accessed 24 June 2024).
113. Zitman FG, Linssen AC, Edelbroek PM, Stijnen T. Low dose amitriptyline in chronic pain: the gain is modest. *Pain* 1990;**42**:35–42.
114. Engel CC Jr, Walker EA, Engel AL, Bullis J, Armstrong A. A randomized, double-blind crossover trial of sertraline in women with chronic pelvic pain. *J Psychosom Res* 1998;**44**:203–7.
115. Kalso E, Tasmuth T, Neuvonen PJ. Amitriptyline effectively relieves neuropathic pain following treatment of breast cancer. *Pain* 1996;**64**:293–302.
116. 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* 2006;**25**:495–7.
117. Sarzi Puttini P, Cazzola M, Boccassini L, Ciniselli G, Santandrea S, Caruso I, Benvenuti C. A comparison of dothiepin versus placebo in the treatment of pain in rheumatoid arthritis and the association of pain with depression. *J Int Med Res* 1988;**16**:331–7.

118. Tasmuth T, Hartel B, Kalso E. Venlafaxine in neuropathic pain following treatment of breast cancer. *Eur J Pain* 2002;**6**:17–24.
119. Ward NG. Tricyclic antidepressants for chronic low-back pain. Mechanisms of action and predictors of response. *Spine* 1986;**11**:661–5.
120. GlaxoSmithKline. *Treatment of Fibromyalgia: A Randomized, Double-blind, Placebo-controlled Study of Paroxetine, a Selective Serotonin re-uptake Inhibitor*. GSK – Clinical Study Register; 2005. URL: [www.gsk-clinicalstudyregister.com](http://www.gsk-clinicalstudyregister.com) (accessed 1 October 2020).
121. NCT. *Effect of Milnacipran in Chronic Neuropathic Low Back Pain*; 2010. URL: <https://clinicaltrials.gov/show/NCT01225068> (accessed 1 October 2020).
122. NCT. *Milnacipran for Chronic Pain in Knee Osteoarthritis*; 2011. URL: <https://clinicaltrials.gov/show/NCT01510457> (accessed 25 June 2024).
123. Bateman L, Palmer RH, Trugman JM, Lin Y. Results of switching to milnacipran in fibromyalgia patients with an inadequate response to duloxetine: a phase IV pilot study. *J Pain Res* 2013;**6**:311–8.
124. Hammody EL, Matloub SY, Shihab SS. Pregabalin versus amitriptyline in the treatment of fibromyalgia patients (a double blind comparative study). *Iraqi Postgrad Med J* 2015;**14**:39–45.
125. Morello CM, Leckband SG, Stoner CP, Moorhouse DF, Sahagian GA. Randomized double-blind study comparing the efficacy of gabapentin with amitriptyline on diabetic peripheral neuropathy pain. *Arch Intern Med* 1999;**159**:1931–7.
126. Razazian N, Baziyar M, Moradian N, Afshari D, Bostani A, Mahmoodi M. Evaluation of the efficacy and safety of pregabalin, venlafaxine, and carbamazepine in patients with painful diabetic peripheral neuropathy. A randomized, double-blind trial. *Neurosciences* 2014;**19**:192–8.
127. Trugman JM, Palmer RH, Ma Y. Milnacipran effects on 24-hour ambulatory blood pressure and heart rate in fibromyalgia patients: a randomized, placebo-controlled, dose-escalation study. *Curr Med Res Opin* 2014;**30**:589–97.
128. Zabihyeganeh M, Amini KA, Vafae AS, Janbozorgi M, Akbari A, Mirzaei A. The effect of cognitive-behavioral therapy versus duloxetine on the laboratory indices of inflammation in fibromyalgia: a randomized controlled trial. *J Ration Emot Cogn Behav Ther* 2021;**1**:1–15.
129. Miki K, Murakami M, Oka H, Onozawa K, Yoshida S, Osada K. Efficacy of mirtazapine for the treatment of fibromyalgia without concomitant depression: a randomized, double-blind, placebo-controlled phase IIa study in Japan. *Pain* 2016;**157**:2089–96.
130. Joint Formulary Committee. *British National Formulary: Duloxetine* (online). London: BMJ Group and Pharmaceutical Press. URL: <https://bnf.nice.org.uk/drugs/duloxetine/> (accessed 29 June 2022).
131. Vrethem M, Boivie J, Arnqvist H, Holmgren H, Lindstrom T, Thorell LH. A comparison of amitriptyline and maprotiline in the treatment of painful polyneuropathy in diabetics and nondiabetics. *Clin J Pain* 1997;**13**:313–23.
132. Piroasca S, Shiely F, Clarke M, Treweek S. Tolerating bad health research: the continuing scandal. *Trials* 2022;**23**:1–8.
133. Chiochia V, Nikolakopoulou A, Higgins J, Page MJ, Papakonstantinou T, Cipriani A, *et al*. ROB-MEN: a tool to assess risk of bias due to missing evidence in network meta-analysis. *BMC Med* 2021;**19**:1–3.
134. Lunn MPT, Hughes RAC, Wiffen PJ. Duloxetine for treating painful neuropathy, chronic pain or fibromyalgia. *Cochrane Database Syst Rev* 2014;**2014**:CD007115. <https://doi.org/10.1002/14651858.CD007115.pub3>

135. Ferreira GE, McLachlan AJ, Lin CW, Zadro JR, Abdel-Shaheed C, O'Keeffe M, *et al.* Efficacy and safety of antidepressants for the treatment of back pain and osteoarthritis: systematic review and meta-analysis. *BMJ* 2021;**372**:1–13.
136. Moore RA, Derry S, McQuay HJ, Straube S, Aldington D, Wiffen P, *et al.*; ACTINPAIN writing group of the IASP Special Interest Group (SIG) on Systematic Reviews in Pain Relief. Clinical effectiveness: an approach to clinical trial design more relevant to clinical practice, acknowledging the importance of individual differences. *Pain* 2010;**149**:173–6. <https://doi.org/10.1016/j.pain.2009.08.007>
137. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *J Pharmacol Pharmacother* 2010;**1**:100–7.
138. Edwards JE, McQuay HJ, Moore RA, Collins SL. Reporting of adverse effects in clinical trials should be improved: lessons from acute postoperative pain. *J Pain Symptom Manage* 1999;**18**:427–37.
139. Phillips R, Hazell L, Sauzet O, Cornelius V. Analysis and reporting of adverse events in randomised controlled trials: a review. *BMJ Open* 2019;**9**:e024537.



# Appendix 1 Search strategies

## MEDLINE

1. pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or breakthrough pain/ or cancer pain/ or exp chest pain/ or chronic pain/ or earache/ or eye pain/ or facial pain/ or flank pain/ or glossalgia/ or exp headache/ or mastodynia/ or metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or neuralgia/ or exp nociceptive pain/ or pain, intractable/ or exp pain, postoperative/ or pain, referred/ or exp pelvic pain/ or renal colic/
2. pain.tw.
3. (headache\* or migraine\* or fibromyalgia\* or neuralgia\*).tw.
4. Fibromyalgia/
5. 1 or 2 or 3 or 4
6. exp ANTIDEPRESSIVE AGENTS/
7. exp MONOAMINE OXIDASE INHIBITORS/
8. exp NEUROTRANSMITTER UPTAKE INHIBITORS/
9. ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or reuptake or re uptake)).tw.
10. (noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic).tw.
11. (antidepress\* or anti-depress\*).tw.
12. (MAOI\* or RIMA).tw.
13. monoamine oxidase inhibit\*.tw.
14. (Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or Befloxatone or Benactyzine or Brofaromin\*).tw.
15. (Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or Cimoxatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine).tw.
16. (Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etooperidone or Femoxetin\* or Fluotracen or Fluoxetine or Fluvoxamin\*).tw.
17. (Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Isocarboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitracen or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide).tw.
18. (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranylcypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone).tw.
19. or/6-18
20. randomized controlled trial.pt.
21. controlled clinical trial.pt.
22. randomized.ab.
23. placebo.ab.
24. drug therapy.fs.
25. randomly.ab.
26. trial.ab.

27. or/20-26
28. exp animals/ not humans.sh.
29. 27 not 28
30. 5 and 19 and 29
31. limit 30 to " all adult (19 plus years)"

## Cochrane Central Register of Controlled Trials (CENTRAL)

- #1 MeSH descriptor: [Antidepressive Agents] explode all trees
- #2 MeSH descriptor: [Monoamine Oxidase Inhibitors] explode all trees
- #3 MeSH descriptor: [Neurotransmitter Uptake Inhibitors] explode all trees
- #4 (((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or reuptake or re uptake)):ti,ab,kw (Word variations have been searched)
- #5 ((noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic)):ti,ab,kw (Word variations have been searched)
- #6 (antidepress\* or anti-depress\*):ti,ab,kw (Word variations have been searched)
- #7 (MAOI\* or RIMA):ti,ab,kw (Word variations have been searched)
- #8 (monoamine oxidase inhibit\*):ti,ab,kw (Word variations have been searched)
- #9 ((Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or Bifloxadone or Benactyzine or Brofaromin\*)):ti,ab,kw (Word variations have been searched)
- #10 ((Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or Cimoxatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine)):ti,ab,kw (Word variations have been searched)
- #11 ((Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etoperidone or Femoxetin\* or Fluotracen or Fluoxetine or Fluvoxamin\*)):ti,ab,kw (Word variations have been searched)
- #12 ((Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Isocarboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitracen or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide)):ti,ab,kw (Word variations have been searched)
- #13 ((Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranlycypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone)):ti,ab,kw (Word variations have been searched)
- #14 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13
- #15 ((headache\* or migraine\* or fibromyalgia\* or neuralgia\*)):ti,ab,kw (Word variations have been searched)
- #16 (pain):ti,ab,kw (Word variations have been searched)
- #17 MeSH descriptor: [Fibromyalgia] this term only
- #18 MeSH descriptor: [Abdominal Pain] explode all trees
- #19 MeSH descriptor: [Arthralgia] explode all trees
- #20 MeSH descriptor: [Back Pain] explode all trees
- #21 MeSH descriptor: [Back Pain] this term only
- #22 MeSH descriptor: [Cancer Pain] this term only
- #23 MeSH descriptor: [Chest Pain] explode all trees

- #24 MeSH descriptor: [Chronic Pain] this term only
- #25 MeSH descriptor: [Earache] this term only
- #26 MeSH descriptor: [Eye Pain] this term only
- #27 MeSH descriptor: [Facial Pain] this term only
- #28 MeSH descriptor: [Flank Pain] this term only
- #29 MeSH descriptor: [Glossalgia] this term only
- #30 MeSH descriptor: [Headache] explode all trees
- #31 MeSH descriptor: [Mastodynia] this term only
- #32 MeSH descriptor: [Metatarsalgia] this term only
- #33 MeSH descriptor: [Musculoskeletal Pain] explode all trees
- #34 MeSH descriptor: [undefined] explode all trees
- #35 MeSH descriptor: [Neuralgia] this term only
- #36 MeSH descriptor: [Nociceptive Pain] explode all trees
- #37 MeSH descriptor: [Pain, Intractable] this term only
- #38 MeSH descriptor: [Pain, Postoperative] explode all trees
- #39 MeSH descriptor: [Pain, Referred] this term only
- #40 MeSH descriptor: [Pelvic Pain] explode all trees
- #41 MeSH descriptor: [Renal Colic] this term only
- #42 #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28  
or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41
- #43 #14 and #42

## EMBASE

1. \*pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or \*breakthrough pain/  
or \*cancer pain/ or exp chest pain/ or \*chronic pain/ or \*earache/ or \*eye pain/ or \*facial  
pain/ or \*flank pain/ or \*glossalgia/ or exp headache/ or \*mastodynia/ or \*metatarsalgia/ or  
exp musculoskeletal pain/ or exp neck pain/ or \*neuralgia/ or exp nociceptive pain/ or \*pain,  
intractable/ or exp pain, postoperative/ or pain, referred/ or exp pelvic pain/ or \*renal colic/
2. pain.tw.
3. (headache\* or migraine\* or fibromyalgia\* or neuralgia\*).tw.
4. Fibromyalgia/
5. 1 or 2 or 3 or 4
6. exp ANTIDEPRESSIVE AGENTS/
7. exp MONOAMINE OXIDASE INHIBITORS/
8. exp NEUROTRANSMITTER UPTAKE INHIBITORS/
9. ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or  
reuptake or re uptake)).tw.
10. (noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or  
TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic).tw.
11. (antidepress\* or anti-depress\*).tw.
12. (MAOI\* or RIMA).tw.
13. monoamine oxidase inhibit\*.tw.
14. (Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or  
Befloxatone or Benactyzine or Brofaromin\*).tw.
15. (Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or  
Cimoxatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine).tw.
16. (Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or  
Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or  
Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram  
or Etoperidone or Femoxetin\* or Fluotracen or Fluoxetine or Fluvoxamin\*).tw.

17. (Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Isocarboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitracen or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide).tw.
18. (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranylcypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone).tw.
19. or/6-18
20. random\$.tw.
21. factorial\$.tw.
22. crossover\$.tw.
23. cross over\$.tw.
24. cross-over\$.tw.
25. placebo\$.tw.
26. (doubl\$ adj blind\$).tw.
27. (singl\$ adj blind\$).tw.
28. assign\$.tw.
29. allocat\$.tw.
30. volunteer\$.tw.
31. Crossover Procedure/
32. double-blind procedure.tw.
33. Randomized Controlled Trial/
34. Single Blind Procedure/
35. 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34
36. (animal/ or nonhuman/) not human/
37. 35 not 36
38. 5 and 19 and 37
39. limit 38 to (adult < 18 to 64 years > or aged < 65+ years >)

## AMED

1. \*pain/ or exp abdominal pain/ or exp arthralgia/ or exp back pain/ or \*breakthrough pain/ or \*cancer pain/ or exp chest pain/ or \*chronic pain/ or \*earache/ or \*eye pain/ or \*facial pain/ or \*flank pain/ or \*glossalgia/ or exp headache/ or \*mastodynia/ or \*metatarsalgia/ or exp musculoskeletal pain/ or exp neck pain/ or \*neuralgia/ or exp nociceptive pain/ or \*pain, intractable/ or exp pain, postoperative/ or pain, referred/ or exp pelvic pain/ or \*renal colic/
2. pain.tw.
3. (headache\* or migraine\* or fibromyalgia\* or neuralgia\*).tw.
4. Fibromyalgia/
5. 1 or 2 or 3 or 4
6. exp ANTIDEPRESSIVE AGENTS/
7. exp MONOAMINE OXIDASE INHIBITORS/
8. exp NEUROTRANSMITTER UPTAKE INHIBITORS/
9. ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or reuptake or re uptake)).tw.
10. (noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic).tw.
11. (antidepress\* or anti-depress\*).tw.

12. (MAOI\* or RIMA).tw.
13. monoamine oxidase inhibit\*.tw.
14. (Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or Befloxatone or Benactyzine or Brofaromin\*).tw.
15. (Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or Cimoxatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine).tw.
16. (Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etoferidone or Femoxetin\* or Fluotraceren or Fluoxetine or Fluvoxamin\*).tw.
17. (Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Isocarboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitraceren or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide).tw.
18. (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranylcypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone).tw.
19. or/6-18
20. (random\* or factorial\* or placebo\* or assign\* or allocat\* or crossover).tw.
21. (cross adj over\*).tw.
22. (trial\* and (control\* or comparative)).tw.
23. ((blind\* or mask\*) and (single or double or triple or treble)).tw.
24. (treatment adj arm\*).tw.
25. (control\* adj group\*).tw.
26. (phase adj (III or three)).tw.
27. (versus or vs).tw.
28. rct.tw.
29. RANDOM ALLOCATION/
30. DOUBLE BLIND METHOD/
31. placebos/
32. randomized controlled trials/
33. 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34. 5 and 19 and 33
35. exp adult/
36. 34 and 35

## PsycINFO

S29 S20 AND S28

S28 S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27

S27 (singl\* OR doubl\* OR trebl\* OR tripl\*) N3 (blind\* OR mask\*)

S26 clinical N3 trial\* OR research N3 design OR evaluat\* N3 stud\* OR prospectiv\* N3 stud\*

S25 placebo\* OR random\* OR "comparative stud\*"

S24 DE "Followup Studies"

S23 DE "Placebo"

S22 DE "Treatment Outcomes" OR DE "Psychotherapeutic Outcomes" OR DE "Side Effects (Treatment)" OR DE "Treatment Compliance" OR DE "Treatment Duration" OR DE "Treatment Refusal" OR DE "Treatment Termination" OR DE "Treatment Withholding"

S21 DE "Treatment Effectiveness Evaluation"

S20 S15 AND S19

S19 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S16 OR S17 OR S18

S18 DE "Neurotransmitter Uptake Inhibitors" OR DE "Atomoxetine" OR DE "Serotonin Norepinephrine Reuptake Inhibitors" OR DE "Serotonin Reuptake Inhibitors"

S17 DE "Monoamine Oxidase Inhibitors" OR DE "Iproniazid" OR DE "Isocarboxazid" OR DE "Moclobemide" OR DE "Nialamide" OR DE "Pargyline" OR DE "Phenelzine" OR DE "Pheniprazine" OR DE "Tranlycypromine"

S16 DE "Antidepressant Drugs" OR DE "Bupropion" OR DE "Citalopram" OR DE "Fluoxetine" OR DE "Fluvoxamine" OR DE "Iproniazid" OR DE "Isocarboxazid" OR DE "Lithium Carbonate" OR DE "Methylphenidate" OR DE "Mianserin" OR DE "Moclobemide" OR DE "Molindone" OR DE "Nefazodone" OR DE "Nialamide" OR DE "Nomifensine" OR DE "Paroxetine" OR DE "Phenelzine" OR DE "Pheniprazine" OR DE "Pipradrol" OR DE "Serotonin Norepinephrine Reuptake Inhibitors" OR DE "Sertraline" OR DE "Sulpiride" OR DE "Tranlycypromine" OR DE "Trazodone" OR DE "Tricyclic Antidepressant Drugs" OR DE "Venlafaxine" OR DE "Zimeldine"

S15 S12 OR S13 OR S14

S14 DE "Fibromyalgia"

S13 pain OR (headache\* or migraine\* or fibromyalgia\* or neuralgia\*)

S12 DE "Pain" OR DE "Aphagia" OR DE "Back Pain" OR DE "Chronic Pain" OR DE "Headache" OR DE "Myofascial Pain" OR DE "Neuralgia" OR DE "Neuropathic Pain" OR DE "Somatoform Pain Disorder"

S11 PAIN

S10 (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Pipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranlycypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone)

S9 (Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Isocarboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitracen or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide)

S8 (Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etoperidone or Femoxetin\* or Fluotracen or Fluoxetine or Fluvoxamin\*)

S7 (Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or Cimoxatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine)

S6 (Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or Befloxadone or Benactyzine or Brofaromin\*)

S5 monoamine oxidase inhibit\*

S4 MAOI\* or RIMA

S3 antidepress\* or anti-depress\*

S2 (noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic)

S1 ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or reuptake or re uptake))

## CINAHL

S31 S4 AND S18 AND S30

S30 S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29

S29 TX allocat\* random\*

S28 (MH "Quantitative Studies")  
 S27 (MH "Placebos")  
 S26 TX placebo\*  
 S25 TX random\* allocat\*  
 S24 (MH "Random Assignment")  
 S23 TX randomi\* control\* trial\*  
 S22 TX ((singl\* n1 blind\*) or (singl\* n1 mask\*)) or TX ((doubl\* n1 blind\*) or (doubl\* n1 mask\*)) or TX ((tripl\* n1 blind\*) or (tripl\* n1 mask\*)) or TX ((trebl\* n1 blind\*) or (trebl\* n1 mask\*))  
 S21 TX clinic\* n1 trial\*  
 S20 PT Clinical trial  
 S19 (MH "Clinical Trials+")  
 S18 S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17  
 S17 (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\* or Norfenfluramin\* or Nortriptylin\* or Noxiptilin\* or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\* or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\* or Toloxatone or Tranylcypromin\* or Trazodone or Trimipramin\* or Tryptophan\* or Venlafaxine or Viloxazine or Vilazodone or Viqualine or Zalospirone)  
 S16 (Hyperforin or Hypericum or St John\* or Imipramin\* or Iprindole or Iproniazid\* or Ipsapirone or Iso-carboxazid\* or Levomilnacipran or Lofepramin\* or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\* or Medifoxamin\* or Melitracen or Metapramin\* or Mianserin or Milnacipran or Minaprin\* or Mirtazapin\* or Moclobemide)  
 S15 (Clorgyline or Clovoxamin\* or "CX157" or Tyrima or Tririma or Demexiptilin\* or Deprenyl or Desipramin\* or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\* or Dimetacrin\* or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etoperidone or Femoxetin\* or Fluotracen or Fluoxetine or Fluvoxamin\*)  
 S14 (Bupropion or Amfebutamone or Butriptylin\* or Caroxazone or Cianopramin\* or Cilobamin\* or Cimoxtatone or Citalopram or Chlorimipramin\* or Clomipramin\* or Chlomipramin\* or Clomipramine)  
 S13 (Agomelatine or Amoxapine or Amineptine or Amitriptylin\* or Amitriptylinoxide or Atomoxetine or Befloxatone or Benactyzine or Brofaromin\*)  
 S12 monoamine oxidase inhibit\*  
 S11 MAOI\* or RIMA  
 S10 antidepress\* or anti-depress\*  
 S9 (noradrenerg\* or antiadrenergic or anti adrenergic or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or heterocyclic or pharmacotherap\* or psychotropic)  
 S8 ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) and (uptake or reuptake or re uptake))  
 S7 (MH "Neurotransmitter Uptake Inhibitors+")  
 S6 (MH "Monoamine Oxidase Inhibitors+")  
 S5 (MH "Antidepressive Agents+")  
 S4 S1 OR S2 OR S3  
 S3 (MH "Fibromyalgia")  
 S2 pain OR (headache\* or migraine\* or fibromyalgia\* or neuralgia\*)  
 S1 (MH "Pain+")

## LILACS

headache\$ or migraine\$ or fibromyalgia\$ or neuralgia\$ or pain [Words] and (Nefazodone or Nialamide or Nitroxazepine or Nomifensin\$ or Norfenfluramin\$ or Nortriptylin\$ or Noxiptilin\$ or Opipramol or Paroxetine or Phenelzine or Pheniprazine or Pipofezine or Pirlindole or Pivagabine or Pizotyline or Propizepine or Protriptylin\$ or Quinupramine or Reboxetine or Rolipram or Scopolamine or Selegiline or Sertraline or Setiptiline or Teciptiline or Thozalinone or Tianeptin\$ or Toloxatone or Tranylcypromin\$

or Trazodone or Trimipramin\$ or Tryptophan\$ or Venlafaxine or Viloxazine or Vilazodone or Viquiline or Zalospirone) or (Hyperforin or Hypericum or St John\$ or Imipramin\$ or Iprindole or Iproniazid\$ or Ipsapirone or Isocarboxazid\$ or Levomilnacipran or Lofepramin\$ or "Lu AA21004" or Vortioxetine or "Lu AA24530" or Tedatioxetine or "LY2216684" or Edivoxetine or Maprotilin\$ or Medifoxamin\$ or Melitracen or Metapramin\$ or Mianserin or Milnacipran or Minaprin\$ or Mirtazapin\$ or Moclobemide) or (Clorgyline or Clovoxamin\$ or "CX157" or Tyrima or Tririma or Demexiptilin\$ or Deprenyl or Desipramin\$ or Pertofrane or Desvenlafaxine or Dibenzepin or Diclofensin\$ or Dimetacrin\$ or Dosulepin or Dothiepin or Doxepin or Duloxetine or Desvenlafaxine or "DVS-233" or Escitalopram or Etoferidone or Femoxetin\$ or Fluotracen or Fluoxetine or Fluvoxamin\$) or (Bupropion or Amfebutamone or Butriptylin\$ or Caroxazone or Cianopramin\$ or Cilobamin\$ or Cimoxatone or Citalopram or Chlorimipramin\$ or Clomipramin\$ or Chlomipramin\$ or Clomipramine) or (Agomelatine or Amoxapine or Amineptine or Amitriptylin\$ or Amitriptylinoxide or Atomoxetine or Befloxatone or Benactyzine or Brofaromin\$) or ((serotonin or norepinephrine or noradrenaline or neurotransmitter\$ or dopamin\$) and (uptake or reuptake or re uptake)) or (noradrenerg\$ or antiadrenergic or anti adrenergic or SSRI\$ or SNRI\$ or NARI\$ or SARI\$ or NDRI\$ or TCA\$ or tricyclic\$ or tetracyclic\$ or heterocyclic or pharmacotherap\$ or psychotropic) or (antidepress\$ or anti-depress\$ or MAOI\$ or RIMA or monoamine oxidase inhibit\$) [Words] and randomised OR randomized OR randomisation OR randomization OR trial OR placebo OR blind OR "phase 3" OR "phase III" [Words]

# Appendix 2 Network meta-analysis reporting decisions

## Overview

This appendix details the decisions made in the reporting of the NMAs in the results section of the review. For each network, we took into account heterogeneity, inconsistency, and network geometry.

## Substantial pain relief (50% reduction)

### *Networks: which model is the best fit?*

Our primary analysis was a Bayesian NMA including treatment. This analysis had high heterogeneity ( $\text{Tau}^2 = 0.26$ ) and inconsistency in both UME and node-splitting models. We also explored networks that separated treatments into different doses, conditions and risk of bias categories and aggregated treatment by class. These networks resulted in models that had similar heterogeneity and variable indications for inconsistency, but the model that included antidepressant dose reduced the estimate of heterogeneity by half ( $\text{Tau}^2 = 0.11$ ) and there was no indication of inconsistency. Therefore, the results are based on the treatment-dose model.

## Pain intensity

### *Change scores and post-intervention scores*

Studies in the review reported pain intensity results in two ways: change scores and post-intervention scores. A total of 50 studies with 14,926 participants reported change scores, while 74 studies with 7703 participants reported post-intervention scores. As these two types of scores cannot be combined directly, we selected model-data combinations on the basis of parsimony, minimisation of inconsistency (identified via UME and node-splitting models), residual deviance and heterogeneity (measured as  $\text{Tau}^2$ ) to minimise the risk of overfitting.

### *Networks: which model is the best fit?*

For both change score and post-intervention analyses, we generated networks and models based on treatment and treatment dose.

### Change scores

The treatment analysis had low heterogeneity ( $\text{Tau}^2 = 0.17$ ) and low inconsistency in the UME model; however, node-splitting models could not be run due to inappropriate network geometry. Models including dose had lower heterogeneity ( $\text{Tau}^2 = 0.10$ ) and no indications for inconsistency in both UME and node-splitting models.

### Post-intervention scores

The treatment analysis had high heterogeneity ( $\text{Tau}^2 = 2.06$ ) compared to change score analysis and inconsistency in the UME model, suggesting it is not possible to fit a robust model to the data. Models including dose continued to have higher heterogeneity than the change score analysis ( $\text{Tau}^2 = 0.46$ ), and high residual deviance across multiple studies, suggesting that a robust model is unlikely to fit the data. UME models continued to show inconsistency between direct and indirect evidence, although node-splitting models showed no inconsistency within studies.

## Mood

### *Change scores and post-intervention scores*

Studies in the review reported pain intensity results in two ways: change scores and post-intervention scores. A total of 38 studies with 12,985 participants reported change scores, while 46 studies with 3885 participants reported post-intervention scores. As these two types of scores cannot be combined, we reported the most appropriate and robust model for the data.

### *Networks: which model is the best fit?*

For both change score and post-intervention analyses, the primary analysis was a Bayesian NMA including treatment.

### **Change scores**

The treatment analysis had low heterogeneity ( $\text{Tau}^2 = 0.09$ ), with no inconsistency in the UME model. We were unable to run node-splitting models due to the network geometry as the majority of the network is formed from two-arm placebo-controlled studies. As the treatment-only analysis had low heterogeneity and no inconsistency, no further analyses were undertaken.

### **Post-intervention scores**

This analysis had moderate heterogeneity ( $\text{Tau}^2 = 0.69$ ), with high residual deviance across multiple studies. UME models showed inconsistency between direct and indirect evidence, although node-splitting models showed no inconsistency within studies. We were unable to run any further analyses including any covariates due to small sample sizes, network geometry and the risk of overfitting.

## Adverse events

### *Networks: which model is the best fit?*

Our primary analysis was a Bayesian NMA including treatment. This analysis had high heterogeneity ( $\text{Tau}^2 = 0.49$ ), with the UME model indicating high inconsistency and divergent transitions within the network. We were unable to run node-splitting models due to network geometry. Models including dose continued to have high heterogeneity ( $\text{Tau}^2 = 0.59$ ), and the UME model showed high inconsistency, similar to the treatment-only model. There continued to be divergent transitions within the network and low effective sample sizes; however, the node-splitting models were able to run and showed no evidence of inconsistency. Due to the network geometry and inappropriateness of running extra models, no further analyses including other covariates were run. The results are based on the treatment-dose model, due to similar levels of heterogeneity and inconsistency, and the ability to run node-splitting models.

## Moderate pain relief

### *Networks: which model is the best fit?*

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.13$ ) and no evidence of inconsistency in both UME and node-splitting models. Therefore, the results are based on a model including treatment only. Divergent transitions suggested unstable models when analysing treatment-dose networks.

## Physical function

### *Change scores and post-intervention scores*

Studies in the review reported physical function results in two ways: change scores and post-intervention scores. A total of 32 studies with 11,760 participants reported change scores, while 30

studies with 3645 participants reported post-intervention scores. As these two types of scores cannot be combined, we reported the most appropriate and robust model for the data.

### **Networks: which model is the best fit?**

For both change score and post-intervention score analyses, the primary analysis was a Bayesian NMA including treatment.

### **Change scores**

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.05$ ), and there was little evidence of inconsistency in the UME model or node-splitting models. Using a model including dose resulted in lower heterogeneity ( $\text{Tau}^2 = 0.04$ ) and no major indications for inconsistency from both UME and node-splitting models.

### **Post-intervention scores**

Our primary analysis was a Bayesian NMA including treatment. This analysis had moderate heterogeneity, higher than that of the change score analysis ( $\text{Tau}^2 = 0.69$ ), with no inconsistency in both UME and node-splitting models. Models including dose increased the heterogeneity ( $\text{Tau}^2 = 0.82$ ) but continued to show no evidence of inconsistency.

## **Sleep**

### **Change scores and post-intervention scores**

Studies in the review reported sleep results in two ways: change scores and post-intervention scores. A total of 18 studies with 6301 participants reported change scores, while 18 studies with 1921 participants reported post-intervention scores. As these two types of scores cannot be combined, we reported the most appropriate and robust model for the data.

### **Networks: which model is the best fit?**

For both change score and post-intervention score analyses, the primary analysis was a Bayesian NMA including treatment.

### **Change scores**

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.06$ ), but due to the star-shaped network geometry we were unable to explore inconsistency using node-splitting models in the treatment-only network. Models including dose also had low heterogeneity ( $\text{Tau}^2 = 0.11$ ) and no indications for inconsistency in UME, but node-splitting models indicated inconsistency, although these parameter estimates may be unreliable due to divergent transitions.

### **Post-intervention scores**

The primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.12$ ) and no inconsistency in both UME and node-splitting models, although there were three divergent transitions. Models including dose had slightly higher heterogeneity ( $\text{Tau}^2 = 0.16$ ), but the network was disconnected, requiring four studies to be removed, and there were 12 divergent transitions.

### **Model used**

Comparing the post-intervention and change score analyses shows that the change score treatment network is more robust and reliable than the post-intervention network as models without divergent transitions were generated. Therefore, the results are based on a model of change scores including both treatment and dose. Results for the treatment-only model are available in the appendices.

## Quality of life

### *Change scores and post-intervention scores*

Studies in the review reported pain intensity results in two ways: change scores and post-intervention scores. A total of 27 studies with 9693 participants reported change scores, while 19 studies with 3103 participants reported post-intervention scores. As these two types of scores cannot be combined, we reported the most appropriate and robust model for the data.

### *Networks: which model is the best fit?*

For both change score and post-intervention analyses, the primary analysis was a Bayesian NMA including treatment.

### **Change scores**

The treatment-only analysis had high heterogeneity ( $\text{Tau}^2 = 0.87$ ), with no evidence of inconsistency in UME and node-splitting models. Models including dose continued to have higher heterogeneity (0.76), with some evidence of inconsistency in the node-splitting models for milnacipran.

### **Post-intervention scores**

The treatment-only analysis had moderate heterogeneity ( $\text{Tau}^2 = 0.55$ ) and no evidence of inconsistency in both UME and node-splitting models, although some residual deviance was present on multiple studies. Models including dose had higher heterogeneity ( $\text{Tau}^2 = 0.67$ ) with similar levels of residual deviance.

### *Model used*

Comparing the post-intervention and change score analyses shows that the post-intervention score treatment network has lower heterogeneity than the change score treatment-dose network. Therefore, the results are based on a model of post-intervention scores including treatment. The results of the change score analyses are included in the appendices.

## Patient Global Impression of Change

### *Patient Global Impression of Change (much/very much improved)*

#### **Networks: which model is the best fit?**

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.12$ ) and no evidence of inconsistency in both UME and node-splitting models. However, there were several divergent transitions. Models including dose reduced the heterogeneity ( $\text{Tau}^2 = 0.08$ ) and continued to show no indications for inconsistency. There was only one divergent transition in this model. Therefore, the results are based on a model including treatment and dose. The results of the treatment-only model are included in the appendices.

### *Patient Global Impression of Change (continuous)*

#### **Networks: which model is the best fit?**

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.05$ ) but some evidence of inconsistency in both UME and node-splitting models. Models including dose continued to have low heterogeneity ( $\text{Tau}^2 = 0.05$ ) and evidence of inconsistency. As the models were very similar, we decided to use the treatment-dose model for clinical utility. The results for the treatment-only model are available in the appendices.

## Serious adverse events

### **Networks: which model is the best fit?**

Our primary analysis was a Bayesian NMA including treatment. This analysis had low heterogeneity ( $\text{Tau}^2 = 0.13$ ) and no inconsistency in both UME and node-splitting models. Including dose into the model did not alter the level of heterogeneity ( $\text{Tau}^2 = 0.16$ ), and continued to have no inconsistency in the UME and node-splitting models. Both treatment-only and treatment-dose models had multiple studies with high residual deviance and imprecision. As both models were very similar, we decided to use the treatment-dose model due to clinical utility. The results for the treatment-only model are included in [Report Supplementary Material 1](#).

## Withdrawal

### **Networks: which model is the best fit?**

Our primary analysis was a Bayesian NMA including treatment. This analysis had high residual deviance and relatively high heterogeneity ( $\text{Tau}^2 = 0.23$ ). We were unable to examine the model using node-splitting models due to the network geometry, as a large proportion of the model was formed of single study connections only. We decided to use this treatment model for the analysis despite the relatively high heterogeneity, as including dose or condition would increase network complexity and dilute already weakly informative edges.



## Appendix 3 Tables of interventions included in network meta-analyses

TABLE 15 Overview of interventions in substantial pain relief analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Desvenlafaxine high dose	2	655
Duloxetine low dose	6	593
Duloxetine standard dose	15	2429
Duloxetine high dose	14	1837
Esreboxetine standard dose	1	553
Esreboxetine high dose	1	280
Milnacipran standard dose	2	644
Milnacipran high dose	1	239
Mirtazapine standard dose	1	211
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline dose unable to be categorised	1	58
Clomipramine standard dose	1	62
Desvenlafaxine standard dose	2	194
Esreboxetine dose unable to be categorised	1	133
Imipramine standard dose	2	113
Mianserin high dose	2	89
Imipramine + pregabalin standard dose	1	69
Venlafaxine standard dose	1	86
Venlafaxine high dose	1	82
Venlafaxine dose unable to be categorised	1	64
Non-antidepressant interventions (excluded from summaries)		
Carbamazepine	1	85
Pregabalin	4	678
Terbutaline	1	39

**TABLE 16** Overview of interventions in pain intensity analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Duloxetine low dose	6	560
Duloxetine standard dose	18	2727
Duloxetine high dose	14	1925
Milnacipran standard dose	4	943
Milnacipran high dose	2	823
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline high dose	1	38
Amitriptyline low dose	1	70
Amitriptyline standard dose	2	130
Amitriptyline dose unable to be categorised	1	24
Citalopram standard dose	2	38
Desipramine standard dose	2	59
Desipramine standard dose + lidocaine	1	30
Desvenlafaxine standard dose	1	49
Desvenlafaxine high dose	1	175
Esreboxetine dose unable to be categorised	1	133
Fluoxetine dose unable to be categorised	1	25
Imipramine low dose	1	18
Milnacipran dose unable to be categorised	2	176
Nortriptyline dose unable to be categorised	1	38
Paroxetine low dose	1	74
Paroxetine dose unable to be categorised	1	58
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	170
Cognitive behavioural therapy	1	15
Gabapentin	1	19
Lidocaine	1	27
Pregabalin	2	550
Psychotherapy	1	74
Usual treatment	1	79

**TABLE 17** Overview of interventions in mood analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Duloxetine	26	4837
Milnacipran	5	1753
Mirtazapine	1	204
Antidepressants with $< 200$ participants (excluded from summaries)		
Citalopram	2	38
Desipramine	1	27
Desipramine + lidocaine	1	32
Esreboxetine	1	126
Fluoxetine	1	25
Imipramine	1	18
Milnacipran + cognitive behavioural therapy	1	17
Nortriptyline	1	38
Paroxetine	1	59
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	166
Cognitive behavioural therapy	1	15
Pregabalin	2	548
Psychotherapy	1	58
Usual treatment	1	63

**TABLE 18** Overview of interventions in the adverse events analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Amitriptyline standard dose	10	518
Desvenlafaxine high dose	2	685
Duloxetine high dose	15	2088
Duloxetine low dose	6	594
Duloxetine standard dose	20	2834
Esreboxetine standard dose	1	556
Milnacipran high dose	7	1573
Milnacipran standard dose	8	1256
Mirtazapine standard dose	1	229
		continued

**TABLE 18** Overview of interventions in the adverse events analysis (*continued*)

Treatment	RCTs	Participants
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline low dose	1	67
Amitriptyline standard dose + melatonin	1	21
Amitriptyline high dose	2	150
Amitriptyline dose unable to be categorised	5	175
Desipramine low dose	1	38
Desipramine low dose + cognitive behavioural therapy	1	37
Desipramine standard dose	1	54
Desvenlafaxine standard dose	2	199
Dothiepin standard dose	1	30
Escitalopram high dose	1	41
Esreboxetine high dose	1	107
Esreboxetine dose unable to be categorised	1	134
Imipramine low dose	2	85
Imipramine standard dose	2	121
Imipramine standard dose + pregabalin	1	69
Imipramine high dose	1	40
Maprotiline low dose	1	33
Milnacipran standard dose + cognitive behavioural therapy	1	20
Milnacipran dose unable to be categorised	2	105
Mirtazapine low dose	1	13
Moclobemide high dose	1	43
Nortriptyline low dose	1	99
Nortriptyline low dose + morphine	1	28
Nortriptyline standard dose	1	28
Nortriptyline dose unable to be categorised	2	61
Nortriptyline dose unable to be categorised + cognitive behavioural therapy	1	41
Nortriptyline dose unable to be categorised + disease management	1	37
Paroxetine unable to be categorised	3	186
Pirlindole low dose	1	45
Sertraline high dose	1	30
Sertraline high dose + coping skills training	1	28
Trazadone low dose + gabapentin	1	94
Venlafaxine low dose	3	123
Venlafaxine standard dose	2	106
Venlafaxine high dose	2	122

**TABLE 18** Overview of interventions in the adverse events analysis (*continued*)

Treatment	RCTs	Participants
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	172
Acetaminophen	1	50
Carbamazepine	2	99
Clonidine	1	20
Cognitive behavioural therapy	4	155
Coping skills training	1	29
Cyclobenzaprine	1	42
Disease management	1	24
Gabapentin	4	175
Lamotrigine	1	46
Lorazepam	1	41
Melatonin	1	21
Morphine	1	28
Naltrexone	1	67
TENS	1	30
Terbutaline	1	51

**TABLE 19** Overview of interventions included in moderate pain relief analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Duloxetine	24	4562
Esreboxetine	2	965
Milnacipran	7	1928
Mirtazapine	2	238
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline	2	80
Desipramine	1	37
Desipramine + cognitive behavioural therapy	1	37
Imipramine	2	113
Imipramine + pregabalin	1	69
Venlafaxine	1	86

continued

**TABLE 19** Overview of interventions included in moderate pain relief analysis (*continued*)

Treatment	RCTs	Participants
Non-antidepressant interventions (excluded from summaries)		
Carbamazepine	2	85
Cognitive behavioural therapy	2	53
Gabapentin	1	22
Pregabalin	4	680
Terbutaline	1	39

**TABLE 20** Overview of all interventions in physical function analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Duloxetine high dose	13	1831
Duloxetine standard dose	14	2157
Milnacipran high dose	2	823
Milnacipran standard dose	3	930
Mirtazapine standard dose	1	204
Antidepressants with < 200 participants (excluded from summaries)		
Citalopram standard dose	2	38
Duloxetine low dose	2	150
Esreboxetine dose unable to be categorised	1	126
Fluoxetine	1	25
Imipramine	1	18
Milnacipran standard + cognitive behavioural therapy	1	17
Nortriptyline dose unable to be categorised	1	38
Paroxetine low dose	1	59
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	166
Cognitive behavioural therapy	1	15
Pregabalin	1	401
Psychotherapy	1	58
Usual treatment	1	63

**TABLE 21** Overview of interventions in sleep analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Duloxetine standard dose	11	1640
Duloxetine high dose	6	891
Milnacipran standard dose	1	398
Milnacipran high dose	1	396
Antidepressants with < 200 participants (excluded from summaries)		
Citalopram standard dose	1	21
Duloxetine low dose	1	141
Esreboxetine unable to be categorised	1	126
Milnacipran unable to be categorised	1	97

**TABLE 22** Overview of all interventions in quality of life analysis

Treatment	RCTs	Participants
Antidepressants with ≥ 200 participants		
Duloxetine	6	306
Esreboxetine	1	736
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline	1	181
Amitriptyline + fluoxetine	1	19
Amitriptyline + melatonin	1	21
Amitriptyline + splint	1	23
Desipramine	1	135
Duloxetine + pregabalin	1	39
Fluoxetine	1	61
Fluoxetine + melatonin	1	50
Imipramine	1	42
Milnacipran	1	53
Nortriptyline	1	36
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	169
Acupuncture	1	28
Cognitive behavioural therapy	1	199
Education	1	66
Melatonin	1	48
Pregabalin	1	63
Saffron	1	23
Terbutaline	1	40
Waitlist	1	21

**TABLE 23** Overview of all interventions in PGIC much/very much improved analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Desvenlafaxine high dose	1	402
Duloxetine high dose	2	287
Duloxetine standard dose	3	481
Esreboxetine high dose	1	275
Esreboxetine standard dose	1	536
Milnacipran high dose	3	1065
Milnacipran standard dose	3	1043
Antidepressants with $< 200$ participants (excluded from summaries)		
Desvenlafaxine standard dose	1	131
Milnacipran dose unable to be categorised	1	79

**TABLE 24** Overview of all interventions in PGIC continuous analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Duloxetine low dose	5	554
Duloxetine standard dose	14	2183
Duloxetine high dose	14	1838
Antidepressants with $< 200$ participants (excluded from summaries)		
Desvenlafaxine high dose	1	184
Desvenlafaxine standard dose	1	54
Non-antidepressant interventions (excluded from summaries)		
ABT-394	1	172
Pregabalin	2	552

**TABLE 25** Overview of all interventions in serious adverse events analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Desvenlafaxine high dose	2	691
Duloxetine high dose	12	1803
Duloxetine low dose	4	473
Duloxetine standard dose	15	2507
Esreboxetine high dose	1	281

**TABLE 25** Overview of all interventions in serious adverse events analysis (*continued*)

Treatment	RCTs	Participants
Esreboxetine standard dose	1	556
Milnacipran high dose	7	1569
Milnacipran standard dose	7	1240
Milnacipran dose unable to be categorised	3	203
Mirtazapine standard dose	3	243
Antidepressants with < 200 participants (excluded from summaries)		
Amitriptyline high dose	1	96
Amitriptyline low dose	1	32
Amitriptyline standard dose	3	114
Amitriptyline dose unable to be categorised	1	25
Bupropion standard dose	1	54
Citalopram standard dose	2	34
Desipramine low dose	1	38
Desipramine + cognitive behavioural therapy	1	37
Desvenlafaxine standard dose	2	199
Esreboxetine dose unable to be categorised	1	134
Imipramine low dose	1	18
Imipramine standard dose	1	51
Milnacipran standard + cognitive behavioural therapy	1	17
Mirtazapine low dose	1	26
Nortriptyline low dose	2	137
Nortriptyline unable to be categorised	1	56
Nortriptyline unable to be categorised + cognitive behavioural therapy	1	41
Nortriptyline unable to be categorised + disease management	1	37
Paroxetine low dose	2	62
Paroxetine dose unable to be categorised	2	152
Reboxetine standard dose	1	18
Sertraline high dose	1	30
Trazadone + gabapentin	1	94
Venlafaxine high dose	1	82
Venlafaxine low dose	1	82
Venlafaxine standard dose	1	86
Non-antidepressant interventions (excluded from summaries)		
Carbamazepine	2	99
Cognitive behavioural therapy	3	72
		continued

**TABLE 25** Overview of all interventions in serious adverse events analysis (*continued*)

Treatment	RCTs	Participants
Coping skills training	1	29
Disease management	1	24
Gabapentin	2	56
Nabilone	1	32
Pregabalin	3	643
Terbutaline	1	51

**TABLE 26** Overview of all interventions in withdrawal analysis

Treatment	RCTs	Participants
Antidepressants with $\geq 200$ participants		
Amitriptyline	34	1326
Desipramine	4	230
Desvenlafaxine	2	885
Duloxetine	45	6082
Esreboxetine	2	978
Imipramine	5	240
Milnacipran	17	3090
Mirtazapine	3	269
Nortriptyline	7	374
Paroxetine	9	389
Venlafaxine	6	409
Antidepressants with $< 200$ participants (excluded from summaries)		
Amitriptyline + fluoxetine	1	31
Amitriptyline + fluphenazine	1	12
Amitriptyline + naproxen	1	19
Amitriptyline + psychotherapy	1	26
Amitriptyline + splint	1	24
Amitriptyline + support	1	26
Bupropion	1	54
Citalopram	4	76
Clomipramine	2	124
Cognitive behavioural therapy and milnacipran	1	20
Cognitive behavioural therapy and amitriptyline	1	12
Coping skills training + sertraline	1	28

TABLE 26 Overview of all interventions in withdrawal analysis (continued)

Treatment	RCTs	Participants
Desipramine + cognitive behavioural therapy	1	37
Desipramine + lidocaine	1	34
Dothiepin	2	55
Doxepin	1	30
Escitalopram	3	86
Fluoxetine	6	140
Fluphenazine	1	13
Gabapentin + nortriptyline	1	56
Maprotiline	3	98
Melatonin + amitriptyline	1	21
Mianserin	2	109
Moclobemide	1	43
Morphine + nortriptyline	1	55
Nortriptyline + cognitive behavioural therapy	1	41
Nortriptyline + disease management	1	37
Nortriptyline + morphine	1	52
Pirlindole	1	45
Pregabalin + duloxetine	1	41
Pregabalin + imipramine	1	73
Reboxetine	1	18
Sertraline	2	66
Trazodone	3	63
Trazodone + gabapentin	2	94
Trimipramine	1	18
Zimeldine	1	17
Non-antidepressant interventions (excluded from summaries)		
ABT-894	1	172
Acetaminophen	1	50
Acupuncture	1	24
Aerobic exercise	1	20
Carbamazepine	2	99
Cognitive behavioural therapy	7	333
Coping skills training	1	29
Cyclobenzaprine	1	42
Disease management	1	24
		continued

**TABLE 26** Overview of all interventions in withdrawal analysis (*continued*)

Treatment	RCTs	Participants
Education	1	71
Gabapentin	6	269
Lamotrigine	1	53
Lidocaine	1	33
Melatonin	1	21
Morphine	2	107
Naltrexone	1	67
Naproxen	1	19
Neurofeedback	1	20
Panax ginseng	1	19
Physical therapy	1	34
Pregabalin	9	919
Psychotherapy	2	116
Saffron/crocin	2	53
Support	1	24
TENS	1	50
Terbutaline	1	51
Usual treatment	2	70
Waitlist	1	24

## Appendix 4 Characteristics of included studies

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Abou-raya 2012	Duloxetine	SNRI	Parallel	Placebo	288	16 weeks	Musculoskeletal	No	No
Agger 2017	Imipramine	TCA	Parallel	Placebo	139	13 weeks	Primary	Partial	Yes
Ahmed 2016	Milnacipran	SNRI	Crossover	Placebo	19	6 weeks	Fibromyalgia	Yes	Yes
Alcoff 1982	Imipramine	TCA	Parallel	Placebo	50	8 weeks	Musculoskeletal	No	Yes
Allen 2014	Desvenlafaxine	SNRI	Parallel	Dose and placebo	412	13 weeks	Neuropathic	Yes	Yes
Allen 2017	Desvenlafaxine	SNRI	Parallel	Dose and placebo	697	15 weeks	Fibromyalgia	Yes	Yes
Anderberg 2000	Citalopram	SSRI	Parallel	Placebo	40	16 weeks	Fibromyalgia	Not reported	Yes
Ang 2013	Milnacipran	SNRI	Parallel	Active and combined	58	21 weeks	Fibromyalgia	No	Yes
Aragona 2005	Citalopram	SSRI	Parallel	AD vs. AD	35	8 weeks	Primary	Not reported	Yes
Arnold 2002	Fluoxetine	SSRI	Parallel	Placebo	60	12 weeks	Fibromyalgia	Yes	Yes
Arnold 2004	Duloxetine	SNRI	Parallel	Placebo	207	12 weeks	Fibromyalgia	Yes	Yes
Arnold 2005	Duloxetine	SNRI	Parallel	Dose and placebo	354	12 weeks	Fibromyalgia	Yes	Yes
Arnold 2010: 1	Esreboxetine	SNRI	Parallel	Placebo	268	8 weeks	Fibromyalgia	Yes	No
Arnold 2010: 2	Milnacipran	SNRI	Parallel	Placebo	1025	12 weeks	Fibromyalgia	Yes	No
Arnold 2010: 3	Duloxetine	SNRI	Parallel	Placebo	530	12 weeks	Fibromyalgia	Yes	No
Arnold 2012: 1	Duloxetine	SNRI	Parallel	Placebo	308	12 weeks	Fibromyalgia	Yes	No
Arnold 2012: 2	Esreboxetine	SNRI	Parallel	Dose and placebo	1122	14 weeks	Fibromyalgia	Yes	Yes
Ash 1999	Dothiepin	TCA	Parallel	Placebo	50	10 weeks	Musculoskeletal	Not reported	Yes
Atkinson 1998	Nortriptyline	TCA	Parallel	Placebo	78	8 weeks	Musculoskeletal	No	Yes
Atkinson 1999	Maprotiline	TeCA	Parallel	AD and active	103	8 weeks	Musculoskeletal	No	Yes
Atkinson 2007	Desipramine	TCA	Parallel	Dose, AD, placebo	121	12 weeks	Musculoskeletal	No	Yes
Bansal 2009	Amitriptyline	TCA	Crossover	Active	51	5 weeks	Neuropathic	No	No
Bateman 2011	Milnacipran	SNRI	Parallel	Placebo	107	10 weeks	Fibromyalgia	Yes	Yes
Bird 2000	Paroxetine	SSRI	Parallel	AD vs. AD	191	8 weeks	Musculoskeletal	Yes	No
Boyle 2012	Amitriptyline	TCA	Parallel	AD and active	83	2 weeks	Neuropathic	Yes	Yes
Branco 2010	Milnacipran	SNRI	Parallel	Placebo	884	16 weeks	Fibromyalgia	Yes	Yes

continued

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Braz 2013	Amitriptyline	TCA	Parallel	AD and active	52	12 weeks	Fibromyalgia	No	Yes
Calderon 2011	Amitriptyline	TCA	Parallel	Placebo, active, combined	47	7 weeks	Musculoskeletal	No	Yes
Cannon 1994	Imipramine	TCA	Parallel	Placebo and active	60	3 weeks	Non-cardiac chest pain	No	No
Cardenas 2002	Amitriptyline	TCA	Parallel	Placebo	84	6 weeks	Neuropathic	No	No
Carette 1986	Amitriptyline	TCA	Parallel	Placebo	70	9 weeks	Fibromyalgia	No	Yes
Carette 1994	Amitriptyline	TCA	Parallel	Placebo and active	208	25 weeks	Fibromyalgia	Yes	Yes
Caruso 1987	Dothiepin	TCA	Parallel	Placebo	60	8 weeks	Fibromyalgia	Not reported	Yes
Chappell 2008	Duloxetine	SNRI	Parallel	Placebo	330	27 weeks	Fibromyalgia	Yes	Yes
Chappell 2009: 1	Duloxetine	SNRI	Parallel	Dose	307	52 weeks	Fibromyalgia	Yes	Yes
Chappell 2009: 2	Duloxetine	SNRI	Parallel	Placebo	231	13 weeks	Musculoskeletal	Yes	No
Chappell 2011	Duloxetine	SNRI	Parallel	Placebo	256	13 weeks	Musculoskeletal	Yes	Yes
Clauw 2008	Milnacipran	SNRI	Parallel	Dose and placebo	1207	15 weeks	Fibromyalgia	Yes	Yes
Creed 2003	Paroxetine	SSRI	Parallel	Active	257	12 weeks	Gastrointestinal	No	Yes
Dezanette 2014	Amitriptyline	TCA	Parallel	Active and combined	63	6 weeks	Fibromyalgia	No	No
Dickens 2000	Paroxetine	SSRI	Parallel	Placebo	98	8 weeks	Musculoskeletal	Yes	No
Drossman 2003	Desipramine	TCA	Parallel	Placebo and active	431	12 weeks	Gastrointestinal	No	Yes
Eberhard 1988	Maprotiline	TeCA	Parallel	AD vs. AD	70	6 weeks	Primary	Not reported	Yes
Engel 1998	Sertraline	SSRI	Crossover	Placebo	25	6 weeks	Musculoskeletal	Not reported	No
Enomoto 2018	Duloxetine	SNRI	Parallel	Active	303	12 weeks	Neuropathic	Yes	No
Enteshari-moghaddam 2019	Duloxetine	SNRI	Parallel	Active	150	12 weeks	Musculoskeletal	No	Yes
Forssell 2004	Venlafaxine	SNRI	Crossover	Placebo	30	4 weeks	Atypical facial pain	No	Yes
Foster 2010: 1	Desipramine	TCA	Parallel	Placebo, active, combined	133	12 weeks	Vulvodynia	No	Yes
Foster 2010: 2	Amitriptyline	TCA	Parallel	Placebo	271	12 weeks	Painful bladder syndrome	Partial	Yes
Frakes 2011	Duloxetine	SNRI	Parallel	Placebo	524	8 weeks	Musculoskeletal	Yes	Yes
Gao 2010	Duloxetine	SNRI	Parallel	Placebo	215	12 weeks	Neuropathic	Yes	Yes
Gao 2015	Duloxetine	SNRI	Parallel	Placebo	405	12 weeks	Neuropathic	Yes	Yes
Gillving 2021	Imipramine	TCA	Crossover	Placebo and active	51	5 weeks	Neuropathic	No	Yes

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Gilron 2009	Nortriptyline	TCA	Crossover	Active and combined	56	6 weeks	Neuropathic	No	Yes
Gilron 2015	Nortriptyline	TCA	Crossover	Active and combined	52	6 weeks	Neuropathic	Yes	Yes
Gilron 2016	Duloxetine	SNRI	Crossover	Placebo, active, combined	41	6 weeks	Fibromyalgia	Yes	Yes
Ginsberg 1996	Amitriptyline	TCA	Parallel	Placebo	51	8 weeks	Fibromyalgia	Not reported	No
Ginsberg 1998	Pirlindole	MAOI reversible	Parallel	Placebo	100	4 weeks	Fibromyalgia	Not reported	Yes
Glaxosmithkline 2005	Paroxetine	SSRI	Parallel	Placebo	52	8 weeks	Fibromyalgia	Yes	Yes
Goldenberg 1986	Amitriptyline	TCA	Parallel	Placebo, active, combined	62	6 weeks	Fibromyalgia	Yes	Yes
Goldenberg 1996	Amitriptyline	TCA	Crossover	Placebo, ad, combined	31	6 weeks	Fibromyalgia	No	Yes
Goldman 2010	Amitriptyline	TCA	Parallel	Placebo	118	6 weeks	Musculoskeletal	No	No
Goldstein 2005	Duloxetine	SNRI	Parallel	Dose and placebo	457	12 weeks	Neuropathic	Yes	Yes
Gonzalez-viejo 2005	Sertraline	SSRI	Parallel	Active	70	24 weeks	Fibromyalgia	Not reported	Yes
Goodkin 1990	Trazodone	SARI	Parallel	Placebo	42	6 weeks	Musculoskeletal	Partial	Yes
Gould 2020	Desipramine	TCA	Parallel	Active and combined	142	12 weeks	Musculoskeletal	No	Yes
Grace 1985	Amitriptyline	TCA	Parallel	Placebo	36	12 weeks	Musculoskeletal	Not reported	Yes
Graff-radford 2000	Amitriptyline	TCA	Parallel	Placebo, active, combined	50	8 weeks	Neuropathic	No	No
Hadianfard 2012	Fluoxetine	SSRI	Parallel	Active	30	8 weeks	Fibromyalgia	No	Yes
Hameroff 1984	Doxepin	TCA	Parallel	Placebo	60	6 weeks	Musculoskeletal	Not reported	Yes
Hammody 2015	Amitriptyline	TCA	Parallel	Active	123	12 weeks	Fibromyalgia	Not reported	Yes
Hannonen 1998	Moclobemide	MAOI reversible	Parallel	AD and placebo	130	12 weeks	Fibromyalgia	Not reported	No
Heymann 2001	Amitriptyline	TCA	Parallel	AD and placebo	118	8 weeks	Fibromyalgia	Not reported	Yes
Holbech 2015	Imipramine	TCA	Crossover	Placebo, active, combined	73	5 weeks	Neuropathic	Yes	Yes
Hudson 2021	Nortriptyline	SNRI	Parallel	Placebo	205	14 weeks	Musculoskeletal	No	No

continued

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Hussain 2011	Fluoxetine	SSRI	Parallel	Active and combined	101	8 weeks	Fibromyalgia	No	No
Iwaki 2020	Duloxetine	SNRI	Parallel	Placebo	47	12 weeks	Neuropathic	Partial	Yes
Johansson 1979	Zimeldine	SSRI	Parallel	Placebo	20	4 weeks	Unspecified	Not reported	Yes
Joharchi 2019	Duloxetine	SNRI	Parallel	Active	180	12 weeks	Neuropathic	No	Yes
Jose 2007	Amitriptyline	TCA	Crossover	Active	75	6 weeks	Neuropathic	Not reported	Yes
Kalso 1996	Amitriptyline	TCA	Crossover	Placebo	20	4 weeks	Neuropathic	No	Yes
Katz 2005	Bupropion	NDRI	Crossover	Placebo	54	7 weeks	Musculoskeletal	Yes	Yes
Kaur 2011	Amitriptyline	TCA	Crossover	AD vs. AD	65	6 weeks	Neuropathic	Not reported	No
Kayiran 2010	Escitalopram	SSRI	Parallel	Active	40	8 weeks	Fibromyalgia	Not reported	Yes
Keefe 2010	Sertraline	SSRI	Parallel	Placebo, active, combined	115	34 weeks	Non-cardiac chest pain	No	Yes
Khoromi 2007	Nortriptyline	TCA	Crossover	Placebo, active, combined	55	9 weeks	Neuropathic	No	Yes
Kim 2013	Milnacipran	SNRI	Crossover	Placebo	20	6 weeks	Fibromyalgia	Yes	No
Lee 2010	Venlafaxine	SNRI	Crossover	Placebo	50	4 weeks	Non-cardiac chest pain	No	No
Lee 2016	Milnacipran	SNRI	Crossover	Placebo	43	6 weeks	Musculoskeletal	Yes	Yes
Leijon 1989	Amitriptyline	TCA	Crossover	Placebo and active	15	4 weeks	Neuropathic	No	No
Loldrup 1989	Clomipramine	TCA	Parallel	AD and placebo	253	6 weeks	Primary	No	Yes
Luo 2009	Fluoxetine	SSRI	Parallel	Placebo	80	8 weeks	Primary	No	No
Maarrawi 2018	Amitriptyline	TCA	Parallel	Placebo	332	8 weeks	Musculoskeletal	No	Yes
Macfarlane 1986	Trimipramine	TCA	Parallel	Placebo	36	12 weeks	Musculoskeletal	Not reported	No
Machado 2018	Amitriptyline	TCA	Parallel	Active and combined	96	16 weeks	Musculoskeletal	No	Yes
Mahmoud 2021	Amitriptyline	TCA	Parallel	Dose	80	16 weeks	Musculoskeletal	No	Yes
Majdinasab 2019	Duloxetine	SNRI	Parallel	Active	104	8 weeks	Neuropathic	No	No
Masand 2009	Paroxetine	SSRI	Parallel	Placebo	72	12 weeks	Gastrointestinal	Yes	Yes
Matthey 2013	Milnacipran	SNRI	Parallel	Placebo	80	7 weeks	Fibromyalgia	Yes	Yes
Max 1988	Amitriptyline	TCA	Crossover	Placebo	58	6 weeks	Neuropathic	Not reported	No
Max 1992	Desipramine	TCA	Crossover	AD and active	54	6 weeks	Neuropathic	Not reported	Yes
Mease 2009	Milnacipran	SNRI	Parallel	Dose and placebo	888	27 weeks	Fibromyalgia	Yes	Yes
Miki 2016	Mirtazapine	NaSSA	Parallel	Placebo	430	12 weeks	Fibromyalgia	Yes	No

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Morello 1999	Amitriptyline	TCA	Crossover	Active	25	6 weeks	Neuropathic	Not reported	Yes
Muller 2008	Escitalopram	SSRI	Parallel	Placebo	51	12 weeks	Primary	Yes	No
Murakami 2015	Duloxetine	SNRI	Parallel	Placebo	393	14 weeks	Fibromyalgia	Yes	No
Nabi 2021	Duloxetine	SNRI	Parallel	Active	72	12 weeks	Neuropathic	No	Yes
Natelson 2015	Milnacipran	SNRI	Parallel	Placebo	34	8 weeks	Fibromyalgia	Yes	Yes
Nct 2011	Milnacipran	SNRI	Parallel	Placebo	46	8 weeks	Musculoskeletal	Yes	Yes
Nct 2010	Milnacipran	SNRI	Parallel	Placebo	40	6 weeks	Neuropathic	Yes	Yes
Nørregaard 1995	Citalopram	SSRI	Parallel	Placebo	43	8 weeks	Fibromyalgia	Yes	No
Otto 2008	Escitalopram	SSRI	Crossover	Placebo	48	5 weeks	Neuropathic	Yes	Yes
Ozerbil 2006	Amitriptyline	TCA	Crossover	AD vs. AD	15	2 weeks	Fibromyalgia	Not reported	No
Pakfetrat 2019	Citalopram	SSRI	Parallel	Active	47	11 weeks	Burning mouth syndrome	no	no
Patkar 2007	Paroxetine	SSRI	Parallel	Placebo	116	12 weeks	Fibromyalgia	Yes	Yes
Petzke 2013	Milnacipran	SNRI	Parallel	Placebo	92	13 weeks	Fibromyalgia	Yes	Yes
Pickering 2018	Milnacipran	SNRI	Parallel	Placebo	54	4 weeks	Fibromyalgia	No	Yes
Rani 1996	Fluoxetine	SSRI	Parallel	AD and placebo	59	4 weeks	Musculoskeletal	Yes	No
Raskin 2005	Duloxetine	SNRI	Parallel	Dose and placebo	348	12 weeks	Neuropathic	Yes	Yes
Razazian 2014	Venlafaxine	SNRI	Parallel	Active	257	4 weeks	Neuropathic	No	Yes
Richards 2015	Venlafaxine	SNRI	Parallel	Placebo	123	12 weeks	Neuropathic	No	No
Rintala 2007	Amitriptyline	TCA	Crossover	Placebo and active	38	8 weeks	Neuropathic	No	Yes
Robinson 2004	Amitriptyline	TCA	Parallel	Placebo	39	6 weeks	Phantom/residual limb pain	No	No
Rowbotham 2004	Venlafaxine	SNRI	Parallel	Dose and placebo	245	6 weeks	Neuropathic	Yes	Yes
Rowbotham 2005	Desipramine	TCA	Parallel	AD vs. AD	47	6 weeks	Neuropathic	No	Yes
Rowbotham 2012	Duloxetine	SNRI	Parallel	Placebo and active	280	8 weeks	Neuropathic	Yes	No
Russell 2008	Duloxetine	SNRI	Parallel	Dose and placebo	520	28 weeks	Fibromyalgia	Yes	Yes
Sarzi Puttini 1988	Dothiepin	TCA	Parallel	Placebo	60	4 weeks	Musculoskeletal	Not reported	Yes
Schukro 2016	Duloxetine	SNRI	Crossover	Placebo	41	4 weeks	Musculoskeletal	No	Yes
Scudds 1989	Amitriptyline	TCA	Crossover	Placebo	39	4 weeks	Fibromyalgia	No	No
Sencan 2004	Paroxetine	SSRI	Parallel	Active	60	6 weeks	Fibromyalgia	Not reported	Yes
Shakiba 2018	Duloxetine	SNRI	Parallel	Active	54	8 weeks	Fibromyalgia	No	No

continued

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Shinichi 2016	Duloxetine	SNRI	Parallel	Placebo	458	14 weeks	Musculoskeletal	Yes	No
Sindrup 2003	Venlafaxine	SNRI	Crossover	AD and placebo	40	4 weeks	Neuropathic	No	Yes
Skljarevski 2009	Duloxetine	SNRI	Parallel	Dose and placebo	404	13 weeks	Musculoskeletal	Yes	Yes
Skljarevski 2010: 1	Duloxetine	SNRI	Parallel	Placebo	236	13 weeks	Musculoskeletal	Yes	Yes
Skljarevski 2010: 2	Duloxetine	SNRI	Parallel	Placebo	401	12 weeks	Musculoskeletal	Yes	No
Smith 2013	Duloxetine	SNRI	Crossover	Placebo	231	5 weeks	Neuropathic	Yes	Yes
Sofat 2016	Duloxetine	SNRI	Parallel	Placebo and active	65	12 weeks	Musculoskeletal	No	Yes
Spinhoven 2010	Paroxetine	SSRI	Parallel	Placebo and active	69	16 weeks	Non-cardiac chest pain	Yes	Yes
Srinivasan 2021	Amitriptyline	TCA	Crossover	Active	67	6 weeks	Neuropathic	No	No
Staud 2015	Milnacipran	SNRI	Parallel	Placebo	61	6 weeks	Fibromyalgia	Yes	Yes
Suttiruksa 2016	Mirtazapine	NaSSA	Parallel	Dose and placebo	40	13 weeks	Fibromyalgia	No	No
Talley 2008	Imipramine	TCA	Parallel	AD and placebo	51	12 weeks	Gastrointestinal	No	Yes
Tammiala-salonen 1999	Trazodone	SARI	Parallel	Placebo	37	8 weeks	Burning mouth syndrome	No	Yes
Tanum 1996	Mianserin	TeCA	Parallel	Placebo	49	7 weeks	Gastrointestinal	No	No
Tasmuth 2002	Venlafaxine	SNRI	Crossover	Placebo	15	4 weeks	Neuropathic	No	No
Tesfaye 2013	Duloxetine	SNRI	Parallel	Active	811	8 weeks	Neuropathic	Yes	Yes
Tetreault 2018	Duloxetine	SNRI	Parallel	Placebo	60	16 weeks	Musculoskeletal	Partial	Yes
Trugman 2014	Milnacipran	SNRI	Parallel	Placebo	321	7 weeks	Fibromyalgia	Yes	Yes
Uchio 2018	Duloxetine	SNRI	Parallel	Placebo	354	14 weeks	Musculoskeletal	Yes	No
Urquhart 2018	Amitriptyline	TCA	Parallel	Placebo	146	24 weeks	Musculoskeletal	No	No
Vahedi 2005	Fluoxetine	SSRI	Parallel	Placebo	44	12 weeks	Gastrointestinal	No	No
Van ophoven 2004	Amitriptyline	TCA	Parallel	Placebo	50	16 weeks	Interstitial cystitis	Not reported	No
Ventafriidda 1987	Amitriptyline	TCA	Parallel	Active	45	2 weeks	Neuropathic	Not reported	Yes
Vitton 2004	Milnacipran	SNRI	Parallel	Dose and placebo	125	12 weeks	Fibromyalgia	Yes	Yes
Vollmer 2014	Duloxetine	SNRI	Parallel	Placebo	239	6 weeks	Neuropathic	Yes	No
Vranken 2011	Duloxetine	SNRI	Parallel	Placebo	48	8 weeks	Neuropathic	No	No
Vrethem 1997: 1	Amitriptyline	TCA	Crossover	AD and placebo	37	4 weeks	Neuropathic	No	No
Vrethem 1997: 2	Amitriptyline	TCA	Crossover	AD and placebo	37	4 weeks	Neuropathic	No	No

Study ID	Antidepressant	Antidepressant class	RCT design	Study design	Total participants randomised	Length of study	Pain condition	Pharmaceutical funding?	High risk of bias?
Wang 2017	Duloxetine	SNRI	Parallel	Placebo	407	13 weeks	Musculoskeletal	Yes	No
Ward 1986	Doxepin	TCA	Parallel	AD vs. AD	35	4 weeks	Musculoskeletal	Not reported	Yes
Ware 2010	Amitriptyline	TCA	Crossover	Active	32	2 weeks	Fibromyalgia	Yes	No
Watson 1992	Amitriptyline	TCA	Crossover	AD vs. AD	35	5 weeks	Neuropathic	No	No
Watson 1998	Amitriptyline	TCA	Crossover	AD vs. AD	33	5 weeks	Neuropathic	Not reported	No
Wernicke 2006	Duloxetine	SNRI	Parallel	Dose and placebo	334	12 weeks	Neuropathic	Yes	Yes
Wolfe 1994	Fluoxetine	SSRI	Parallel	Placebo	42	6 weeks	Fibromyalgia	Yes	Yes
Yasuda 2011	Duloxetine	SNRI	Parallel	Dose and placebo	339	12 weeks	Neuropathic	Yes	No
Yeephu 2013	Mirtazapine	NaSSA	Parallel	Dose and placebo	40	13 weeks	Fibromyalgia	No	No
Yucel 2005	Venlafaxine	SNRI	Parallel	Dose and placebo	60	8 weeks	Neuropathic	Yes	No
Zabihyeganeh 2021	Duloxetine	SNRI	Parallel	Active	128	10 weeks	Fibromyalgia	No	Yes





EME  
HSDR  
**HTA**  
PGfAR  
PHR

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).  
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
Department of Health and Social Care*

***Published by the NIHR Journals Library***