Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

D.1 Systemic pharmacotherapies

Overview of the PICO structure

Definition of the intervention
Systemic pharmacotherapies are medicines that act on the whole body or body systems that involve the entire body, such as the endocrine or/and cardiovascular systems. Systemic pharmacotherapies delivered for short-term and long-term treatment durations were considered. Systemic pharmacotherapies with long- and short-term treatment duration included: • Opioid analgesics and mixed agents: short term < 4 weeks, long term ≥ 4 weeks • Non-steroidal anti-inflammatory drugs (NSAIDs), including cyclo-oxygenase-2 [COX-2] inhibitors: short term < 12 weeks, long term ≥ 12 weeks • Serotonin norepinephrine reuptake inhibitor (SNRI) antidepressants: short term < 12 weeks, long term ≥ 12 weeks • Tricyclic antidepressants (TCAs): short term < 12 weeks, long term ≥ 12 weeks • Anticonvulsants: short term < 12 weeks, long term ≥ 12 weeks • Skeletal muscle relaxants (SMRs): short term < 12 weeks, long term ≥ 12 weeks • Glucocorticoids (systemically administered, i.e. not including epidural steroids): no treatment duration restriction applied • Acetaminophen/Paracetamol: short term < 12 weeks, long term ≥ 12 weeks • Benzodiazepines: short term < 12 weeks, long term ≥ 12 weeks.
PICO question

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Population and subgroups	Community-dwelling adults (aged 20 years and over) experiencing chronic primary low back pain, with or without leg pain, including older people (aged 60 years and older). Subgroups: • Age (all adults and those aged 60 years and over) • Gender and/or sex • Presence of leg pain (radicular, non-radicular, mixed) • Race/ethnicity - studies of populations who were historically marginalized compared with studies of those who were not • Regional economic development - studies carried out in high-income countries compared with studies in low- to middle-income countries
Comparators	a) Placebo/sham b) No drug

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Change in the use of medications • Adverse events (as reported in trials) Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Adverse events (as reported in trials) • Change in the use of medications • Falls Critical outcomes constructs (older adults, aged ≥ 60 years)

Other Evidence-to-Decision (EtD) considerations across all systemic pharmacotherapies

Summary of *values and preferences*	
All adults	**Older people**

No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	**# Review findings GRADE-CERQual Assessment of confidence** 6 Many participants experienced that medication was often the only thing that made a difference to the severity of their pain. However, they were apprehensive of, or dissatisfied with, medication for a number of reasons, often viewing it as a quick fix, temporary relief or that it just masked the pain. Many participants were apprehensive of taking too many medications, the side effects, addiction or did not like how the medications made them feel. Some avoided taking medication all together, did not fill their prescriptions or adjusted medication themselves because of this. MODERATE

Summary of *resource considerations*	
All adults	**Older people**
No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	**# Review findings GRADE-CERQual Assessment of confidence** 8 In rural Nigeria, participants considered medicines as a legitimate form of treatment (cultural norm that disease was treated and 'cured' with medication) and depended on them to be able to perform daily tasks. Other treatments were looked down on or stigmatized, such as exercise. Some participants took medication only to comply with this cultural norm. However, there was a constant struggle to be able to afford the drugs on which they depended to function normally. LOW

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of *equity and human rights considerations*	
All adults	**Older people**
No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	No evidence identified

Summary of *acceptability considerations*	
All adults	**Older people**
No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	**# Review findings GRADE-CERQual Assessment of confidence** 9 Many participants expressed fear of addiction to medication, **especially to opioids**. This led them to not fill prescriptions, to adjust the dosage or stop taking the medication often without consulting their health care provider. MODERATE 10 Some participants in rural Nigeria stated that when the locally produced drugs did not work (they felt that they were substandard or counterfeit), they believed they were fake or substandard. These participants believed that foreign imported drugs were stronger and could lead to a cure. LOW

Summary of *feasibility considerations*	
All adults	**Older people**
No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	No evidence identified

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of judgements by agent

D.1.1 Opioids

Domain	All adults	Older people
Benefits	Small; moderate	Small; moderate
Harms	Small; moderate; large	Small; moderate; large
Balance benefits to harms	Probably favours opioids; probably does not favour opioids; does not favour opioids	Probably favours opioids; probably does not favour opioids; does not favour opioids
Overall certainty	Moderate	Moderate
Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.2 NSAIDs

Benefits	Small; moderate	Small; moderate
Harms	Small; moderate	Small; moderate
Balance benefits to harms	Favours NSAIDs; probably favours NSAIDs	Favours NSAIDs; probably favours NSAIDs
Overall certainty	Moderate	Moderate
Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies

Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.3 SNRI antidepressants

Benefits	Small; trivial	Small; trivial
Harms	Small; moderate	Small; moderate
Balance benefits to harms	Probably favours SNRI antidepressants; probably does not favour SNRI antidepressants	Probably favours SNRI antidepressants; probably does not favour SNRI antidepressants
Overall certainty	Low	Low
Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.4 Tricyclic antidepressants

Benefits	Trivial; uncertain	Trivial; uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Probably does not favour tricyclic antidepressants; does not favour tricyclic antidepressants	Probably does not favour tricyclic antidepressants; does not favour tricyclic antidepressants
Overall certainty	Very low	Very low

Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.5 Anticonvulsants

Benefits	Trivial; uncertain; small	Trivial; uncertain
Harms	Uncertain; moderate	Uncertain; moderate
Balance benefits to harms	Does not favour anticonvulsants	Does not favour anticonvulsants
Overall certainty	Very low	Very low
Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.6 Skeletal muscle relaxants

Benefits	Small; trivial; uncertain	Small; trivial; uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Uncertain	Uncertain
Overall certainty	Low; very low	Low; very low

Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.7 Glucocorticoids

Benefits	Uncertain	Uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Does not favour glucocorticoids; uncertain	Does not favour glucocorticoids; uncertain
Overall certainty	Very low	Very low
Values and preferences	No important uncertainty or variability; varies	No important uncertainty or variability; varies
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	No impact; probably reduced	No impact; probably reduced
Acceptability	Yes; probably no	Yes; probably no
Feasibility	Yes	Yes

D.1.8 Paracetamol (acetaminophen)

ETD process not completed since no trials were available.

D.1.9 Benzodiazepines

ETD process not completed since no trials were available.

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***Opioid analgesics (treatment duration ≥ 1 month) for chronic primary low back pain at 1 to 6 months versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (mean difference on 0 to 10 scale at 1 to 6 months)											
25	RCT	Low	Serious inconsistency (-1)[a]	No indirectness	No imprecision	None noted	4416	3689	NA	MD -0.81 (-1.00 to -0.62)	Moderate
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	NA	MD -0.3 (95% CI NR)	Very low
Pain (proportion with ≥30% or at least moderate improvement at 1 to 6 months)											
18	RCT	Low	Serious inconsistency (-1)[e]	No indirectness	No imprecision	None noted	3474	2964	RR 1.35 (1.22 to 1.52)	ARD 16% (11 to 21)	Moderate
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	32	33	RR 1.16 (0.58 to 2.30)	ARD 7.3% (-16 to 31)	Very low
Function (standardized mean difference at 1 to 6 months)											
16	RCT	Low	Serious inconsistency (-1)[f]	No indirectness	No imprecision	None noted	2874	2592	NA	SMD -0.21 (-0.32 to -0.11)	Moderate
Population subgroup: Presence of radicular leg pain											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	NA	SMD -0.29 (-0.82 to 0.23)	Very low
Function (proportion with ≥30% improvement or Roland Morris Disability Questionnaire (scale 0 to 24) score <14 at 1 to 6 months)											
2	RCT	Moderate (-1)[g]	Consistent	No indirectness	Serious imprecision (-1)[h]	None noted	384	409	RR 1.14 (1.04 to 1.25) and RR 1.13 (0.97 to 1.32)	ARD 10% (3 to 17) and 8.7 (-2.4 to 19.7)	Low
Population subgroup: Presence of radicular leg pain											
No studies											
Quality of life (mean difference on Short-Form-36 or -12 Physical Component Score or Physical Function Subscale [scale 0 to 100])											
7	RCT	Low	No inconsistency	No indirectness	No imprecision	None noted	1014	1065	NA	Mean difference 2.63 (1.62 to 3.86)	High
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	NA	Mean difference 4.7 (-9.4 to 18.8)	Very low
Quality of life (mean difference on Short-Form-36 or -12 Mental Component Score or Mental Health Subscale [scale 0 to 100])											
7	RCT	Low	Serious inconsistency (-1)[i]	No indirectness	No imprecision	None noted	1015	1065	NA	Mean difference -0.11 (-2.02 to 1.96)	Moderate
Population subgroup: Presence of radicular leg pain											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	NA	Mean difference -1.0 (-13.1 to 11.1)	Very low
Psychological well-being (mean difference on Beck Depression Inventory [scale 0 to 63])											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	48	55	NA	Mean change from baseline +13% vs -5.8% (NS)	Very low
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	NA	Mean difference 0.6 (-4.0 to 5.2)	Very low
Serious adverse events (proportion with serious adverse events at 1 to 6 months)											
17	RCT	Low	Consistent	No indirectness	Very serious imprecision (-2)[j]	None noted	3762	3100	RR 1.43 (0.95 to 2.15)	ARD 1% (0 to 1)	Low
Population subgroup: Presence of radicular leg pain											
No studies											
Treatment discontinuation due to adverse events (proportion with treatment discontinuation due to adverse events at 1 to 6 months)											
24	RCT	Low	Serious inconsistency (-1)[k]	No indirectness	No imprecision	None noted	4724	3825	RR 1.52 (1.06 to 2.16)	ARD 4% (1 to 8)	Moderate
Population subgroup: Presence of radicular leg pain											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	41	39	RR 3.80 (0.44 to 32.57)	ARD 7% (-3 to 18)	Very low
Constipation (proportion with constipation at 1 to 6 months)											
22	RCT	Low	Consistent	No indirectness	No imprecision	None noted	4523	3621	RR 2.74 (2.16 to 3.58)	ARD 7% (4 to 10)	High
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	RR 9.00 (2.30 to 35.20)	ARD 57% (37 to 77)	Very low
Headache (proportion with headache at 1 to 6 months)											
20	RCT	Low	Consistent	No indirectness	Serious imprecision (-1)[h]	None noted	4177	3374	RR 1.16 (0.91 to 1.40)	ARD 0% (-1 to 1)	Moderate
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	RR 1.00 (0.28 to 3.61)	ARD 0% (-18 to 18)	Very low
Nausea (proportion with nausea at 1 to 6 months)											
23	RCT	Low	Serious inconsistency (-1)[i]	No indirectness	No imprecision	None noted	4650	3748	RR 2.06 (1.63 to 2.62)	ARD 9% (5 to 12)	Moderate
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)[b]	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	RR 5.00 (0.25 to 99.67)	ARD 7% (-2 to 17)	Very low

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
Vomiting (proportion with vomiting at 1 to 6 months)											
19	RCT	Low	No inconsistency	No indirectness	No imprecision	None noted	3471	2887	RR 2.69 (1.99 to 3.72)	ARD 5% (3 to 7)	High
Population subgroup: Presence of radicular leg pain											
No studies											
Pruritus (proportion with pruritus at 1 to 6 months)											
8	RCT	Low	Serious inconsistency (-1)[m]	No indirectness	No imprecision	None noted	1510	1038	RR 2.63 (1.14 to 6.21)	ARD 7% (-3 to 17)	Moderate
Population subgroup: Presence of radicular leg pain											
No studies											
Somnolence (proportion with somnolence at 1 to 6 months)											
18	RCT	Low	Consistent	No indirectness	No imprecision	None noted	3217	2631	RR 2.36 (1.66 to 3.43)	ARD 5% (2 to 8)	High
Population subgroup: Presence of radicular leg pain											
1	RCT	Moderate (-1)	Unable to assess (-1)[c]	No indirectness	Very serious imprecision (-2)[d]	None noted	28	28	RR 7.00 (0.92 to 53.23)	ARD 21% (4 to 39)	Very low
All outcomes											
Population subgroup: Gender and/or sex											
Two RCTs stated no treatment interaction by sex (data not provided in the trials)											
Population subgroup: Race/ethnicity											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
Two RCTs stated no treatment interaction by race (data not provided in the trials)											
Population subgroup: Regional economic development											
No data. All trials were conducted in very high income settings											
Older adults (aged 60 years and over)											
All outcomes: No RCT restricted enrolment to persons 60 years or older; 3 RCTs reported no interaction by age (one trial reported similar effects on pain intensity in persons ≥65 years and persons <65 years and reported increased likelihood of experiencing ≥30% improvement in pain in both age groups; two trials reported no interaction by age but did not provide data)											

Explanations

a. Downgraded one level for inconsistency because I^2=68%.
b. Downgraded one level for risk of bias because the only trial was rated fair quality.
c. Downgraded one level for inconsistency because there was only 1 trial (unable to assess consistency).
d. Downgraded two levels for imprecision because the number of participants was <100.
e. Downgraded one level for inconsistency because I^2=78%.
f. Downgraded one level for inconsistency because I^2=67%.
g. Downgraded one level for risk of bias because both one trial was rated poor quality and the other trial was rated fair quality.
h. Downgraded one level for imprecision because the confidence interval for the RR estimate included “no effect” and crossed the threshold for a small effect.
i. Downgraded one level for inconsistency because I^2=65%.
j. Downgraded two levels for imprecision because the confidence interval for the RR estimate included “no effect” and crossed the threshold for a large effect.
k. Downgraded one level for inconsistency because I^2=73%.
l. Downgraded one level for inconsistency because I^2=58%.
m. Downgraded one level for inconsistency because I^2=72%.

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***Opioid analgesics (treatment duration <1 month) for chronic primary low back pain at 1 month versus placebo***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Opioids	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month											
1[51,a]	RCT	not serious	serious[b]	not serious	very serious[c]	none	13	12	-	MD **2.74 lower** (4.21 lower to 1.27 lower)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pain intensity at 1-3 months											
No data											
Function, health-related quality of life, psychological well-being, social participation, change in use of medication or adverse events											
No data											
Older adults (aged 60 years and over)											
No data (age range from 20 to 60 years)											

Explanations

a. One parallel randomized trial (Ionescu 2016), conducted in Romania, of adults 20-60 years with chronic low back pain. Tramadol (100 mg/day) for seven days compared to placebo. Pain intensity measured as mean difference on a 1-6 visual analogue scale [data transformed to 0-10] at 7 days.

b. Inconsistency. We downgraded once. This was because there is only one study and inconsistency cannot be assessed.

c. Imprecision. We downgraded twice. This was because there were fewer than 100 participants in the analysis.

References

[51] Ionescu et al. Effects of tramadol treatment on aerobic exercise capacity in subjects with chronic non-specific low back pain. Palestrica of the third millennium – Civilization and Sport; 2015.

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 3. ***Nonsteroidal anti-inflammatory drugs (treatment duration ≥ 12 weeks) for chronic primary low back pain at 3 to 6 months versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (mean improvement on 0 to 10 scale at 3 to 6 months)											
4	RCT	Low	Serious inconsistency[a]	No indirectness	No imprecision	None noted	805	488	NA	Mean difference -0.76 (-1.31 to -0.24)	Moderate
Pain (proportion with ≥30% improvement in pain at 3 to 6 months)											
2	RCT	Low	No inconsistency	No indirectness	Serious imprecision (-1)[b]	None noted	383	271	RR 1.27 (0.87 to 1.71)	ARD 9% (-3 to 18)	Moderate
Function (mean improvement on Roland Morris Disability Questionnaire [0 to 24 scale] at 3 to 6 months)											
4	RCT	Low	Serious inconsistency (-1)[c]	No indirectness	No imprecision	None noted	805	488	NA	Mean difference -1.33 (-2.67 to -0.09)	High
Quality of life (mean improvement on SF-12 Mental Component Summary [0 to 100 scale] at 3 to 6 months)											
2	RCT	Low	No inconsistency	No indirectness	No imprecision[d]	None noted	422	217	NA	Mean difference 0.20 (-1.36 to 1.76)	High
Quality of life (mean improvement on SF-12 Physical Component Summary [0 to 100 scale] at 3 to 6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
2	RCT	Low	No inconsistency	No indirectness	No imprecision[d]	None noted	422	217	NA	Mean difference 2.56 (0.76 to 4.32)	High
Serious adverse events (proportion with serious adverse events at 3 to 6 months)											
3	RCT	Low	No inconsistency	No indirectness	Very serious imprecision (-2)[e]	None noted	598	381	RR 1.13 (0.38 to 6.81)	ARD 1% (-1 to 3)	Low
Discontinuation due to adverse events (proportion with discontinuation due to adverse events at 3 to 6 months)											
4	RCT	Low	No inconsistency	No indirectness	Very serious imprecision (-2)	None noted	808	490	RR 1.10 (0.51 to 2.31)	ARD 1% (-3% to 5)	Low
Nausea (proportion with nausea at 3 to 6 months)											
3	RCT	Low	No inconsistency	No indirectness	Very serious imprecision (-1)	None noted	720	449	RR 1.88 (0.81 to 4.85)	ARD 2% (0 to 4)	Low
Population subgroups, for all outcomes:											
Population subgroup 1: Gender and/or sex											
No data (proportion female ranged from 50% to 62%)											
Population subgroup 2: Race/ethnicity											
No data											
Population subgroup 3: Presence of radicular leg pain											
Patients with radicular pain were excluded from all of the trials											
Population subgroup 4: Regional economic development											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All trials were conducted in the United States											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trials ranged from 52 to 53 years)											

Explanations

a. Downgraded one level for indirectness because I^2=73%.
b. Downgraded one level for imprecision because the confidence interval for the RR estimate included "no effect" and crossed the threshold for a moderate effect (RR ≥1.5).
c. Downgraded one level for inconsistency because I^2=81%.
d. Not downgraded for imprecision; although the confidence interval for the mean difference estimate included "no effect," it did not cross the threshold a small effect (mean difference ≥5 points on a 0 to 100 scale).
e. Downgraded two levels for imprecision because the confidence interval for the RR estimate included "no effect" and crossed the threshold for a large effect (RR ≥2.0).

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 4. ***Nonsteroidal anti-inflammatory drugs (treatment duration < 12 weeks) for chronic primary low back pain at < 1 to 3 months versus placebo***

para							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NSAIDs	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month (mean difference on a 0-10 or 0-100 visual analogue scale at 2-3 weeks)											
5[56-58,a]	RCT	not serious	serious[b] I^2 = 69%	not serious	not serious	We downgraded the evidence by one level because of imputation. [a]	180	117	-	MD **0.77 lower** (1.44 lower to 0.1 lower)	⨁⨁◯◯ Low
Subgroup: gender/sex – not performed (41% female but no stratified analyses)											
Subgroup: radicular pain – not performed (radicular pain excluded or not reported)											
Subgroup: race/ethnicity – not performed (74-99% White but no stratified analyses)											
Subgroup: economic development – not performed (All trials were conducted in high income settings (Australia, USA, Germany, United Kingdom)											
Pain intensity at 1-3 months (mean difference on a 0-10 scale at 4 weeks)											
2[40,58,c]	RCT	not serious	not serious	not serious	not serious	none	173	168	-	MD **0.44 lower** (0.8 lower to 0.07 lower)	⨁⨁⨁⨁ High
Subgroup: gender/sex – not performed (41% female but no stratified analyses)											
Subgroup: radicular pain – not performed (radicular pain excluded or not reported)											
Subgroup: race/ethnicity – not performed (74-99% White but no stratified analyses)											
Subgroup: economic development – not performed (All trials were conducted in high income settings (Australia, USA, Germany, United Kingdom)											
Function at <1 month											
No data											
Function at 1-3 months (mean difference on the 0-24 Roland Morris Disability Questionnaire at 4 weeks)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

para							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NSAIDs	placebo	Relative (95% CI)	Absolute (95% CI)	
1[58,d]	RCT	not serious	serious[e]	not serious	serious[f]	none	64	58	-	MD **1.43 lower** (2.6 lower to 0.26 lower)	⨁⨁◯◯ Low
Subgroup: gender/sex – not performed (41% female but no stratified analyses)											
Subgroup: radicular pain – not performed (radicular pain excluded or not reported)											
Subgroup: race/ethnicity – not performed (74-99% White but no stratified analyses)											
Subgroup: economic development – not performed (All trials were conducted in high income settings (Australia, USA, Germany, United Kingdom)											
Psychological well-being, social participation											
No data											
Change in medication use											
One trial[56] reported no statistically significant difference between groups for the consumption of rescue paracetamol and the other trial[40] significantly lower percentage of patients on flupirtine (70/109, 64.2%) versus placebo (83/110, 75.5%; p = 0.048) used rescue medication.											Unable to evaluate
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Adverse events											
4[40,56,58,g]	RCT	not serious	not serious	not serious	serious[h]	none	79/267 (29.6%)	52/229 (22.7%)	**RR 1.10** (0.83 to 1.46)	**23 more per 1000** (from 39 fewer to 104 more)	⨁⨁⨁◯ Moderate
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Discontinuation due to adverse events											
2[40,58,c]	RCT	not serious	serious[e]	not serious	very serious[i]	none	4/193 (2.1%)	4/194 (2.1%)	**RR 1.01** (0.26 to 3.94)	**0 fewer per 1000** (from 15 fewer to 61 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pruritus											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

para							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NSAIDs	placebo	Relative (95% CI)	Absolute (95% CI)	
1[58,d]	RCT	not serious	serious[e]	not serious	serious[f]	none	0/74 (0.0%)	1/74 (1.4%)	**RR 0.33** (0.01 to 8.05)	**9 fewer per 1000** (from 13 fewer to 95 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Nausea											
2[40,58c]	RCT	not serious	not serious	not serious	very serious[i]	none	5/193 (2.6%)	3/194 (1.5%)	**RR 1.62** (0.17 to 15.79)	**10 more per 1000** (from 13 fewer to 229 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Constipation											
2[40,58,c]	RCT	not serious	not serious	not serious	very serious[i]	none	4/193 (2.1%)	3/194 (1.5%)	**RR 1.26** (0.20 to 7.94)	**4 more per 1000** (from 12 fewer to 107 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dizziness											
2[40,58,c]	RCT	not serious	not serious	not serious	very serious[i]	none	7/193 (3.6%)	5/194 (2.6%)	**RR 1.43** (0.47 to 4.41)	**11 more per 1000** (from 14 fewer to 88 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Somnolence											
1[58,d]	RCT	not serious	serious[e]	not serious	serious[f]	none	1/74 (1.4%)	1/74 (1.4%)	**RR 1.00** (0.06 to 15.69)	**0 fewer per 1000** (from 13 fewer to 199 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dry mouth											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

para							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	NSAIDs	placebo	Relative (95% CI)	Absolute (95% CI)	
2[40,58,d]	RCT	not serious	serious[e]	not serious	very serious[i]	none	0/193 (0.0%)	2/194 (1.0%)	**RR 0.20** (0.01 to 4.16)	**8 fewer per 1000** (from 10 fewer to 33 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Headache											
2[40,58,c]	RCT	not serious	not serious	not serious	very serious[i]	none	2/193 (1.0%)	7/194 (3.6%)	**RR 0.30** (0.06 to 1.47)	**25 fewer per 1000** (from 34 fewer to 17 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Vomiting											
2[40,58,c]	RCT	not serious	serious[e]	not serious	very serious[i]	none	0/193 (0.0%)	1/194 (0.5%)	**RR 0.34** (0.01 to 8.17)	**3 fewer per 1000** (from 5 fewer to 37 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trials ranged from 51 to 59 years)											

Explanations

a. Three trials (Berry 1982, Ghosh 1981, Gurrell 2018), conducted in high-income countries, of adults with chronic low back pain (radicular pain excluded or not reported) with mean ages of 51-55. NSAIDs included naproxen (1100 mg/day), diflunisal (100 mg/day), flurbiprofen (300 mg/day), indomethacin (150 mg/day), and naproxen (1000 mg/day). Pain intensity was measured at 2-3 weeks. The two crossover trials each analysed two NSAIDs; therefore, we split the control sample to avoid over-weighting. The two crossover trials only reported group-level data, which we analysed in the same way as parallel studies. Imputation of the standard deviation was required for the crossover trials, which was taken from the parallel trial. We downgraded the evidence by one level because of this imputation.

b. Inconsistency. We downgraded once. This was because I2 is greater than 50% and there was insufficient data to conduct stratified/sensitivity analyses (I2 = 69%).

c. Two parallel trials (Gurrell 2018, Uberall 2012), conducted in high-income countries, of adults with chronic low back pain (radicular pain excluded) with mean ages of 51-59. NSAIDs naproxen (1000 mg/day) and flupirtine modified release (400 mg/day). Outcome measured at 4 weeks.

d. One parallel trial (Gurrell 2018), conducted in the United States, of adults with chronic low back pain (radicular pain excluded) with mean age of 51. Naproxen (1000 mg/day).

e. Inconsistency. We downgraded once. This was because there is only one study and inconsistency cannot be assessed.

f. Imprecision. We downgraded once. This was because there were fewer than 200 participants in analysis.

g. Three trials (Berry 1982, Gurrell 2018, Uberall 2012), conducted in high-income countries, of adults with chronic low back pain (radicular pain excluded or not reported) with mean ages of 51-59. NSAIDs included naproxen (1100 mg/day), diflunisal (100 mg/day), naproxen (1000 mg/day), and flupirtine modified release (400 mg/day). The crossover trial analysed two NSAIDs; therefore, we split the control sample to avoid over-weighting. The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
h. Imprecision. We downgraded once This was because the pooled estimate crosses the null and the threshold a small effect.
i. Imprecision. We downgraded twice. This was because there are more than 200 participants in the single study, but the estimate crosses the null and the threshold for a large effect.
j. Imprecision. We downgraded twice. This was because the pooled estimate crosses the null and the threshold for a large effect.

References

[56] Berry et al. Naproxen sodium, diflunisal, and placebo in the treatment of chronic back pain. Annals of the Rheumatic Diseases; 1982.
[57] Ghosh et al. A double-blind crossover trial of indomethacin flurbiprofen and placebo in the management of lumbar spondylosis. Current Therapeutic Research, Clinical and Experimental; 1981.
[58] Gurrell et al. A randomized, placebo-controlled clinical trial with the α2/3/5 subunit selective GABAA positive allosteric modulator PF-06372865 in patients with chronic low back pain. PAIN; 2018.
[40] Uberall et al. Efficacy and safety of flupirtine modified release for the management of moderate to severe chronic low back pain: results of SUPREME, a prospective randomized, double-blind, placebo- and active-controlled parallel-group phase IV study. Current Medical Research and Opinion; 2012.

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 5. ***SNRI antidepressants (treatment duration ≥ 12 weeks) for chronic primary low back pain at 3 to 6 months versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (mean difference on 0 to 10 scale at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	808	654	NA	Mean difference -0.54 (-0.76 to -0.34)	Moderate
Pain (proportion with ≥30% improvement in pain intensity at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	812	659	**≥30%:** RR 1.26 (1.13 to 1.39)	ARD 12% (7 to 17)	Moderate
Function (mean difference on Brief Pain Inventory Pain Interference [0 to 10 scale] at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	784	653	NA	Mean difference -0.42 (-0.77 to -0.14) on 0 to 10 scale	Moderate
Quality of life (mean difference in EuroQoL [0 to 1 scale] at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	Serious imprecision (-1)[b]	None noted	830	667	NA	Mean difference ranged from 0 to 0.05 in 3 RCTs (1 RCT reported no difference; data not provided)	Low

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
Psychological well-being (mean differences on SF-36 Mental Health score [0 to 100 scale] at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	Serious imprecision (-1)[b]	None noted	830	667	NA	Mean difference ranged from no difference to 4.88 points in 4 RCTs	Low
Work (mean differences on the Work Productivity and Activity Impairment absenteeism scale at 3 to <6 months)											
3	RCT	Moderate (-1)[c]	No inconsistency	Direct	Serious imprecision (-1)	None noted	543	550	NA	No differences	Low
Serious adverse event (proportion with serious adverse event at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	Very serious imprecision (-2)	None noted	832	667	RR 1.33 (0.55 to 5.86)	ARD 1% (-1 to 3)	Very low
Discontinuation due to adverse events (proportion with discontinuation due to adverse event at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	832	667	RR 2.33 (1.62 to 3.36)	ARD 7% (3 to 12)	Moderate
Nausea (proportion with nausea at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	834	665	RR 4.59 (2.80 to 7.48)	ARD 10% (6 to 15)	Moderate
Constipation (proportion with constipation at 3 to <6 months)											
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	834	665	RR 2.59 (1.22 to 5.89)	ARD 4% (0 to 7)	Moderate
Dizziness (proportion with dizziness at 3 to <6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
4	RCT	Moderate (-1)[a]	No inconsistency	Direct	No imprecision	None noted	834	665	RR 2.28 (1.14 to 5.98)	ARD 3% (0 to 5)	Moderate
Somnolence (proportion with somnolence at 3 to <6 months)											
3	RCT	Moderate (-1)[d]	No inconsistency	Direct	No imprecision	None noted	719	544	RR 2.67 (1.38 to 5.01)	ARD 5% (-2 to 13)	Moderate
Population subgroups, for all outcomes:											
Population subgroup 1: Gender and/or sex											
No data (proportion female in the trials ranged from 11% to 61%)											
Population subgroup 2: Race/ethnicity											
No data											
Population subgroup 3: Presence of radicular leg pain											
No data (all trials excluded patients with radicular leg pain except one trial in which 12% had radicular low back pain and one trial that did not report inclusion of persons with radicular pain											

Population subgroup 4: Regional economic development
All trials were conducted in high income settings

Older adults (aged 60 years and over)
No data, for all outcomes (mean age in the trials ranged from 46 to 59 years)

Explanations:

a. Downgraded 1 level for risk of bias because 3 of 4 trials (encompassing 70% of participants) were rated fair quality.
b. Downgraded 1 level for imprecision because the risk estimates in the trials included "no effect" and crossed the threshold for a small effect.
c. Downgraded 1 level for risk of bias because 2 of 3 trials (encompassing 63% of participants) were rated fair quality.
d. Downgraded 1 level for risk of bias because 2 of 3 trials (encompassing 64% of participants) were rated fair quality.

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 6. ***SNRI antidepressants (treatment duration < 12 weeks) for chronic primary low back pain at < 1 to 3 months versus placebo***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month (mean difference on a 0-10 scale at 3 weeks)											
2[67,70,a]	RCT	very serious[b]	serious[c] I^2 = 65%	not serious	very serious[s]	none	69	73	-	MD **1.1 lower** (2.62 lower to 0.42 higher)	⨁◯◯◯ Very low
Subgroup: gender/sex – not performed (0 to 58% female but no stratified analyses) Subgroup: radicular pain – not performed (some studies included radicular pain but no stratified analyses) Subgroup: race/ethnicity – not performed (White ranged from 85% to 98% but no stratified analyses) Subgroup: economic development – not performed (All trials were conducted in high income settings)											
Pain intensity at 1-3 months (mean difference on a 0-10 scale at 4-8 weeks)											
4[66,67,69,70,e]	RCT	very serious[b]	serious[c] I^2 = 51%	not serious	serious[f]	none	107	124	-	MD **0.23 lower** (1.18 lower to 0.71 higher)	⨁◯◯◯ Very low
Subgroup: gender/sex – not performed (0 to 58% female but no stratified analyses) Subgroup: radicular pain – not performed (some studies included radicular pain but no stratified analyses) Subgroup: race/ethnicity – not performed (White ranged from 85% to 98% but no stratified analyses) Subgroup: economic development – not performed (All trials were conducted in high income settings)											
Function at <1 month											
No data											
Function at 1-3 months (standardized mean difference on the 0-100 Oswestry Disability Index at 8 weeks)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
1[67,h]	RCT	very serious[b]	serious[i]	not serious	very serious[i]	none	41	46	-	SMD **0.15 lower** (0.57 lower to 0.27 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Quality of life at <1 month											
No data											
Quality of life at 1-3 months (standardized mean difference on the Physical Health sub-scale of the Short-Form 36 at 4 weeks)											
1[70,l]	RCT	very serious[b]	serious[i]	not serious	very serious[i]	none	21	21	-	SMD **0.46 higher** (0.16 lower to 1.07 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Psychological well-being at <1 month (mean difference on the 0-60 Montgomery Asberg Depression Rating Scale at 3 weeks)											
1[67,h]	RCT	very serious[b]	serious[i]	not serious	very serious[i]	none	35	37	-	MD **0.5 lower** (3.5 lower to 2.5 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Psychological well-being at 1-3 months (standardized mean difference [questionnaires include 0-60 Montgomery Asberg Depression Rating Scale, Mental Health sub-scale of the Short-Form 36] at 8 weeks)											
2[67,70,a]	RCT	very serious[b]	not serious	not serious	serious[s]	none	65	69	-	SMD **0.08 higher** (0.26 lower to 0.42 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
Social participation											
No data											
Medication use											
One trial[70] reported that rescue medication use did not differ between groups.											Not evaluated
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Adverse events											
4[66,67,69,70,e]	RCT	very serious[b]	serious[c]	not serious	serious[f]	none	83/118 (70.3%)	82/129 (63.6%)	**RR 1.12** (0.85 to 1.48)	**76 more per 1000** (from 95 fewer to 305 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Serious adverse events											
3[67-69,n]	RCT	very serious[b]	not serious	not serious	very serious[d]	none	0/79 (0.0%)	2/82 (2.4%)	**RR 0.34** (0.04 to 3.21)	**16 fewer per 1000** (from 23 fewer to 54 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Discontinuation due to adverse events											
3[66-68,o]	RCT	very serious[b]	not serious	not serious	serious[s]	none	13/93 (14.0%)	3/98 (3.1%)	**RR 4.50** (1.32 to 15.28)	**107 more per 1000** (from 10 more to 437 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Nausea											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
3[67,69,70,p]	RCT	very serious[b]	not serious	not serious	serious[s]	none	20/96 (20.8%)	6/97 (6.2%)	**RR 3.21** (1.33 to 7.73)	**137 more per 1000** (from 20 more to 416 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Constipation											
4[66,67,69,70,e]	RCT	very serious[b]	not serious	not serious	very serious[d]	none	15/118 (12.7%)	10/129 (7.8%)	**RR 1.75** (0.84 to 3.65)	**58 more per 1000** (from 12 fewer to 205 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dizziness											
3[67,69,70,p]	RCT	very serious[b]	not serious	not serious	very serious[d]	none	7/96 (7.3%)	6/97 (6.2%)	**RR 1.17** (0.22 to 6.19)	**11 more per 1000** (from 48 fewer to 321 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Somnolence											
3[66,67,69,q]	RCT	very serious[b]	not serious	not serious	serious[f]	none	15/87 (17.2%)	24/100 (24.0%)	**RR 0.85** (0.55 to 1.31)	**36 fewer per 1000** (from 108 fewer to 74 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dry mouth											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
4[66,67,69,70,e]	RCT	very serious[b]	serious[c]	not serious	very serious[d]	none	25/118 (21.2%)	21/129 (16.3%)	**RR 2.65** (0.45 to 15.76)	**269 more per 1000** (from 90 fewer to 1000 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Headache											
2[67,69,r]	RCT	very serious[b]	not serious	not serious	serious[s]	none	4/65 (6.2%)	15/68 (22.1%)	**RR 0.28** (0.10 to 0.78)	**159 fewer per 1000** (from 199 fewer to 49 fewer)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Vomiting											
1[67,h]	RCT	very serious[b]	serious[i]	not serious	very serious[i]	none	4/45 (8.9%)	0/48 (0.0%)	**RR 9.59** (0.53 to 173.18)	**0 fewer per 1000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pruritus											
2[67,69,r]	RCT	very serious[b]	not serious	not serious	very serious[d]	none	1/65 (1.5%)	1/68 (1.5%)	**RR 1.02** (0.11 to 9.52)	**0 fewer per 1000** (from 13 fewer to 125 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trials ranged from 52 to 59 years)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Explanations

a. One parallel trial (Dickens 2000) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day) and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
b. Risk of bias. We downgraded two levels. This was because more than 50% of participants come from studies with high risk of bias.
c. Inconsistency. We downgraded one level. This was because I2 is greater than 50% and not explained by stratified/sensitivity analyses due to limited data.
d. Imprecision. We downgraded two levels. This was because the pooled estimate crosses the null and the threshold for a large effect.
e. Three parallel trials (Atkinson 1999, Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20-30 mg/day), milnacipran (up to 200 mg/day), and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
f. Imprecision. We downgraded one level. This was because the pooled estimate crosses the null and the threshold for a small effect.
g. Two parallel trials (Atkinson 1998, Atkinson 1999), conducted in the USA, of adults with chronic low back pain with mean ages of 46-49. TCA antidepressants included nortriptyline (up 100 mg/day) and maprotiline (up to 150 mg/day).
h. One parallel trial (Dickens 2000), conducted in the United Kingdom, of adults with chronic low back pain with a mean age of 45. Paroxetine (20 mg/day).
i. Inconsistency. We downgraded one level. This was because there is only one study and inconsistency cannot be assessed.
j. Imprecision. We downgraded two levels. This was because there were fewer than 100 participants in the analysis.
k. Two parallel trials (Atkinson 1998, Pheasant 1983), conducted in the USA, of adults with chronic low back pain with mean ages of 46-47. TCA antidepressants included nortriptyline (up to 100 mg/day) and amitriptyline (up to 150 mg/day).
l. One crossover trial (Schukro 2016), conducted in Austria, of adults with chronic low back pain with a mean age of 58 years. The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
m. One parallel trial (Atkinson 1998), conducted in the USA, of adults with chronic low back pain with mean age of 46 years. TCA antidepressant was nortriptyline (up to 100 mg/day).
n. Two parallel trials (Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20 mg/day), milnacipran (up to 200 mg/day), and duloxetine (60 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
o. Two parallel trials (Atkinson 1999, Dickens 2000) and one crossover trial (Johnson 2011), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20-30 mg/day) and duloxetine (60 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
p. Two parallel trials (Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day), milnacipran (up to 200 mg/day), and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
q. Three parallel trials (Atkinson 1999, Dickens 2000, NCT01226068), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20-30 mg/day) and milnacipran (up to 200 mg/day).
r. Two parallel trials (Dickens 2000, NCT01225068), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day) and milnacipran (up to 200 mg/day).
s. Imprecision. We downgraded one level. This was because there were fewer than 200 participants in the analysis.

References

[67] Dickens et al. The relationship between pain and depression in a trial using paroxetine in sufferers of chronic low back pain. Psychosomatics; 2000.
[70] Schukro et al. Efficacy of duloxetine in chronic low back pain with a neuropathic component: a randomized, double-blind, placebo-controlled crossover trial. Anesthesiology; 2016.
[69] NCT01225068. Effect of milnacipran in chronic neuropathic low back pain. 2012.
[66] Atkinson et al. Effects of noradrenergic and serotonergic antidepressants on chronic low back. PAIN; 1999.
[71] Atkinson et al. A placebo-controlled randomized clinical trial of nortriptyline for chronic low. PAIN; 1998.
[72]Pheasant et al. Amitriptyline and chronic low-back pain. A randomized double-blind crossover study. Spine; 1983.
[68]Johnson et al. Effects of duloxetine and placebo in patients with chronic low back pain. The Journal of Pain; 2011.

GRADE Table 7. ***Tricyclic antidepressants (treatment duration ≥ 12 weeks) for chronic primary low back pain at 3 to 6 months versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (mean difference on 0 to 10 scale at 3 to <6 months)											
3	RCT	Moderate (-1)[a]	No inconsistency	Direct	Serious imprecision (-1)[b]	None noted	161	133	NA	Mean difference -0.58 (-1.89 to 0.72), -0.40 (-0.56 to 1.36), and -0.10 (-0.79 to 5.78)	Low
Pain (mean difference on 0 to 10 scale at 6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[d]	None noted	72	74	NA	Mean difference -0.78 (-1.6 to 0.01)	Low
Pain (proportion with ≥30% or >75% improvement in pain intensity at 3 to <6 months)											
2	RCT	Moderate (-1)[e]	No inconsistency	Direct	Serious imprecision (-1)[f]	None noted	67	55	**≥30%:** RR 1.23 (0.72 to 2.11) **>75%:** RR 1.28 (0.43 to 3.85)	**≥30%:** ARD 10% (-13 to 33) 1.23 (0.72 to 2.11) **>75%:** ARD 5% (-17 to 27)	Low
Function (mean difference on Brief Pain Inventory Pain Interference [0 to 10 scale] or Roland Morris Disability Questionnaire [0 to 24 scale] at 3 to <6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
2	RCT	Moderate (-1)[e]	No inconsistency	Direct	Serious imprecision (-1)[f]	None noted	109	107	NA	Mean difference -0.77 (-1.87 to 0.33) on BPI and -1.62 (-2.88 to -0.36) on RDQ	Low
Function (mean difference on Roland Morris Disability Questionnaire [0 to 24 scale] at 6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	NA	Mean difference -0.98 (-2.42 to 0.46)	Low
Quality of life (mean difference in EuroQoL [0 to 1 scale] at 3 to <6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	NA	Mean difference -0.03 (-0.11 to 0.07)	Low
Quality of life (mean difference in EuroQoL [0 to 1 scale] at 6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	NA	Mean difference -0.05 (-0.004 to 0.10)	Low
Psychological well-being (mean differences on Beck Depression Inventory [0 to 63] at 3 to <6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	NA	Mean difference -0.84 (-2.42 to 0.74)	Low
Psychological well-being (mean difference on Beck Depression Inventory [0 to 63] at 6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	NA	Mean difference -0.93 (-3.34 to 1.49)	Low
Work (proportion with work absence at 3 to <6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	51	50	NA	Adjusted OR 0.86 (0.32 to 2.31)	Low
Work (proportion with work absence at 6 months)											
1	RCT	Low	Unable to assess (-1)[c]	Direct	Very serious imprecision (-2)[h]	None noted	44	43	NA	Adjusted OR 1.51 (0.43 to 5.38)	Very low
Serious adverse event (proportion with serious adverse event at 3 to <6 months)											
1	RCT	Moderate (-1)[i]	Unable to assess (-1)[c]	Direct	Very serious imprecision (-2)[h]	None noted	38	33	RR 2.62 (0.11 to 62.10)	ARD 3% (-5 to 10)	Very low
Moderate to severe adverse events (proportion with any moderate to severe adverse event at 6 months)											
1	RCT	Moderate (-1)[i]	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	26.5% vs 31.8% (p=0.58)	ARD -5% (CI not available)	Very low
Discontinuation due to adverse events (proportion with discontinuation due to adverse event at 3 to <6 months)											
2	RCT	Moderate (-1)[i]	Serious inconsistency (-1)[j]	Direct	Very serious imprecision (-2)[k]	None noted	90	59	RR 3.15 (0.45 to 21.94)	ARD 15% (-12 to 42)	Very low
Discontinuation due to adverse events (proportion with discontinuation due to adverse event at 6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Low	Unable to assess (-1)[c]	Direct	Serious imprecision (-1)[g]	None noted	72	74	RR 1.03 (0.43 to 2.44)	ARD 0% (-10 to 11)	Very low
Nausea (proportion with nausea at 3 to <6 months)											
1	RCT	Moderate (-1)[i]	Unable to assess (-1)[c]	Direct	Very serious imprecision (-2)[h]	None noted	38	33	RR 0.29 (0.01 to 6.90)	ARD -3% (-11 to 5)	Very low
Constipation (proportion with constipation at 3 to <6 months)											
2	RCT	Moderate (-1)[i]	No inconsistency	Direct	Very serious imprecision (-2)[k]	None noted	68	55	RR 7.24 (0.95 to 55.39)	ARD 12% (3 to 20)	Very low
Somnolence (proportion with somnolence at 3 to <6 months)											
1	RCT	Moderate (-1)[i]	Unable to assess (-1)c	Direct	Very serious imprecision (-2)[h]	None noted	38	33	RR 0.87 (0.06 to 13.35)	ARD 0% (-8 to 7)	Very low
Dry mouth (proportion with dry mouth at 3 to <6 months)											
2	RCT	Moderate (-1)[i]	No inconsistency	Direct	No imprecision	None noted	68	55	RR 3.87 (1.20 to 12.49)	ARD 15% (1 to 29)	Moderate
Population subgroups, for all outcomes:											
Population subgroup 1: Gender and/or sex											
No data (proportion female in the trials ranged from 11% to 61%)											
Population subgroup 2: Race/ethnicity											
No data											
Population subgroup 3: Presence of radicular leg pain											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
No data (all trials excluded patients with radicular leg pain except one trial in which 12% had radicular low back pain and one trial that did not report inclusion of persons with radicular pain											

Population subgroup 4: Regional economic development
All trials were conducted in high income settings

Older adults (aged 60 years and over)
No data, for all outcomes (mean age in the trials ranged from 46 to 59 years)

Explanations

a. Downgraded 1 level for risk of bias because two of three trials (encompassing 50% of participants) were rated fair quality.
b. Downgraded 1 level for imprecision because the confidence intervals for the estimates in the individual trials included "no effect" and crossed the threshold for a small (≥0.5 on a 0 to 10 scale) or moderate (≥1 on a 0 to 10 scale) effect.
c. Downgraded 1 level for inconsistency because there was only 1 trial (unable to assess inconsistency).
d. Downgraded 1 level for imprecision because the confidence interval for the estimate included "no effect" and crossed the threshold for a moderate effect.
e. Downgraded 1 level for risk of bias because both trials were rated fair quality.
f. Downgraded 1 level for imprecision because the confidence intervals for the estimates in the individual trials included "no effect" and crossed the threshold for clinically relevant (greater than small) effects.
g. Downgraded 1 level for imprecision because there were <200 participants.
h. Downgraded 2 levels for imprecision because there were <100 participants.
i. Downgraded 1 level for risk of bias because the only trial was rated fair quality.
j. Downgraded 1 level for inconsistency because I^2=88%.
k. Downgraded 2 levels for imprecision because the confidence interval for the estimate included "no effect" and crossed the threshold for a large effect (RR ≥2.0).

GRADE Table 8. ***Tricyclic antidepressants (treatment duration < 12 weeks) for chronic primary low back pain at <1 to 3 months versus placebo***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month											
No data											
Pain intensity at 1-3 months (mean difference on a 0-10 scale at 8 weeks)											
2[66,71,g]	RCT	very serious[b]	not serious	not serious	serious[s]	none	58	72	-	MD **0.69 lower** (1.36 lower to 0.03 lower)	⨁◯◯◯ Very low
Subgroup: gender/sex – not performed (female ranged from 0% to 75% but no stratified analyses)											
Subgroup: radicular pain – not performed (radicular pain ranged from 8 to 19% in three trials but no stratified analyses)											
Subgroup: race/ethnicity – not performed (White ranged from 78% to 85% but no stratified analyses)											
Subgroup: economic development – not performed (all trials were conducted in high-income countries)											
Function at <1 month											
No data											
Function at 1-3 months (standardized mean difference [questionnaires include Sickness Impact Profile, 5-question ordinal scale] at 6-8 weeks)											
2[71,72,k]	RCT	very serious[b]	not serious	not serious	very serious[t]	none	47	49	-	SMD **0.16 lower** (0.91 lower to 0.58 higher)	⨁◯◯◯ Very low
Subgroup: gender/sex – not performed (female ranged from 0% to 75% but no stratified analyses)											
Subgroup: radicular pain – not performed (radicular pain ranged from 8 to 19% in three trials but no stratified analyses)											
Subgroup: race/ethnicity – not performed (White ranged from 78% to 85% but no stratified analyses)											
Subgroup: economic development – not performed (all trials were conducted in high-income countries)											
Quality of life at <1 month											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
No data											
Quality of life at 1-3 months (standardized mean difference on the Quality of Wellbeing scale at 8 weeks)											
1[71,m]	RCT	very serious[b]	serious[i]	not serious	very serious[t]	none	38	40	-	SMD **0.2 higher** (0.25 lower to 0.64 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Psychological well-being at <1 month											
No data											
Psychological well-being at 1-3 months (standardized mean difference on the Beck Depression Inventory at 8 weeks)											
1[71,m]	RCT	very serious[b]	serious[i]	not serious	very serious[t]	none	38	40	-	SMD **0.4 lower** (0.85 lower to 0.05 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Social participation											
No data											
Change in medication use											
One trial[72] reported that average analgesic usage was significantly lower during on amitriptyline compared to placebo (4.7 versus 8.7 per week, p < 0.005).											Not evaluated
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Adverse events											
2[66,71,g]	RCT	very serious[b]	not serious	not serious	serious[s]	none	46/48 (95.8%)	60/61 (98.4%)	**RR 0.99** (0.91 to 1.06)	**10 fewer per 1000** (from 89 fewer to 59 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Discontinuation due to adverse events											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-depressants	placebo	Relative (95% CI)	Absolute (95% CI)	
2[66,71,g]	RCT	very serious[b]	serious[c] I^2 = 75%	not serious	very serious[s]	none	11/71 (15.5%)	4/76 (5.3%)	**RR 2.50** (0.18 to 35.62)	**79 more per 1000** (from 43 fewer to 1000 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Constipation											
2[66,71,g]	RCT	very serious[b]	not serious	not serious	serious[s]	none	22/48 (45.8%)	13/61 (21.3%)	**RR 2.14** (1.21 to 3.78)	**243 more per 1000** (from 45 more to 592 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Somnolence											
2[66,71,g]	RCT	very serious[b]	not serious	not serious	serious[s]	none	33/48 (68.8%)	35/61 (57.4%)	**RR 1.23** (0.94 to 1.62)	**132 more per 1000** (from 34 fewer to 356 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dry mouth											
2[66,71,g]	RCT	very serious[b]	not serious	not serious	serious[s]	none	40/48 (83.3%)	37/61 (60.7%)	**RR 1.38** (1.08 to 1.74)	**230 more per 1000** (from 49 more to 449 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trials ranged from 30 to 49 years)											

Explanations

a. One parallel trial (Dickens 2000) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day) and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.

b. Risk of bias. We downgraded two levels. This was because more than 50% of participants come from studies with high risk of bias.

c. Inconsistency. We downgraded one level. This was because I2 is greater than 50% and not explained by stratified/sensitivity analyses due to limited data.
d. Imprecision. We downgraded two levels. This was because the pooled estimate crosses the null and the threshold for a large effect.
e. Three parallel trials (Atkinson 1999, Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20-30 mg/day), milnacipran (up to 200 mg/day), and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
f. Imprecision. We downgraded one level. This was because the pooled estimate crosses the null and the threshold for a small effect.
g. Two parallel trials (Atkinson 1998, Atkinson 1999), conducted in the USA, of adults with chronic low back pain with mean ages of 46-49. TCA antidepressants included nortriptyline (up 100 mg/day) and maprotiline (up to 150 mg/day).
h. One parallel trial (Dickens 2000), conducted in the United Kingdom, of adults with chronic low back pain with a mean age of 45. Paroxetine (20 mg/day).
i. Inconsistency. We downgraded one level. This was because there is only one study and inconsistency cannot be assessed.
j. Imprecision. We downgraded two levels. This was because there is no pooled estimate and fewer than 100 participants in the study.
k. Two parallel trials (Atkinson 1998, Pheasant 1983), conducted in the USA, of adults with chronic low back pain with mean ages of 46-47. TCA antidepressants included nortriptyline (up to 100 mg/day) and amitriptyline (up to 150 mg/day).
l. One crossover trial (Schukro 2016), conducted in Austria, of adults with chronic low back pain with a mean age of 58 years. The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
m. One parallel trial (Atkinson 1998), conducted in the USA, of adults with chronic low back pain with mean age of 46 years. TCA antidepressant was nortriptyline (up to 100 mg/day).
n. Two parallel trials (Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20 mg/day), milnacipran (up to 200 mg/day), and duloxetine (60 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
o. Two parallel trials (Atkinson 1999, Dickens 2000) and one crossover trial (Johnson 2011), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20-30 mg/day) and duloxetine (60 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
p. Two parallel trials (Dickens 2000, NCT01225068) and one crossover trial (Schukro 2016), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day), milnacipran (up to 200 mg/day), and duloxetine (up to 120 mg/day). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.
q. Three parallel trials (Atkinson 1999, Dickens 2000, NCT01226068), conducted in high-income countries, of adults with chronic low back pain with mean ages of 37-58. SNRI antidepressants included paroxetine (20-30 mg/day) and milnacipran (up to 200 mg/day).
r. Two parallel trials (Dickens 2000, NCT01225068), conducted in high-income countries, of adults with chronic low back pain with mean ages of 45-58. SNRI antidepressants included paroxetine (20 mg/day) and milnacipran (up to 200 mg/day).
s. Imprecision. We downgraded one level. This was because there were fewer than 200 participants in the analysis.
t. Imprecision. We downgraded two levels. This was because there were fewer than 100 participants in the analysis.

References

[67] Dickens et al. The relationship between pain and depression in a trial using paroxetine in sufferers of chronic low back pain. Psychosomatics; 2000.
[70] Schukro et al. Efficacy of duloxetine in chronic low back pain with a neuropathic component: a randomized, double-blind, placebo-controlled crossover trial. Anesthesiology; 2016.
[69] NCT01225068. Effect of milnacipran in chronic neuropathic low back pain. 2012.
[66] Atkinson et al. Effects of noradrenergic and serotonergic antidepressants on chronic low back. PAIN; 1999.
[71] Atkinson et al. A placebo-controlled randomized clinical trial of nortriptyline for chronic low. PAIN; 1998.
[72]Pheasant et al. Amitriptyline and chronic low-back pain. A randomized double-blind crossover study. Spine; 1983.
[68]Johnson et al. Effects of duloxetine and placebo in patients with chronic low back pain. The Journal of Pain; 2011

GRADE Table 9. ***Anticonvulsants (gabapentin) with treatment duration ≥ 12 weeks for chronic primary low back pain at 3 to < 6 months versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs/ studies	Types of RCTs or other study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (eg publication bias)	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (mean difference on 0 to 10 scale at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	NA	No difference (p=0.42, data otherwise not provided)	Very low
Pain (proportion with ≥30% improvement in pain at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	36% vs 36% (p=1.00, CI NR)	ARD 0% (CI NR)	Very low
Psychological well-being (mean difference on Beck Depression Inventory [0 to 63 scale] at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	NA	No difference (p=0.52), data otherwise not provided)	Very low
Serious adverse event (proportion with "marked" adverse event at 3 to <6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs/ studies	Types of RCTs or other study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (eg publication bias)	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	RR 0.19 (0.02 to 1.60)	ARD -8% (-16 to 1)	Very low
Discontinuation due to adverse events (proportion with discontinuation due to adverse event at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[d]	None noted	55	53	RR 1.35 (0.46 to 3.99)	ARD 3% (-9 to 15)	Very low
Concentration difficulties (proportion with concentration difficulties at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	RR 3.37 (1.48 to 7.70)	ARD 27% (11 to 42)	Very low
Dizziness (proportion with dizziness at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[d]	None noted	55	53	RR 1.65 (0.96 to 2.84)	ARD 17% (-0.5 to 35)	Very low
Dry mouth (proportion with dry mouth at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	RR 2.12 (1.11 to 4.04)	ARD 21% (4 to 38)	Very low
Sedation (proportion with sedation at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[d]	None noted	55	53	RR 1.84 (0.99 to 3.43)	ARD 17% (0.6 to 34)	Very low
Loss of balance (proportion with loss of balance at 3 to <6 months)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs/ studies	Types of RCTs or other study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (eg publication bias)	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Serious imprecision (-1)[c]	None noted	55	53	RR 8.67 (2.11 to 35.57)	ARD 29% (16 to 42)	Very low
Nausea/vomiting (proportion with nausea/vomiting at 3 to <6 months)											
1	RCT	Moderate (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)	None noted	55	53	RR 0.84 (0.33 to 2.16)	ARD -2% (-15 to 11)	Very low
Population subgroups, for all outcomes:											
Population subgroup 1: Gender and/or sex											
No data (proportion female in the trial was 23%)											
Population subgroup 2: Race/ethnicity											
No data											
Population subgroup 3: Presence of radicular leg pain											
No data (43% of patients had radicular pain; no analysis stratified by presence of radicular pain)											
Population subgroup 4: Regional economic development											
The single trial was conducted in the United States											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trial was 56 years)											

Explanations

a. Downgraded one level for risk of bias because the only trial was rated fair quality.
b. Downgraded one level for inconsistency because there was only one trial (unable to assess consistency).
c. Downgraded one level for imprecision because the number of participants was <100.
d. Downgraded two levels for imprecision because the confidence interval for the estimate included “no effect” and crossed the threshold for a large effect (RR ≥2).

GRADE Table 10. ***Anticonvulsants (treatment duration < 12 weeks) for chronic primary low back pain at < 1 to 3 months versus placebo***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-convulsants	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month (measured on a 0-10 scale at 3 weeks)											
2[77,79,a]	RCT	not serious	not serious	not serious	serious[b]	none	72	72	-	MD **0.16 lower** (1.05 lower to 0.72 higher)	⨁⨁⨁◯ Moderate
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pain intensity at 1-3 months (measured on a 0-10 scale at 6-10 weeks)											
3[77-79,c]	RCT	not serious	serious[d] I^2 = 53%	not serious	not serious	none	103	106	-	MD **0.89 lower** (1.72 lower to 0.06 lower)	⨁⨁⨁◯ Moderate
Subgroup: gender/sex – not performed (38%-55% female but no stratified analyses) Trials in subgroups stratified by race/ethnicity, presence of radicular pain or economic development not identified											
Function at <1 month											
No data											
Function at 1-3 months (measured on the 0-50 Oswestry Disability Index at 10 weeks)											
1[79,e]	RCT	not serious	serious[f]	not serious	very serious[g]	none	48	48	-	MD **4.9 lower** (7 lower to 2.8 lower)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Quality of life at < 1 month											
No data											
Quality of life at 1-3 months (measured on the General Health Perceptions sub-scale of the Short-Form 36 at 10 weeks)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-convulsants	placebo	Relative (95% CI)	Absolute (95% CI)	
1[79,e]	RCT	not serious	serious[f]	not serious	very serious[g]	none	48	48	-	MD **3.5 higher** (0.88 higher to 6.12 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Psychological well-being at < 1 month											
No data											
Psychological well-being at 1-3 months (measured on the Mental Health Perceptions sub-scale of the Short-Form 36 at 10 weeks)											
1[79,e]	RCT	not serious	serious[f]	not serious	very serious[g]	none	48	48	-	MD **5.4 higher** (3.14 higher to 7.66 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Social participation											
No data											
Change in medication use											
One trial[77] reported that mean analgesic consumption increased from 5.41 tablets to 6.07 tablets in the placebo phase and fell from 5.14 tablets to 5.09 tablets in the gabapentin phase. Another trial[78] reported that average number of concomitant analgesics taken fell from 4.72 to 4.27 in the gabapentin group and there was a small but statistically insignificant increase in analgesic consumption in the placebo group. Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											Not evaluated
Adverse events											
1[77,h]	RCT	not serious	serious[f]	not serious	very serious[g]	none	9/24 (37.5%)	2/24 (8.3%)	**RR 4.50** (1.08 to 18.69)	**292 more per 1000** (from 7 more to 1000 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Discontinuation due to adverse events											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-convulsants	placebo	Relative (95% CI)	Absolute (95% CI)	
1[77,h]	RCT	not serious	serious[f]	not serious	very serious[g]	none	1/24 (4.2%)	0/24 (0.0%)	**RR 3.00** (0.13 to 70.16)	**0 fewer per 1000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Nausea											
2[77,78,i]	RCT	not serious	not serious	not serious	very serious[j]	none	8/55 (14.5%)	7/58 (12.1%)	**RR 1.23** (0.48 to 3.14)	**28 more per 1000** (from 63 fewer to 258 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Constipation											
2[77,78,i]	RCT	not serious	not serious	not serious	very serious[j]	none	1/55 (1.8%)	1/58 (1.7%)	**RR 1.05** (0.11 to 9.80)	**1 more per 1000** (from 15 fewer to 152 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Dizziness											
2[77,78,i]	RCT	not serious	not serious	not serious	very serious[j]	none	10/79 (12.7%)	3/82 (3.7%)	**RR 3.08** (0.47 to 20.20)	**76 more per 1000** (from 19 fewer to 702 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Headache											
3[77-79,c]	RCT	not serious	not serious	not serious	very serious[j]	none	7/103 (6.8%)	4/106 (3.8%)	**RR 1.58** (0.49 to 5.10)	**22 more per 1000** (from 19 fewer to 155 more)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Somnolence											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-convulsants	placebo	Relative (95% CI)	Absolute (95% CI)	
3[77-79,c]	RCT	not serious	not serious	not serious	very serious[i]	none	6/103 (5.8%)	0/106 (0.0%)	**RR 5.15** (0.91 to 29.08)	**0 fewer per 1000** (from 0 fewer to 0 fewer)	⨁⨁◯◯ Low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pruritus											
1[78,k]	RCT	not serious	serious[f]	not serious	very serious[g]	none	0/31 (0.0%)	1/34 (2.9%)	**RR 0.36** (0.02 to 8.63)	**19 fewer per 1000** (from 29 fewer to 224 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Older adults (aged 60 years and over)											
No data, for all outcomes (mean age in the trials ranged from 42 to 49 years)											

Explanations

a. One parallel trial (Muehlbacher 2006) and one crossover trial (McCleane 2000), conducted in high-income countries, of adults with chronic low back pain with mean ages of 42-49. Anticonvulsants included topiramate (up to 300 mg/day) and gabapentin (individualized dosage of 15 mg/kg to the nearest 300 mg). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.

b. Imprecision. We downgraded one level. This was because the pooled estimate crosses the null and the threshold for a small effect.

c. Two parallel trials (Muehlbacher 2006, McCleane 2001) and one crossover trial (McCleane 2000), conducted in high-income countries, of adults with chronic low back pain with mean ages of 42-49. Anticonvulsants included topiramate (up to 300 mg/day) and gabapentin (one trial used a dosage of up to 1200 mg/day, and one trial used an individualized dosage of 15 mg/kg to the nearest 300 mg). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.

d. Inconsistency. We downgraded one level. This was because I2 is greater than 50% and is not explained by stratified/sensitivity analyses.

e. One parallel trial (Muehlbacher 2006), conducted in Germany, of adults with chronic low back pain with mean age of 49 years. Anticonvulsant was topiramate (up to 300 mg/day).

f. Inconsistency. We downgraded one level. This was because there is only one study and inconsistency cannot be assessed.

g. Imprecision. We downgraded two levels. This was because there is no pooled estimate and fewer than 100 participants in the study.

h. One crossover trial (McCleane 2000), conducted in Ireland, of adults with chronic low back pain with mean age of 42 years. Anticonvulsant was gabapentin (individualized dosage of 15 mg/kg to the nearest 300 mg). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.

i. One parallel trial (McCleane 2001) and one crossover trial (McCleane 2000), conducted in Ireland, of adults with chronic low back pain with mean ages of 42-44. Anticonvulsants included gabapentin (one trial used a dosage of up to 1200 mg/day, and one trial used an individualized dosage of 15 mg/kg to the nearest 300 mg). The crossover trial only reported group-level data, which we analysed in the same way as parallel studies.

j. Imprecision. We downgraded two levels. This was because the pooled estimate crosses the null and the threshold for a large effect.

k. One parallel trial (McCleane 2001), conducted in Ireland, of adults with chronic low back pain with mean age of 44 years. Anticonvulsant was gabapentin (dosage of up to 1200 mg/day).

References

[79] Muehlbacher et al. Topiramate in treatment of patients with chronic low back pain: a randomized, double-blind, placebo-controlled study. Clinical Journal of Pain; 2006.

[77] McCleane. Gabapentin reduces chronic benign nociceptive pain: a double-blind, placebo-controlled cross-over study. The Pain Clinic; 2000.

[78] McCleane. Does gabapentin have an analgesic effect on background, movement and referred pain? A randomized, double-blind, placebo controlled study. The Pain Clinic; 2001.

GRADE Table 11. ***Skeletal muscle relaxants (treatment duration < 12 weeks) for chronic primary low back pain at < 1 to 4 months versus placebo***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Muscle Relaxants	placebo	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at <1 month (proportion of participants at 3 weeks with ≥50% difference in pre- and post-treatment scores on a 0-10 scale)											
1[81,a]	RCT	not serious	serious[b]	not serious	very serious[c]	none	11/15 (73.3%)	4/16 (25.0%)	**RR 2.93** (1.19 to 7.23)	**483 more per 1000** (from 47 more to 1000 more)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pain intensity at 1-4 months (mean difference on 0 to 10 scale at 16 weeks)											
1[80,d]	RCT	not serious	serious[b]	not serious	very serious[c]	none	15	16	-	MD **0.5 higher** (1.59 lower to 2.59 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pain intensity at 1-4 months (proportion of participants at 8-16 weeks with ≥50% in pre- and post-treatment scores [two trials] or <4 out of 10 [one trial])											
3[81-83,e]	RCT	not serious	not serious	not serious	serious[h]	none	30/58 (51.7%)	9/60 (15.0%)	**RR 3.18** (1.27 to 7.95)	**327 more per 1000** (from 41 more to 1000 more)	⨁⨁⨁◯ Moderate
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Function at <1 month											
No data											
Function at 1-4 months (standardized mean difference on the Roland Morris Disability Questionnaire at 16 weeks)											
1[80,d]	RCT	not serious	serious[b]	not serious	very serious[c]	none	16	16	-	SMD **0.43 SD higher** (0.28 lower to 1.13 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Muscle Relaxants	placebo	Relative (95% CI)	Absolute (95% CI)	
Function at 1-4 months (proportion of participants at 8-16 weeks with "significant improvement" [defined differently across studies] on the Oswestry Disability Index)											
3[1,3,4,e]	RCT	not serious	not serious	not serious	serious[h]	none	37/58 (63.8%)	10/58 (17.2%)	**RR 3.49** (1.92 to 6.35)	**429 more per 1000** (from 159 more to 922 more)	⨁⨁⨁◯ Moderate
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Quality of life at <1 month											
No data											
Quality of life at 1-4 months (mean difference on 0 to 100 visual analogue scale [lower scores better] at 16 weeks)											
1[80,d]	RCT	not serious	serious[b]	not serious	very serious[c]	none	15	16	-	MD **0.33 higher** (20.68 lower to 21.34 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Psychological well-being											
No data											
Inability to work at 1-4 months (mean difference in number of sick leave days due to low back pain at 16 weeks)											
1[80,d]	RCT	not serious	serious[b]	not serious	very serious[c]	none	15	16	-	MD **4 lower** (14.37 lower to 6.37 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Change in medication use											
No data											
Adverse events (proportion of participants with any adverse event up to 16 weeks)											
4[80-83,f]	RCT	not serious	not serious	not serious	very serious[g]	none	3/76 (3.9%)	4/77 (5.2%)	**RR 0.81** (0.12 to 5.60)	**10 fewer per 1000** (from 46 fewer to 239 more)	⨁⨁◯◯ Low

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Muscle Relaxants	placebo	Relative (95% CI)	Absolute (95% CI)	
Older adults (aged 60 years and over)											
No data (mean ages in the trial ranged from 38 to 50 years)											

Explanations

a. One parallel randomized trial (Foster 2001), conducted in the USA, of adults with chronic low back pain with a mean age of 47 years. Botulinum toxin A delivered via single administration in paravertebral muscles.
b. Inconsistency. We downgraded once. This was because there is only one study and inconsistency cannot be assessed.
c. Imprecision. We downgraded twice. This was because there is no pooled estimate and fewer than 100 participants in the single study.
d. One crossover randomized trial (Cogne 2017), conducted in France, of adults with chronic low back pain with a mean age of 38 years. Botulinum toxin A delivered via single administration in paravertebral muscles. The crossover trial was analysed like a parallel trial.
e. Three parallel randomized trials (Foster 2001, Jazayeri 2011, Machado 2016). Two conducted in high-income countries (USA) and one conducted in Iran, including adults with chronic low back pain with mean ages ranging from 42 to 50 years. Botulinum toxin A delivered via single administration in paravertebral muscles.
f. Three parallel randomized trials (Foster 2001, Jazayeri 2011, Machado 2016) and one crossover trial (Cogne 2017). Three conducted in high-income countries (USA, France) and one conducted in Iran, including adults with chronic low back pain with mean ages ranging from 38 to 50 years. Botulinum toxin A delivered via single administration in paravertebral muscles. The crossover trial was analysed like a parallel trial.
g. Imprecision. We downgraded twice. This was because the pooled estimate crosses the null and the threshold for a large effect.
h. Imprecision. We downgraded one. This was because there were fewer than 200 participants in the analysis.

References

[81] Foster et al. Botulinum toxin A and chronic low back pain: a randomized, double-blind study. Neurology; 2001.
[80] Cogné et al. Are paraspinous intramuscular injections of botulinum toxin a (BoNT-A) efficient. BMC Musculoskeletal Disorders; 2017.
[83] Machado et al. Abobotulinum toxin A in the treatment of chronic low back pain. Toxins; 2016.
[82] Jazayeriet al. Efficacy of botulinum toxin type A for treating chronic low back pain. Anesthesiology and Pain Medicine; 2011.

GRADE Table 12. ***Skeletal muscle relaxants (treatment duration < 12 weeks) for chronic primary low back pain at < 1–3 months versus no treatment***

Certainty assessment							№ of patients		Effect		Certainty
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Muscle relaxants	No treatment	Relative (95% CI)	Absolute (95% CI)	
Pain intensity at < 1 month (mean difference on 0 to 10 scale at 3 weeks)											
1[84,a]	RCT	very serious[b]	serious[c]	not serious	very serious[d]	none	20	20	-	MD **0.2 lower** (1.48 lower to 1.08 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Pain intensity at 1-3 months (mean difference on 0 to 10 scale at 10 weeks)											
1[84,a]	RCT	very serious[b]	serious[c]	not serious	very serious[d]	none	15	16	-	MD **0.5 higher** (1.59 lower to 2.59 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Function at <1 month											
No data											
Function at 1-3 months (mean difference on the 0-24 Roland Morris Disability Questionnaire at 10 weeks)											
1[84,a]	RCT	very serious[b]	serious[c]	not serious	very serious[d]	none	16	16		SMD **0.43 SD higher** (0.28 lower to 1.13 higher)	⨁◯◯◯ Very low
Trials in subgroups stratified by gender/sex, race/ethnicity, presence of radicular pain or economic development not identified											
Quality of life, psychological well-being, social participation, change in use of medications or adverse events											
No data or not reported											
Older adults (aged 60 years and over)											
No data (mean age in the trial was 55 years)											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Explanations

a. One parallel randomized trial (Zaringhalam 2010), conducted in Iran, of male adults with chronic low back pain with a mean age of 55 years. Baclofen (30 mg/day) for 5 weeks compared to no treatment.
b. Risk of bias. We downgraded twice. This was because all participants were from a trial rated at high risk of bias due to lack of blinding of participants and care givers.
c. Inconsistency. We downgraded once. This was because there is only one study and inconsistency cannot be assessed.
d. Imprecision. We downgraded twice. This was because there is no pooled estimate and fewer than 100 participants in the single study.

References

[84] Zaringhalam et al. Reduction of chronic non-specific low back pain: a randomized controlled clinical trial on acupuncture and baclofen. Chinese Medicine; 2010.

GRADE Table 13. ***Systemic glucocorticoids (any treatment duration) for chronic primary low back pain versus placebo***

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs/ studies	Types of RCTs or other study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (eg publication bias)	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
All Adults											
Pain (proportion with full symptom relief or greatly improved symptoms at <2 weeks)											
1	RCT	Unclear (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[c]	None noted	38	53	RR 1.30 (0.94 to 1.78)	16% (-3.4 to 36)	Very low
Psychological wellbeing (proportion with worse mood at <2 weeks)											
1	RCT	Unclear (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[c]	None noted	38	53	RR 1.39 (0.90 to 2.16)	16% (-4.9 to 36)	Very low
Hyperglycaemia (proportion with blood sugar increase of at least 50 mg/dL at <2 weeks)											
1	RCT	Unclear (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[c]	None noted	38	53	RR 0.95 (0.54 to 1.69)	-1.6% (-21 to 18)	Very low
Weight gain (proportion with weight gain ≥1.5 kg at <2 weeks)											
1	RCT	Unclear (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[c]	None noted	38	53	RR 0.99 (0.63 to 1.57)	-0.5% (-21 to 20)	Very low
Gastrointestinal symptoms (proportion with gastrointestinal symptoms at <2 weeks)											
1	RCT	Unclear (-1)[a]	Unable to assess (-1)[b]	Direct	Very serious imprecision (-2)[c]	None noted	38	53	RR 3.49 (0.71 to 17.03)	9.4% (-2.5 to 21)	Very low
Population subgroups, for all outcomes:											

Web Annex D.D1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							Summary of findings				
							No. of participants		Effect		Certainty
No. of RCTs/ studies	Types of RCTs or other study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (eg publication bias)	Intervention	Comparator	Relative (95%CI)	Absolute (95%CI)	
Population subgroup 1: Gender and/or sex											
No data (population 31% female)											
Population subgroup 2: Race/ethnicity											
No data											
Population subgroup 3: Presence of radicular leg pain											
All patients had radicular leg pain											

Population subgroup 4: Regional economic development
The only trial was conducted in Germany

Older adults (aged 60 years and over)
No data, for all outcomes (mean age in the single trial was 47 years)

Explanations:

a. Downgraded one level for risk of bias because the only trial had unclear risk of bias.
b. Downgraded one level for inconsistency because there was only one trial (unable to assess consistency).
c. Downgraded two levels for imprecision because there were fewer than 100 participants.