World Health Organization

WHO guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings

Web Annex D: Evidence-to-Decision summaries for each intervention

WHO guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings. Web Annex D. Evidence-to-decision summaries.

ISBN 978-92-4-008570-1 (electronic version)

Suggested citation. Web Annex D. Evidence-to-decision summaries. In: WHO guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at https://iris.who.int/.

This publication forms part of the WHO guideline entitled *WHO guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings*. It is being made publicly available for transparency purposes and information, in accordance with the *WHO handbook for guideline development*, 2nd edition (2014).

Contents

Introduction

This web annex contains a summary of the benefits and harms for each intervention, by comparator and by age sub-group, for each intervention. Evidence to inform the judgements for each of the EtD domains is also included, where relevant to the intervention. Evidence related to EtD domains that is generic in nature (i.e. not related specifically to the intervention) is summarized in the guideline document, Section 4.2.

An overall summary of EtD judgements made by the GDG is provided, along with the GRADE Evidence Profile Tables.

EtD: Evidence-to-Decision

Web Annex D.A1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

A.1 Structured and standardized education/advice

Overview of the PICO structure

Definition of the intervention	
"Education and/or advice" aims to improve the understanding of the pain experience for a person with CPLBP and guide their self-management and well-being. Evidence reviewed for the guideline included "structured and standardized education and/or advice", defined as the provision of structured/standardized information delivered by health workers(s) to a person with CPLBP. This is distinct and separate from education/advice provided by a health worker to a person with CPLBP as part of a clinical encounter. Structured/standardized advice may not be tailored or personalized. Among the trials identified to inform the guideline, this intervention was delivered by health practitioners.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.A1: 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 • Health literacy • Adverse events (as reported in trials) • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Change in the use of medications • Adverse events (as reported in trials) • Falls	Critical outcomes constructs (older adults, aged ≥ 60 years) Pain

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	Peer support interventions appeared to be acceptable and sought after by some participants. They were seen as an acceptable way of gaining support and sharing information or advice. **# Review findings GRADE-CERQual Assessment of confidence** 21 Participants broadly had positive views of peer support although they found it was difficult to access and did not know of support groups in their area. Empathy and "being believed" through common experience were the most important attributes in a peer supporter. Participants believed it would be helpful to share information and receive or exchange support and advice. LOW

Summary of *feasibility considerations*	
All adults	**Older people**

Web Annex D.A1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	No evidence identified

Summary of judgements

Domain	All adults	Older people
Benefits	Small; trivial	Uncertain
Harms	Trivial; uncertain	Uncertain
Balance benefits to harms	Probably favours the intervention	Probably favours the intervention
Overall certainty	Very low	Very low
Values and preferences	Possibly important uncertainty or variability; no important uncertainty or variability	Possibly important uncertainty or variability; no important uncertainty or variability
Resource considerations	Moderate costs; varies	Moderate costs; varies
Equity and human rights	Probably increased	Probably increased
Acceptability	Yes	Yes
Feasibility	Yes; probably yes	Yes; probably yes

GRADE evidence profile tables by comparator

GRADE Table 1: ***What are the benefits and harms of education/advice in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Sham	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (high-income country, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[1]	randomized trials	very serious[a]	not serious[b]	serious[c]	very serious[d]	none	40	40	-	MD **0.22 higher** (0.05 higher to 0.39 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function (high-income country, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
1[1]	randomized trials	very serious[a]	not serious[b]	serious[c]	very serious[d]	none	40	40	-	MD **0.2 higher** (5.7 lower to 6.1 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Fear avoidance (high-income country, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: FABQ-PA; benefit indicated by lower values; scale: 0 to 24)												
1[1]	randomized trials	very serious[a]	not serious[b]	serious[c]	very serious[d]	none	40	40	-	MD **5.41 higher** (0.28 higher to 10.54 higher)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance (high-income country, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: FABQ-W; benefit indicated by lower values; scale: 0 to 42)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Sham	Relative (95% CI)	Absolute (95% CI)		
1[1]	randomized trials	very serious[a]	not serious[b]	serious[c]	very serious[d]	none	40	40	-	MD **2.64 higher** (0.54 lower to 5.82 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on fear avoidance stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
Trials on health-related quality of life, depression, catastrophizing, anxiety or self-efficacy not identified												
0												
Trials on social participation, change in use of medications, adverse events/harms or health literacy not identified												
0												
OLDER ADULTS (aged 60 years or more)												
Trials on pain, function, health-related quality of life, psychological functioning, change in use of medications, falls or adverse events/harms not identified												
0												

CI: confidence interval; **FABQ-PA:** Fear Avoidance Beliefs Questionnaire-Physical Activity outcomes**; FABQ-W:** Fear Avoidance Beliefs Questionnaire-Work outcomes**; MD:** mean difference; **NRS:** numerical rating scale; **ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size

The following was used to guide the ratings.
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).
Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.
Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.
Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).
Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations
a. We downgraded twice due to two risk of bias domains with high risk and greater than two domains with unclear risk.
b. Inconsistency: We did not downgrade; however, there are no additional studies with which to compare these findings.

c. Indirectness: We downgraded once. This is a single trial from a single country (high-income).
d. Imprecision: We downgraded twice due to small sample size (OIS would have not been reached).

References

1.Jassi FJ, Del Antonio TT,Azevedo BO,Moraes R,George SZ,Chaves TC. Star-Shape Kinesio Taping Is Not Better Than a Minimal Intervention or Sham Kinesio Taping for Pain Intensity and Postural Control in Chronic Low Back Pain: A Randomized Controlled Trial. Arch Phys Med Rehabil; 2021.

GRADE Table 2: ***What are the benefits and harms of education/advice in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to no intervention or interventions where the effect of education/ advice could be isolated?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (follow-up: closest to 3 months; assessed with: NRS, VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
10[1,2,3,4,5,6,7,8,9,10]	randomized trials	very serious[a]	serious[b]	not serious[c]	serious[d]	none	430	428	-	MD **1.1 lower** (1.63 lower to 0.56 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in males (follow-up: closest to 3 months; assessed with: VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
2[1,4]	randomized trials	very serious[a]	not serious[e]	not serious[f]	serious	none	225	225	-	MD **1.12 lower** (1.5 lower to 0.74 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in females and males (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
7[2,3,6,7,8,9,10]	randomized trials	very serious[a]	serious[g]	not serious[c]	serious[h]	none	187	186	-	MD **1.16 lower** (2.08 lower to 0.23 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in females (follow-up: closest to 3 months; assessed with: NRS, VAS; benefit indicated by lower values; scale: 0 to 10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[5]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	18	17	-	MD **0.69 lower** (1.56 lower to 0.18 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: NRS, VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
6[1,3,4,6,8,10]	randomized trials	very serious[a]	serious[l]	not serious[c]	serious[d]	none	349	351	-	MD **1.01 lower** (1.85 lower to 0.17 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in people without leg pain (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
2[2,9]	randomized trials	very serious[a]	serious[m]	serious[n]	very serious[k]	none	34	34	-	MD **1.33 lower** (12.08 lower to 9.42 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people either with or without non-radicular leg pain (follow-up: closest to 3 months; assessed with: NRS, VAS; benefit indicated by lower values; scale: 0 to 10)												
2[5,7]	randomized trials	very serious[a]	serious[b]	not serious[o]	very serious[k]	none	49	43	-	MD **1.15 lower** (7.99 lower to 5.69 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials undertaken in low- or lower middle-income countries (follow-up: closest to 3 months; assessed with: VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
2[1,4]	randomized trials	very serious[a]	not serious[e]	not serious[f]	serious[d]	none	225	225	-	MD **1.12 lower** (1.5 lower to 0.74 lower)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Pain in trials undertaken in high to upper-middle income countries (follow-up: closest to 3 months; assessed with: NRS, VAS, benefit indicated by lower values; scale: 0 to 10)												
8[2,3,5,6,7,8,9,10]	randomized trials	very serious[a]	serious[p]	not serious[c]	serious[d]	none	205	203	-	MD **1.09 lower** (1.86 lower to 0.31 lower)	⨁◯◯◯ Very low	CRITICAL
Pain stratified by race/ethnicity												
0												
Pain (education intervention: mixed content) (follow-up: closest to 3 months; assessed with: NRS, VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
5[1,3,4,6,10]	randomized trials	very serious[a]	not serious[q]	not serious[c]	serious[r]	none	329	332	-	MD **0.8 lower** (1.41 lower to 0.19 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (education intervention: pain neuroscience) (follow-up: closest to 3 months; assessed with: NRS, VAS; benefit indicated by lower values; scale: 0 to 10)												
5[2,5,7,8,9]	randomized trials	very serious[a]	serious[p]	not serious[o]	serious[h]	none	101	96	-	MD **1.47 lower** (2.57 lower to 0.37 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (education intervention delivery mode: combined verbal and written and/or electronic) (follow-up: closest to 3 months; assessed with: NRS, VAS, Chronic Pain Questionnaire; benefit indicated by lower values; scale: 0 to 10)												
7[1,2,4,5,8,9,10]	randomized trials	very serious[a]	serious[s]	not serious[c]	serious[t]	none	322	319	-	MD **1.21 lower** (1.84 lower to 0.57 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (education intervention delivery mode: verbal) (follow-up: closest to 3 months; assessed with: NRS, VAS; benefit indicated by lower values; scale: 0 to 10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
3[3,6,7]	randomized trials	very serious[a]	serious[u]	not serious[c]	very serious[v]	none	108	109	-	MD **0.68 lower** (3.19 lower to 1.83 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (after removing high risk of bias studies) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
2[2,6]	randomized trials	very serious[a]	very serious[w]	not serious[c]	very serious[x]	none	102	102	-	MD **1.1 lower** (13.41 lower to 11.22 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 6 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[6]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	74	74	-	MD **0.55 lower** (1.49 lower to 0.39 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[6]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	74	74	-	MD **1.35 lower** (2.34 lower to 0.36 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: 2 years; assessed with: VAS; benefit indicated by lower values; scale: 0 to 100)												
1[11]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	40	50	-	MD **8 lower** (18.14 lower to 2.14 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Function (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Chronic Pain Questionnaire, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
10[1,2,3,4,5,6,7,8,9,10]	randomized trials	very serious[a]	serious[p]	not serious[c]	serious[d]	none	430	428	-	SMD **0.51 lower** (0.89 lower to 0.12 lower)	⨁◯◯◯ Very low	CRITICAL
Function in males (follow-up: closest to 3 months; assessed with: RMDQ, Chronic Pain Questionnaire; benefit indicated by lower values)												
2[1,4]	randomized trials	very serious[a]	not serious[e]	not serious[f]	serious[y]	none	225	225	-	SMD **0.4 lower** (0.79 lower to 0)	⨁◯◯◯ Very low	CRITICAL
Function in females and males (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Chronic Pain Questionnaire, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
7[2,3,6,7,8,9,10]	randomized trials	very serious[a]	serious[z]	not serious[o]	serious[aa]	none	187	186	-	SMD **0.55 lower** (1.22 lower to 0.13 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
1[5]	randomized trials	very serious[a]	not serious[i]	serious[i]	very serious[k]	none	18	17	-	SMD **0.58 lower** (1.26 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Chronic Pain Questionnaire; benefit indicated by lower values)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
6[1,3,4,6,8,10]	randomized trials	very serious[a]	not serious[ab]	not serious[c]	serious[d]	none	349	351	-	SMD **0.35 lower** (0.62 lower to 0.07 lower)	⨁◯◯◯ Very low	CRITICAL
Function in people without leg pain (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
2[2,9]	randomized trials	very serious[a]	not serious[e]	serious[n]	very serious[k]	none	34	34	-	SMD **1.46 lower** (3.33 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people either with or without non-radicular leg pain (follow-up: closest to 3 months; assessed with: RMDQ, Chronic Pain Questionnaire; benefit indicated by lower values)												
2[5,7]	randomized trials	very serious[a]	not serious[e]	not serious[o]	very serious[k]	none	47	43	-	SMD **0.49 lower** (1.41 lower to 0.43 higher)	⨁◯◯◯ Very low	CRITICAL
Function in trials undertaken in low- or lower middle-income countries (follow-up: closest to 3 months; assessed with: RMDQ, Chronic Pain Questionnaire; benefit indicated by lower values)												
2[1,4]	randomized trials	very serious[a]	not serious[q]	not serious[f]	serious[y]	none	225	225	-	SMD **0.4 lower** (0.79 lower to 0)	⨁◯◯◯ Very low	CRITICAL
Function in trials undertaken in high to upper-middle income countries (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
8[2,3,5,6,7,8,9,10]	randomized trials	very serious[a]	serious[ac]	not serious[o]	serious[ad]	none	205	203	-	SMD **0.55 lower** (1.1 lower to 0)	⨁◯◯◯ Very low	CRITICAL
Function stratified by race/ethnicity												
0												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Function (education intervention: mixed content) (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Chronic Pain Questionnaire; benefit indicated by lower values)												
5[1,3,4,6,10]	randomized trials	very serious[a]	not serious[ae]	not serious[c]	serious[y]	none	329	332	-	SMD **0.28 lower** (0.68 lower to 0.11 higher)	⨁◯◯◯ Very low	CRITICAL
Function (education intervention: pain neuroscience) (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
5[2,5,7,8,9]	randomized trials	very serious[a]	not serious[af]	not serious[o]	serious[ag]	none	101	96	-	SMD **0.87 lower** (1.46 lower to 0.28 lower)	⨁◯◯◯ Very low	CRITICAL
Function (education intervention delivery mode: combined verbal, written, and/or electronic) (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Chronic Pain Questionnaire; benefit indicated by lower values)												
7[1,2,4,5,8,9,10]	randomized trials	very serious[a]	serious[ah]	not serious[c]	serious[ai]	none	322	319	-	SMD **0.68 lower** (1.08 lower to 0.28 lower)	⨁◯◯◯ Very low	CRITICAL
Function (education intervention delivery mode: verbal) (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
3[3,6,7]	randomized trials	very serious[a]	serious[aj]	not serious[o]	very serious[v]	none	108	109	-	SMD **0.08 lower** (1.52 lower to 1.36 higher)	⨁◯◯◯ Very low	CRITICAL
Function (after removing high risk of bias studies) (follow-up: closest to 3 months; assessed with: RMDQ, ODI; benefit indicated by lower values)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[2,6]	randomized trials	very serious[a]	very serious[w]	not serious[o]	very serious[x]	none	102	102	-	SMD **0.74 lower** (9.46 lower to 7.98 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 6 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 100)												
1[6]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	74	74	-	MD **2.86 lower** (7.51 lower to 1.79 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 12 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 100)												
1[6]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	74	74	-	MD **4.66 lower** (9.68 lower to 0.36 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: 2 years; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[11]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	40	50	-	MD **1.5 lower** (3.42 lower to 0.42 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
2[4,10]	randomized trials	very serious[a]	not serious[e]	not serious[c]	serious[ak]	none	150	149	-	MD **24.27 higher** (12.93 higher to 35.61 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life (unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
2[4,10]	randomized trials	very serious[a]	very serious[al]	not serious[c]	very serious[x]	none	125	125	-	MD **13.99 higher** (62.04 lower to 90.03 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (follow-up: closest to 3 months; assessed with: WHOQOL-BREF; benefit indicated by higher values; scale: 26 to 130)												
1[3]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	8	9	-	MD **9.4 lower** (17 lower to 1.8 lower)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance (high-income country) (follow-up: closest to 3 months; assessed with: TSK, TSK-11; benefit indicated by lower values)												
5[2,5,7,8,9]	randomized trials	very serious[a]	serious[am]	not serious[o]	serious[ag]	none	72	70	-	SMD **1.4 lower** (2.51 lower to 0.29 lower)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance in females and males (follow-up: closest to 3 months; assessed with: TSK, TSK-11; benefit indicated by lower values)												
4[2,7,8,9]	randomized trials	very serious[a]	serious[an]	not serious[o]	serious[aa]	none	83	79	-	SMD **1.57 lower** (3.21 lower to 0.07 higher)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance in females (follow-up: closest to 3 months; assessed with: TSK-11; benefit indicated by lower values; scale: 11 to 44)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[5]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	18	17	-	MD **7.59 lower** (12.63 lower to 2.55 lower)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance in people without leg pain (follow-up: closest to 3 months; assessed with: TSK, TSK-11; benefit indicated by lower values)												
2[2,9]	randomized trials	very serious[a]	not serious[ao]	not serious[o]	very serious[k]	none	34	34	-	SMD **2.12 lower** (7.61 lower to 3.37 higher)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance in people either with or without non-radicular leg pain (follow-up: closest to 3 months; assessed with: TSK, TSK-11; benefit indicated by lower values)												
2[5,7]	randomized trials	very serious[a]	not serious[ap]	not serious[o]	very serious[k]	none	47	43	-	SMD **0.67 lower** (3.89 lower to 2.55 higher)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance in people with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: TSK; benefit indicated by lower values)												
1[8]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	20	19	-	SMD **1.52 lower** (2.24 lower to 0.8 lower)	⨁◯◯◯ Very low	CRITICAL
Fear avoidance (after removing high risk of bias studies) (follow-up: closest to 3 months; assessed with: TSK, TSK-11; benefit indicated by lower values)												
1[2]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	28	28	-	SMD **1.95 lower** (2.59 lower to 1.31 lower)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Trials on fear avoidance stratified by race/ethnicity or low- or lower middle-income countries not identified												
0												
Fear avoidance (follow-up: 2 years; assessed with: FABQ; benefit indicated by lower values; scale: 13 to 78)												
1[11]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	40	50	-	MD **1 lower** (7.13 lower to 5.13 higher)	⨁◯◯◯ Very low	CRITICAL
Catastrophizing (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												
2[2,5]	randomized trials	very serious[a]	serious[aq]	not serious[o]	very serious[k]	none	46	45	-	MD **10.19 lower** (55.46 lower to 35.07 higher)	⨁◯◯◯ Very low	CRITICAL
Catastrophizing (females and males, no leg pain) (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												
1[2]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	28	28	-	MD **13.9 lower** (17.16 lower to 10.64 lower)	⨁◯◯◯ Very low	CRITICAL
Catastrophizing (females, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												
1[5]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	18	17	-	MD **6.77 lower** (8.48 lower to 5.06 lower)	⨁◯◯◯ Very low	CRITICAL
Catastrophizing in trials undertaken in high to upper-middle income countries (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[2,5]	randomized trials	very serious[a]	serious[aq]	not serious[o]	very serious[k]	none	46	45	-	MD **10.19 lower** (55.46 lower to 35.07 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on catastrophizing stratified by race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Catastrophizing (after removing high risk of bias studies) (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												
1[2]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	28	28	-	MD **13.9 lower** (17.16 lower to 10.64 lower)	⨁◯◯◯ Very low	CRITICAL
Depression (females and males, low-income country, either with or without leg pain unclassified) (follow-up: closest to 2 weeks; assessed with: Multidisciplinary Work-related LBP Predictor Questionnaire, Emotional Coping subscale; benefit indicated by higher values; scale: 4 to 20)												
1[12]	randomized trials	very serious[a]	not serious[i]	serious[j]	serious[ag]	none	63	62	-	MD **2.1 higher** (1.05 higher to 3.15 higher)	⨁◯◯◯ Very low	CRITICAL
Depression (females and males, low-income country, either with or without leg pain unclassified) (follow-up: closest to 6 months; assessed with: Multidisciplinary Work-related LBP Predictor Questionnaire, Emotional Coping subscale; benefit indicated by higher values; scale: 4 to 20)												
1[12]	randomized trials	very serious[a]	not serious[i]	serious[j]	serious[ag]	none	63	62	-	MD **1.5 higher** (0.5 higher to 2.5 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on anxiety, depression stratified by gender, race/ethnicity or in high to upper middle-income countries not identified												
0												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
Self-efficacy (females and males, low-income country, either with or without leg pain unclassified) (follow-up: closest to 2 weeks; assessed with: Multidisciplinary Work-related LBP Predictor Questionnaire, Self-efficacy subscale; benefit indicated by higher values; scale: 7 to 35)												
1[12]	randomized trials	very serious[a]	not serious[i]	serious[j]	serious[ag]	none	63	62	-	MD **4.4 higher** (2.77 higher to 6.03 higher)	⨁◯◯◯ Very low	CRITICAL
Self-efficacy (females and males, low-income country, either with or without leg pain unclassified) (follow-up: closest to 6 months; assessed with: Multidisciplinary Work-related LBP Predictor Questionnaire, Self-efficacy subscale; benefit indicated by higher values; scale: 7 to 35)												
1[12,ar]	randomized trials	very serious[a]	not serious[i]	serious[j]	serious[ag]	none	63	62	-	MD **1.6 higher** (0.04 higher to 3.16 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on elf-efficacy stratified by gender, race/ethnicity or in high to upper middle-income countries not identified												
0												
Social participation (paid work) (females and males, high-income country, unclassified presence of leg pain) (follow-up: 2 years; assessed with: number of sickness absence days; benefit indicated by lower values)												
1[11]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	40	50	-	MD **11 lower** (44 lower to 22 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on social participation stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on change in use of medications or health literacy not identified												
0												
Adverse events/harms (people with uncertain presence of leg pain, high-income country) (follow-up: 2 years)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[11]	randomized trials	very serious[a]	not serious[i]	serious[j]	serious[ag]	none	The trial author reported that no adverse events were reported by participants (n=90) during the interventions.				⨁◯◯◯ Very low	CRITICAL

OLDER ADULTS (aged 60 years or more)

Pain (high-income country) (follow-up: closest to 3 months; assessed with: NRS, VAS; benefit indicated by lower values; scale: 0 to 10)

2[3,5]	randomized trials	very serious[a]	not serious[e]	not serious[o]	very serious[k]	none	23	26	-	MD **0.5 lower** (5.42 lower to 4.41 higher)	⨁◯◯◯ Very low	CRITICAL

Pain (females, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)

1[5]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	18	17	-	MD **0.69 lower** (1.56 lower to 0.18 higher)	⨁◯◯◯ Very low	CRITICAL

Pain (females and males, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)

1[3]	randomized trials	very serious[a]	not serious[i]	serious[j]	very serious[k]	none	5	9	-	**0.3 higher** (2.38 lower to 2.98 higher)	⨁◯◯◯ Very low	CRITICAL

Trials on pain stratified by race/ethnicity or in adults in low- or lower middle-income countries not identified

0												

Function (high-income country) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[3,5]	randomized trials	very serious[a]	very serious[as]	not serious[c]	very serious[k]	none	23	26	-	SMD **0.02 lower** (9.79 lower to 9.76 higher)	⨁◯◯◯ Very low	CRITICAL
Function (females, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[5]	randomized trials	very serious[a]	not serious[i]	serious[i]	very serious[k]	none	18	17	-	MD **1.12 lower** (2.37 lower to 0.13 higher)	⨁◯◯◯ Very low	CRITICAL
Function (females and males, unclassified presence of leg pain) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[3]	randomized trials	very serious[a]	not serious[i]	serious[i]	very serious[k]	none	5	9	-	MD **4.52 higher** (0.46 higher to 8.58 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Fear avoidance (females, high-income country, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: TSK-11; benefit indicated by lower values)												
1[5]	randomized trials	very serious[a]	not serious[i]	serious[i]	very serious[k]	none	18	17	-	SMD **0.97 lower** (1.68 lower to 0.27 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on fear avoidance in males, stratified by race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on health-related quality of life, depression, catastrophizing, anxiety, self-efficacy, change in use of medications, falls or adverse events/harms not identified												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	No treatment	Relative (95% CI)	Absolute (95% CI)		
0												

CI: confidence interval; **FABQ:** Fear Avoidance Beliefs Questionnaire; **LBP**: low back pain; **MCS:** mental component summary; **MD:** mean difference; **n/a:** non-applicable; **NRS:** numerical rating scale; **ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size; **PCS:** Physical Component Summary; **RMDQ:** Rolland Morris Disability Questionnaire; **SF-36:** short form health survey; **SMD:** standardized mean difference; **TSK:**Tampa Scale of Kinesiophopia; **VAS:** Visual Analogue Scale; **WHOQOL-BREF:** World Health Organization Quality of Life Questionnaire – Brief version

The following was used to guide the ratings.
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).
Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.
Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.
Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).
Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations
a. Risk of bias: We downgraded twice. All of the trials were rated as overall high or unclear risk of bias.
b. Inconsistency: We downgraded once. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 54%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
c. Indirectness: We did not downgrade because the trials were conducted in different countries (high and low- or lower middle-income).
d. Imprecision: We downgraded once (studies have small sample sizes ranging from 5 to 125 participants per group). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval does not cross null; however, one of the boundaries crosses the pre-specified threshold (≥ 10% scale range or SMD ≥ 0.2).
e. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
f. Indirectness: We did not downgrade because the trials were conducted in different countries (low- or lower middle-income).
g. Inconsistency: We downgraded once. The point estimates are mostly in the same direction with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 68%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
h. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval does not cross the null; however, one of the boundaries crosses the pre-specified threshold (≥ 10% scale range or SMD ≥ 0.2).
i. Inconsistency: We did not downgrade; however, there are no additional studies with which to compare these findings.
j. Indirectness: We downgraded once. This is a single trial from a single centre (high or upper-middle income).
k. Imprecision: We downgraded twice due to small sample size (the OIS would not have been reached).
l. Inconsistency: We downgraded once. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 58%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
m. Inconsistency: We downgraded once. There is some overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 79%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
n. Indirectness: We downgraded once because the trials were conducted in the same country (high-income).
o. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).

p. Inconsistency: We downgraded once. The point estimates are or are mostly in the same direction with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 64%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
q. Inconsistency: We did not downgrade. The point estimates are mostly similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 30%).
r. Imprecision: We downgraded once (studies have small sample sizes ranging from 5 to 125 participants per group). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval does not cross null; however, one of the boundaries crosses the pre-specified threshold (≥ 10% scale range or SMD ≥ 0.2).
s. Inconsistency: We downgraded once. There is similarity in most or all of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 52%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
t. Imprecision: We downgraded once (studies have small sample sizes ranging from 6 to 125 participants per group). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval does not cross null; however, one of the boundaries crosses the pre-specified threshold (≥ 10% scale range or SMD ≥ 0.2).
u. Inconsistency: We downgraded once. There are overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 57%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
v. Imprecision: We downgraded twice. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crossed the null with the boundaries crossing the thresholds for what may be considered appreciable benefit and harm (MD ≥ 10% scale range or SMD ≥ 0.2).
w. Inconsistency: We downgraded twice. The point estimates differ without overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 94%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
x. Imprecision: We downgraded twice. The sample size is small (OIS would not have been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crossed the null with the boundaries crossing the thresholds for what may be considered appreciable benefit and harm (MD ≥ 10% scale range or SMD ≥ 0.2).
y. Imprecision: We downgraded once (studies have sample sizes ranging from 100 to 125 participants per group). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crosses the null.
z. Inconsistency: We downgraded once. There similarity is some of the point estimates with some overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 76%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
aa. Imprecision: We downgraded once . The sample size is small (OIS would not have been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crosses the null.
ab. Inconsistency: We downgraded once. There is similarity in most or all of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 49%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ac. Inconsistency: We downgraded once. The point estimates are mostly in the same direction with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 72%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
ad. Imprecision: We downgraded once (studies have sample sizes ranging from 5 to 74 participants per group). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crosses the null.
ae. Inconsistency: We did not downgrade. There is similarity in most of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 43%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
af. Inconsistency: We did not downgrade. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 50%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ag. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached).
ah. Inconsistency: We downgraded once. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 60%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ai. Imprecision: We downgraded once (studies have small sample sizes ranging from 6 to 125 participants per group).
aj. Inconsistency: We downgraded once. There is similarity in most or all of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 59%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ak. Imprecision: We downgraded once (studies have sample sizes ranging from 24 to 125 participants per group). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval does not cross the null.
al. Inconsistency: We downgraded twice. The point estimates are in the same direction with no overlap of confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 89%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
am. Inconsistency: We downgraded once. There is similarity in most of the point estimates and overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 78%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
an. Inconsistency: We downgraded once. There is similarity in some of the point estimates and some overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 83%). This could not be explained due to small subgroups and may represent substantial heterogeneity.

ao. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 34%).
ap. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 34%).
aq. Inconsistency: We downgraded once. The point estimates differ without overlapping confidence intervals, but are in the same direction. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 93%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
ar. An additional report of the same trial (Shojaei 2017, Ref. ID 22030) also assessed self-efficacy at 6 months with another scale (The Behaviour Questionnaire). We reported the estimate obtained with the Multidisciplinary Work-related LBP Predictor Questionnaire (self-efficacy subscale), since it was also used to assess self-efficacy in the immediate term (closest to 2 weeks) (Shojaei 2017, Ref. ID 25009).
as. We downgraded twice because there was high statistical heterogeneity (I2 = 81%) which could not be explained due to small subgroups. Education was favoured in Kim 2022 (SMD = -0.59; 95% CI -1.26 to 0.10); no treatment was favoured in da Silva 2014 (SMD =1.03; 95% CI -0.15 to 2.21).

References

1.Ayanniyi O, Ige GO. Back care education on peasant farmers suffering from chronic mechanical low back pain. Journal of Experimental and Integrative Medicine; 2015.
2.Bodes Pardo G, Lluch Girbes E,Roussel NA,Gallego Izquierdo T,Jimenez Penick V,Pecos Martin D. Pain Neurophysiology Education and Therapeutic Exercise for Patients With Chronic Low Back Pain: A Single-Blind Randomized Controlled Trial. Arch Phys Med Rehabil; 2018.
3.da Silva TMJC, da Silva NN,de Souza Rocha SH,et al. Back school program for back pain: education or physical exercise?. ConScientiae Saúde; 2014.
4.Ibrahimi Ghavamabadi L, Mohammadi A,Behzadi A,Dehaghi BF. Effectiveness of the training program on the low back pain and functional disability in industrial workers. International Journal of Health Promotion and Education; 2022.
5.Kim KS, An J,Kim JO,Lee MY,Lee BH. Effects of Pain Neuroscience Education Combined with Lumbar Stabilization Exercise on Strength and Pain in Patients with Chronic Low Back Pain: Randomized Controlled Trial. J Pers Med; 2022.
6.Miyamoto GC, Fagundes FRC,de Melo do Espirito Santo C,et al. Education With Therapeutic Alliance Did Not Improve Symptoms in Patients With Chronic Low Back Pain and Low Risk of Poor Prognosis Compared to Education Without Therapeutic Alliance: A Randomized Controlled Trial. 2021.
7.Pires D, Cruz EB,Caeiro C. Aquatic exercise and pain neurophysiology education versus aquatic exercise alone for patients with chronic low back pain: a randomized controlled trial. Clin Rehabil; 2015.
8.Saracoglu I, Arik MI,Afsar E,Gokpinar HH. The effectiveness of pain neuroscience education combined with manual therapy and home exercise for chronic low back pain: A single-blind randomized controlled trial. Physiother Theory Pract; 2020.
9.Tellez-Garcia M, de-la-Llave-Rincon AI,Salom-Moreno J,Palacios-Cena M,Ortega-Santiago R,Fernandez-de-Las-Penas C. Neuroscience education in addition to trigger point dry needling for the management of patients with mechanical chronic low back pain: A preliminary clinical trial. J Bodyw Mov Ther; 2015.
10.Zhang Y, Wan L,Wang X. The effect of health education in patients with chronic low back pain. J Int Med Res; 2014.
11.Rantonen J, Karppinen J,Vehtari A,et al. Effectiveness of three interventions for secondary prevention of low back pain in the occupational health setting - a randomized controlled trial with a natural course control. 2018.
12.Shojaei S, Sadat Tavafian S,Reza Jamshidi A,Wagner J,Reza Sepahvandi M. Social Cognitive Theory-Based Intervention and Low Back Pain among Health Care Workers in Qom Hospitals of Iran. 2017.

GRADE Table 3: ***What are the benefits and harms of education/advice in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Usual care	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (high or upper-middle income country) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
2[1,2]	randomized trials	very serious[a]	serious[b]	not serious[c]	very serious[d]	none	83	77	-	MD **2.49 lower** (10.73 lower to 5.75 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people with and without radicular leg pain (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[1]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	42	48	-	MD **1.8 lower** (3.03 lower to 0.57 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in people with and without non-radicular leg pain (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **3.1 lower** (4.14 lower to 2.06 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 6 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Usual care	Relative (95% CI)	Absolute (95% CI)		
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **2.1 lower** (3.13 lower to 1.07 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **7.8 lower** (14.28 lower to 1.32 lower)	⨁◯◯◯ Very low	CRITICAL
Function (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 6 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **9.2 lower** (16.5 lower to 1.9 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **2.5 higher** (1.41 lower to 6.41 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Usual care	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **9.4 higher** (2.7 higher to 16.1 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **2.4 higher** (1.56 lower to 6.36 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (high-income country, either with or without non-radicular leg pain) (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
1[2]	randomized trials	very serious[a]	not serious[e]	serious[f]	very serious[g]	none	41	29	-	MD **7.2 higher** (0.53 higher to 13.87 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on psychological functioning, social participation, change in use of medications, health literacy or adverse events/harms not identified												
0												

OLDER ADULTS (aged 60 years or more)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Education or advice	Usual care	Relative (95% CI)	Absolute (95% CI)		
Trials on pain, function, health-related quality of life, psychological functioning, change in use of medications, falls or adverse events/harms not identified												
0												

CI: confidence interval; **MD:** mean difference; **MCS:** mental component summary**; ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size; **PCS:** Physical Component Summary**; SF-36:** short form health survey; **VAS:** Visual Analogue Scale

The following was used to guide the ratings.
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).
Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.
Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.
Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).
Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations
a. Risk of bias: We downgraded twice. Trials were rated as overall high or unclear risk of bias.
b. Inconsistency: We downgraded once. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 60%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
c. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).
d. Imprecision: We downgraded twice due to small sample size (OIS would not have been reached). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 10% scale range or SMD ≥ 0.2). The confidence interval crosses the null.
e. Inconsistency: We did not downgrade; however, there are no additional studies with which to compare these findings.
f. Indirectness: We downgraded once. This is a single trial from a single centre (high-income country).
g. Imprecision: We downgraded twice due to small sample size (the OIS would not have been reached).

References
1.Akca NK, Aydin G, Gumus K. Effect of Body Mechanics Brief Education in the Clinical Setting on Pain Patients with Lumbar Disc Hernia: A Randomized Controlled Trial. International Journal of Caring Sciences; 2017.
2.Morone G, Paolucci T, Alcuri MR et al. Quality of life improved by multidisciplinary back school program in patients with chronic non-specific low back pain: a single blind randomized controlled trial. 2011.

B.1 Structured exercise therapies or programmes

Overview of the PICO structure

Definition of the intervention	
Exercise is a subcategory of physical activity that is planned, structured, repetitive and purposeful in the sense that improvement or maintenance of one or more components of physical fitness is its objective. Structured exercise therapies or programmes are prescribed or planned by health workers, often delivered with instruction and supervision and may be standardized or individualized. These therapies are broadly defined as "a series of specific movements with the aim of training or developing physical capacity (e.g. muscle and joint strength and function, range of motion or aerobic capacity) by repetition or as physical training to promote good physical health" with the goal of reducing pain and functional limitations *(1)*. They include adopting postures, movements or activities, or a combination (e.g. strengthening, stretching, aerobic exercise) of varying duration, frequency and intensity. Exercise modalities considered for the guideline included: aerobic exercise; muscle strength training; stretching, flexibility or mobilizing exercises; Yoga; core strengthening; motor control exercise; functional restoration exercise; Pilates; Tai Chi; Qigong; aquatic/hydrotherapy; and mixed exercise therapies (i.e. two or more types of exercise in which one did not clearly predominate). Among the trials identified to inform the guideline, this intervention was delivered by health practitioners.	
PICO question	
Population and subgroups	Community-dwelling adults (age 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) • Exercise type • Risk of bias judgement (low vs. not low) • 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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Outcomes	• Pain • Function • Harms/adverse events

Other Evidence-to-Decision (EtD) considerations

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** 12 Participants emphasized the importance of continuity of physical exercises to maintain mobility and to reduce pain. A lack of continuity of physical exercise and instruction could have adverse effects, such as injuries. LOW 12 Participants wanted educational materials for physical interventions which had drawings and descriptions of the exercises. LOW

Summary of *resource 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 *equity and human rights 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** 14 Participants saw the need to reduce the stigma associated with doing exercises as treatment for LBP as this was not regarded as legitimate treatment in rural Nigeria. They suggested that changes at the community level such as increasing awareness about the benefits of exercise could change negative community beliefs about exercises to legitimize exercise as treatment for back pain thereby reduce the current stigma associated with it. LOW

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** 15 Many participants liked a group format for physical exercise classes as these facilitated social support, collaborative learning and social activities, which encouraged increased attendance. MODERATE

Summary of *feasibility 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** 16 Some participants adopted physical exercise or physical supports as a part of their self-management approach to supplement conventional treatments, or when conventional treatments failed or were insufficient. Some viewed this as experimenting to find a solution. MODERATE 17 Participants requested shorter sessions of physical exercises on specific days to fit in with their daily schedule. VERY LOW

Summary of judgements

Domain	All adults	Older people
Benefits	Small; moderate; trivial; uncertain	Small; moderate
Harms	Trivial; uncertain	Uncertain
Balance benefits to harms	Favours exercise; probably favours exercise; uncertain	Probably favours exercise; uncertain
Overall certainty	Low; very low	Very low
Values and preferences	Possibly important uncertainty or variability; no important uncertainty or variability	Possibly important uncertainty or variability; no important uncertainty or variability
Resource considerations	Moderate costs; negligible costs and savings; varies (according to country and health system)	Moderate costs; negligible costs and savings; varies (according to country and health system)
Equity and human rights	Probably increased; probably reduced; no impact; varies	Probably increased; probably reduced; no impact; varies
Acceptability	Yes; probably yes; uncertain; varies	Probably yes; uncertain; varies

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Feasibility	Yes	Yes

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1: ***What are the benefits and harms of exercise in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
4[1,2,3,4,a]	randomized trials	very serious[b]	serious[c]	not serious[d]	serious[e]	none	192	152	-	MD **1.51 lower** (3.02 lower to 0)	⨁◯◯◯ Very low	CRITICAL
Pain in adults (excluding those aged 60+ years) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
3[1,2,4,a]	randomized trials	very serious[b]	not serious[f]	not serious[d]	serious[e]	none	152	112	-	MD **0.61 lower** (0.91 lower to 0.31 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[3]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **5.54 lower** (6.43 lower to 4.65 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults in low- or lower middle-income countries (follow-up: closest to 2 weeks)												
0												
Pain (core strengthening) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
1[3]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **5.54 lower** (6.43 lower to 4.65 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[1,a]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	22	10	-	MD **0.55 lower** (1.03 lower to 0.07 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (motor control exercise) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
2[2,4,a]	randomized trials	very serious[b]	serious[j]	not serious[d]	serious[e]	none	106	92	-	MD **0.87 lower** (1.66 lower to 0.09 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (stretching or flexibility/mobilizing exercise) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[1,a]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	24	10	-	MD **0.55 lower** (1.01 lower to 0.09 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (low ROB) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
1[2,a]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	77	77	-	MD **1 lower** (1.85 lower to 0.15 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (high or unclear ROB) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
3[1,3,4,a]	randomized trials	very serious[b]	serious[l]	not serious[d]	serious[e]	none	115	75	-	MD **1.6 lower** (3.44 lower to 0.24 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (motor control exercise, low ROB trial) (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[2,a]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	77	77	-	MD **1.3 lower** (2.13 lower to 0.47 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on pain in older adults (aged 60+ years) or in adults in low- or lower middle-income countries not identified												
0												
Function (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values; scale: 0 to 24)												
3[1,2,3,a]	randomized trials	very serious[b]	not serious[m]	not serious[d]	serious[e]	none	163	137	-	MD **3.29 lower** (6.22 lower to 0.36 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults (excluding those aged 60+ years) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
2[1,2,a]	randomized trials	very serious[b]	not serious[f]	not serious[d]	serious[e]	none	123	97	-	MD **2.04 lower** (2.86 lower to 1.22 lower)	⨁◯◯◯ Very low	CRITICAL
Function in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values; scale: 0 to 24)												
1[3]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **6.69 lower** (7.38 lower to 6 lower)	⨁◯◯◯ Very low	CRITICAL
Trial on function in adults in low- or lower middle-income countries not identified												
0												
Function (core strengthening) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values; scale: 0 to 24)												
1[3]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **6.69 lower** (7.38 lower to 6 lower)	⨁◯◯◯ Very low	CRITICAL
Function (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[1,a]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	22	10	-	MD **2.01 lower** (3.32 lower to 0.7 lower)	⨁◯◯◯ Very low	CRITICAL
Function (motor control exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
1[2,a]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	77	77	-	MD **2.3 lower** (4.26 lower to 0.34 lower)	⨁◯◯◯ Very low	CRITICAL
Function (stretching, or flexibility/mobilizing exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[1,a]	randomized trials	very serious[b]	not serious[g]	serious[h]	very serious[i]	none	24	10	-	MD **1.97 lower** (3.22 lower to 0.72 lower)	⨁◯◯◯ Very low	CRITICAL
Function (low ROB) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[2]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	77	77	-	MD **2.3 lower** (4.26 lower to 0.34 lower)	⨁◯◯◯ Very low	CRITICAL
Function (high or unclear ROB) (follow-up: closest to 2 weeks; assessed with: RMDQ; ODI; benefit indicated by lower values; scale: 0 to 24)												
2[1,3,a]	randomized trials	very serious[b]	not serious[n]	serious[h]	very serious[i]	none	86	60	-	MD **3.59 lower** (7.11 lower to 0.07 lower)	⨁◯◯◯ Very low	CRITICAL
Function (motor control exercise, low ROB trial) (follow-up: closest to 12 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	sham	Relative (95% CI)	Absolute (95% CI)		
1[2,a]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	77	77	-	MD **0.9 lower** (3.15 lower to 1.35 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function in older adults (aged 60+ years) or in adults in low to lower middle-income countries not identified												
0												
Harms												
1[2,o]	randomized trials	not serious[k]	not serious[g]	serious[h]	very serious[i]	none	3/77 (3.9%)	2/77 (2.6%)	**OR 1.52** (0.25 to 9.36)	**13 more per 1,000** (from 19 fewer to 174 more)	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval; **MD:** mean difference; **NRS:** Numerical Rating Scale; **ODI:** Oswestry Disability Index; **OR:** odds ration; **PSFS:** Patient-Specific Functional Scale; **RMDQ:** Roland Morris Disability Questionnaire; **VAS:** Visual Analog Scale

Explanations

a. Comparison groups were split in half for trials with multiple comparisons.
b. Risk of bias: We downgraded twice. Most or all trials were rated as overall high risk of bias.
c. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 95%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
d. Indirectness: We did not downgrade. Trials conducted in different high-income countries.
e. Imprecision: We downgraded once due to low sample size (OIS would not have been reached).
f. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
g. Inconsistency: We did not downgrade; however, there are no additional trials with which to compare findings.
h. Indirectness: We downgraded once. Trial(s) conducted in one country (high income).
i. Imprecision: We downgraded twice due to low sample size (OIS would not have been reached).
j. Inconsistency: We downgraded once. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 6%).
k. Risk of bias: We did not downgrade. Trial(s) rated as overall low risk of bias.
l. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 96%); this could not be explained due to small subgroups and may represent considerable heterogeneity.

m. Inconsistency: We did not downgrade. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 96%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
n. Inconsistency: We did not downgrade. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 99%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
o. Costa 2009: motor control exercise. Does not include older adults (60+ years). All adverse events were temporary exacerbations of pain.

References

1.Kim. Core Stability and Hip Exercises Improve Physical Function and Activity in Patients with Non-Specific Low Back Pain: A Randomized Controlled Trial. 2020.
2.Costa. Motor control exercise for chronic low back pain: a randomized placebo-controlled trial. 2009.
3.Park. A Randomized Controlled Trial Investigating the Effects of Equine Simulator Riding on Low Back Pain, Morphological Changes, and Trunk Musculature in Elderly Women. 2020.
4.Xu. Effect of Transversus abdominis muscle training on pressure-pain threshold in patients with chronic low Back pain. 2021.

GRADE Table 2: ***What are the benefits and harms of exercise in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no treatment/no additional treatment?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, ODI, MPQ; benefit indicated by lower values; scale: 0 to 10)												
41[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,a,b,c,d,e,f,g,h]	randomized trials	very serious[i]	not serious[j]	not serious[k]	not serious[l]	none	1109	959	-	MD **1.32 lower** (1.8 lower to 0.85 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults (excluding aged 60+ years) (follow-up: closest to 2 weeks; assessed with: VAS, NRS, ODI; benefit indicated by lower values; scale: 0 to 10)												
35[1,2,3,4,5,6,7,9,10,12,13,14,15,17,18,19,20,21,22,23,24,25,26,28,29,30,31,32,33,34,35,36,37,40,41,a,b,c,d,e,f,g,h]	randomized trials	very serious[i]	not serious[j]	not serious[k]	not serious[l]	none	943	793	-	MD **1.2 lower** (1.7 lower to 0.69 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: VAS, NRS, MPQ; benefit indicated by lower values; scale: 0 to 10)												
6[8,11,16,27,38,39]	randomized trials	very serious[i]	serious[m]	not serious[k]	serious[n]	none	166	166	-	MD **2.31 lower** (3.37 lower to 1.24 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults in high or upper-middle income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS, ODI, MPQ; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)	Certainty	Importance
22[1,8,9,10,11,12,14,15,16,20,22,23,24,25,26,28,29,31,33,36,37,38,a,b,c,d]	randomized trials	very serious[i]	not serious[o]	not serious[p]	not serious[l]	none	708	595	-	MD **1.23 lower** (1.57 lower to 0.89 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
19[2,3,4,5,6,7,13,17,18,19,21,27,30,32,34,35,39,40,41,a,e,f,g,h]	randomized trials	very serious[i]	serious[q]	not serious[r]	not serious[l]	none	401	364	-	MD **1.41 lower** (2.23 lower to 0.59 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (aerobic exercise) (follow-up: closest to 2 weeks; assessed with: VAS, NRS, ODI; benefit indicated by lower values; scale: 0 to 10)												
9[1,6,8,9,19,23,29,33,36,a]	randomized trials	very serious[i]	serious[s]	not serious[k]	serious[t]	none	253	214	-	MD **1.61 lower** (3.41 lower to 0.19 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (core strengthening) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
12[4,7,10,16,18,20,21,22,26,30,32,40,a,f,h]	randomized trials	very serious[i]	serious[u]	not serious[k]	serious[n]	none	196	177	-	MD **1.52 lower** (2.02 lower to 1.01 lower)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Pain (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
3[3,14,34,a]	randomized trials	serious[v]	serious[w]	not serious[k]	very serious[x]	none	92	84	-	MD **0.61 higher** (1.62 lower to 2.84 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (mixed exercise) (follow-up: closest to 2 weeks; assessed with: VAS, MPQ; benefit indicated by lower values; scale: 0 to 10)												
7[11,12,27,36,37,38,39,a,b,c,d]	randomized trials	very serious[i]	serious[y]	not serious[k]	not serious[l]	none	250	203	-	MD **1.52 lower** (2.58 lower to 0.47 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (motor control exercise) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
5[2,13,25,35,41,a]	randomized trials	very serious[i]	serious[z]	not serious[k]	very serious[x]	none	104	92	-	MD **0.78 lower** (1.79 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (Pilates) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1[28,e]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	43	43	-	MD **2.1 lower** (3.07 lower to 1.13 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (Qigong) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
2[15,24]	randomized trials	very serious[i]	not serious[ac]	serious[ab]	very serious[x]	none	60	60	-	MD **0.93 lower** (1.45 lower to 0.4 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (stretching or flexibility/mobilizing exercise) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
5[5,17,31,34,40,a,g]	randomized trials	very serious[i]	not serious[ad]	not serious[k]	very serious[x]	none	96	79	-	MD **1.52 lower** (2.08 lower to 0.95 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (Tai Chi) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[26]	randomized trials	very serious[i]	not serious[aa]	serious[ab]	very serious[x]	none	15	7	-	MD **2.38 lower** (3.16 lower to 1.6 lower)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Pain (low ROB trials) (follow-up: closest to 2 weeks; assessed with: VAS 0 to 100; benefit indicated by lower values)												
1[42]	randomized trials	not serious[ae]	not serious[aa]	serious[ab]	very serious[x]	none	Smeets 2008: 119 participants total. Mixed exercise vs no/no additional treatment. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem-solving training) vs graded activity with problem solving training alone. Between-group MD (VAS 0-100) graded activity with problem-solving training alone vs combination treatment = 5.35, 95% CI -3.73 to 14.42.				⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 3 months; assessed with: VAS, ODI; benefit indicated by lower values; scale: 0 to 10)												
5[23,33,36,37,43,a]	randomized trials	very serious[i]	not serious[af]	not serious[p]	serious[n]	none	191	156	-	MD **0.54 lower** (0.88 lower to 0.2 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on pain in older adults or in adults in low- or lower middle-income countries not identified												
0												
Pain (aerobic exercise) (follow-up: closest to 3 months; assessed with: VAS, ODI; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
3[23,33,36,a]	randomized trials	very serious[i]	not serious[af]	not serious[p]	serious[n]	none	111	70	-	MD **0.73 lower** (1.35 lower to 0.11 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (core strengthening) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[43]	randomized trials	very serious[i]	not serious[aa]	serious[ab]	very serious[x]	none	47	47	-	MD **0.53 lower** (0.97 lower to 0.09 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (mixed exercise) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
2[36,37,a]	randomized trials	serious[v]	not serious[af]	not serious[p]	very serious[x]	none	33	39	-	MD **0.05 lower** (1.13 lower to 1.02 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (low ROB trials) (follow-up: closest to 3 months)												
0												
Pain (follow-up: closest to 12 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect			
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)	Certainty	Importance
1[14,ag]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	35	35	-	MD **0.1 lower** (1.32 lower to 1.12 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain in older adults or in adults in low- or lower middle-income countries not identified												
0												
Pain (general/muscle strength training) (follow-up: closest to 12 months; assessed with: benefit indicated by lower values; scale: 0 to 10)												
1[14]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	35	35	-	MD **0.1 lower** (1.32 lower to 1.12 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (mixed exercise, low ROB trial) (follow-up: closest to 12 months; assessed with: VAS 0-100; benefit indicated by lower values)												
1[42]	randomized trials	not serious[ae]	not serious[aa]	serious[ab]	very serious[x]	none	Smeets 2008 (119 participants). Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (VAS 0-100) graded activity with problem solving training alone vs combination treatment = 6.25, 95% CI -2.94 to 15.44.			⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI, Quebec Back Pain Disability Scale, Hannover, PROMIS, WI; benefit indicated by lower values)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
39[1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18,19,21,23,24,25,27,28,29,30,31,32,33,34,35,36,37,38,40,41,44,45,46,a,ah,ai,aj,ak,al,am,an,ao,ap,aq]	randomized trials	very serious[i]	serious[ar]	not serious[k]	not serious[l]	none	1077	956	-	SMD **0.8 lower** (1.07 lower to 0.53 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults (excluding aged 60+ years) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI, Quebec Back Pain Disability Scale, Hannover, PROMIS, WI; benefit indicated by lower values)												
35[1,2,3,4,5,6,7,9,10,12,13,14,15,17,18,19,21,23,24,25,28,29,30,31,32,33,34,35,36,37,40,41,44,45,46,a,ah,ai,aj,ak,al,am,an,ao,ap,aq]	randomized trials	very serious[i]	serious[ar]	not serious[k]	not serious[l]	none	933	811	-	SMD **0.8 lower** (1.1 lower to 0.5 lower)	⨁◯◯◯ Very low	CRITICAL
Function in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												
4[8,16,27,38,a]	randomized trials	very serious[i]	serious[as]	not serious[k]	serious[n]	none	144	145	-	SMD **0.85 lower** (1.66 lower to 0.04 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults in high or upper-middle income countries (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, Hannover, PROMIS, WI; benefit indicated by lower values)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
18[1,8,9,10,12,14,15,16,23,24,25,28,29,31,33,36,37,38,a,ah,ai,aj,am,ap]	randomized trials	very serious[i]	not serious[o]	not serious[p]	not serious[l]	none	637	544	-	SMD **0.48 lower** (0.7 lower to 0.27 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI, Quebec Back Pain Disability Scale; benefit indicated by lower values)												
21[2,3,4,5,6,7,13,17,18,19,21,27,30,32,34,35,40,41,44,45,46,a,ak,al,an,ao,aq]	randomized trials	very serious[i]	not serious[at]	not serious[r]	not serious[l]	none	440	412	-	SMD **1.19 lower** (1.74 lower to 0.64 lower)	⨁⨁◯◯ Low	CRITICAL
Function (aerobic exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, Quebec Back Pain Disability Scale, Hannover, PROMIS; benefit indicated by lower values)												
10[1,6,8,9,19,23,29,33,36,44,a]	randomized trials	very serious[i]	not serious[au]	not serious[k]	not serious[l]	none	263	224	-	SMD **0.98 lower** (1.51 lower to 0.45 lower)	⨁⨁◯◯ Low	CRITICAL
Function (core strengthening) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
10[4,7,10,16,18,21,30,32,40,45,a,ak,ap,aq]	randomized trials	very serious[i]	not serious[av]	not serious[k]	serious[n]	none	186	178	-	SMD **1.08 lower** (1.47 lower to 0.69 lower)	⨁◯◯◯ Very low	CRITICAL
Function (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												
3[3,14,34,a]	randomized trials	serious[v]	serious[aw]	not serious[k]	very serious[x]	none	92	84	-	SMD **1.09 higher** (0.99 lower to 3.17 higher)	⨁◯◯◯ Very low	CRITICAL
Function (mixed exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, WI; benefit indicated by lower values)												
6[12,27,36,37,38,46,a,ah,ai,aj,am,an,ao]	randomized trials	very serious[i]	serious[ax]	not serious[k]	not serious[l]	none	233	196	-	SMD **0.83 lower** (1.38 lower to 0.29 lower)	⨁◯◯◯ Very low	CRITICAL
Function (motor control exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI; benefit indicated by lower values)												
5[2,13,25,35,41,a]	randomized trials	very serious[i]	serious[ay]	not serious[k]	very serious[x]	none	104	92	-	SMD **0.82 lower** (1.65 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Function (Pilates) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
1[28]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	43	43	-	SMD **0.74 lower** (1.18 lower to 0.3 lower)	⨁◯◯◯ Very low	CRITICAL
Function (Qigong) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
2[15,24]	randomized trials	very serious[i]	not serious[az]	serious[ab]	very serious[x]	none	60	60	-	SMD **1.16 lower** (1.87 lower to 0.45 lower)	⨁◯◯◯ Very low	CRITICAL
Function (stretching or flexibility/mobilizing exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												
5[5,17,31,34,40,a,al,ao]	randomized trials	very serious[i]	serious[ba]	not serious[k]	very serious[x]	none	96	79	-	SMD **0.62 lower** (1.36 lower to 0.13 higher)	⨁◯◯◯ Very low	CRITICAL
Function (Tai Chi) (follow-up: closest to 2 weeks; assessed with: ODI 0-50; benefit indicated by lower values)												
1[47]	randomized trials	very serious[i]	serious[aa]	serious[ab]	very serious[x]	none	Liu 2018: 43 participants total. Authors reported the average ODI score in each domain of Tai Chi group decreased significantly compared to comparison group (overall scores not reported).				⨁◯◯◯ Very low	

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Function (low ROB trials) (follow-up: closest to 2 weeks)												
1[42]	randomized trials	not serious[ae]	not serious[aa]	serious[ab]	very serious[x]	none	Smeets 2008 (119 participants). Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (RMDQ 0-24) graded activity with problem solving training alone vs combination treatment = 0.58, 95% CI -1.08 to 2.24.				⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 3 months; assessed with: ODI, Hannover, Functional Rating Test, WI; benefit indicated by lower values)												
5[23,33,37,43,48,a]	randomized trials	very serious[i]	serious[as]	not serious[k]	serious[n]	none	211	163	-	SMD **0.99 lower** (1.69 lower to 0.3 lower)	⨁◯◯◯ Very low	CRITICAL
Function in older adults (aged 60+ years) (follow-up: closest to 3 months)												
0												
Function in adults in high or upper-middle income countries (follow-up: closest to 3 months; assessed with: ODI, Hannover, WI; benefit indicated by lower values)												
4[23,33,37,43]	randomized trials	very serious[i]	not serious[af]	not serious[p]	serious[n]	none	173	129	-	SMD **0.43 lower** (0.66 lower to 0.19 lower)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Function in adults in low- or lower middle-income countries (follow-up: closest to 3 months; assessed with: Functional Rating Test; benefit indicated by lower values)												
1[48,a]	randomized trials	very serious[i]	not serious[aa]	serious[bb]	very serious[x]	none	38	34	-	SMD **2.87 lower** (6.68 lower to 0.93 higher)	⨁◯◯◯ Very low	CRITICAL
Function (aerobic exercise) (follow-up: closest to 3 months; assessed with: ODI, Hannover; benefit indicated by lower values)												
2[23,33]	randomized trials	very serious[i]	not serious[af]	not serious[p]	very serious[x]	none	102	56	-	SMD **0.27 lower** (0.6 lower to 0.07 higher)	⨁◯◯◯ Very low	CRITICAL
Function (core strengthening) (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values)												
1[43]	randomized trials	very serious[i]	not serious[aa]	serious[ab]	very serious[x]	none	47	47	-	SMD **0.66 lower** (1.07 lower to 0.24 lower)	⨁◯◯◯ Very low	CRITICAL
Function (mixed exercise) (follow-up: closest to 3 months; assessed with: WI; benefit indicated by lower values)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1[37]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	24	26	-	SMD **0.44 lower** (1.01 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL
Function (stretching or flexibility/mobilizing exercise) (follow-up: closest to 3 months; assessed with: Functional Rating Scale (unspecified scale range); benefit indicated by lower values)												
1[48,a]	randomized trials	very serious[i]	not serious[aa]	serious[bb]	very serious[x]	none	38	34	-	SMD **2.87 lower** (6.68 lower to 0.93 higher)	⨁◯◯◯ Very low	CRITICAL
Function (low ROB trials) (follow-up: closest to 3 months)												
0												
Function (follow-up: closest to 12 months; assessed with: RMDQ; benefit indicated by lower vales; scale: 0 to 24)												
1[14,bc]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	35	35	-	MD **0.2 lower** (2.73 lower to 2.33 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function in older adults or in adults in low- or lower middle-income countries not identified												
0												
Function (general strength training) (follow-up: closest to 12 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1[14]	randomized trials	serious[v]	not serious[aa]	serious[ab]	very serious[x]	none	35	35	-	MD **0.2 lower** (2.73 lower to 2.33 higher)	⨁◯◯◯ Very low	CRITICAL
Function (mixed exercise, low ROB trial) (follow-up: closest to 12 months; assessed with: RMDQ 0-24; benefit indicated by lower values)												
1[42]	randomized trials	not serious[ae]	not serious[aa]	serious[ab]	very serious[x]	none	Smeets 2008: 119 participants. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (RMDQ 0-24) graded activity with problem solving training alone vs combination treatment = 1.11, 95% CI -0.56 to 2.79.				⨁◯◯◯ Very low	CRITICAL
Harms												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
6[23,28,32,33,38,42]	randomized trials	serious[v]	not serious	not serious	not serious	none	Lang 2021 (aerobic exercise; 174 participants total): no harms reported. Miyamoto 2013 (Pilates; 86 participants total): no harms reported. Rahbar 2018 (core strengthening; 80 participants total): no harms reported. Rotter 2022 (aerobic exercise; 55 participants total): no harms reported. Smeets 2008 (mixed exercise; 119 participants total): 3 (5%) of participants in exercise group had increased back pain. Weiner 2008 (older adults) (mixed exercise; 200 participants total): no significant intervention-associated adverse events reported. One participant (2%) had increased back pain. One participant (2%) had decreased functional status.				⨁⨁⨁◯ Moderate	CRITICAL

CI: confidence interval; **Hannover:** Hannover Functional Ability Questionnaire; **MD:** mean difference; **MPQ:** McGill Pain Questionnaire; **NRS:** Numerical Rating Scale; **ODI:** Oswestry Disability Index; **PROMIS:** Patient-Reported Outcomes Measurement Information System; **PSFS:** Patient-Specific Functional Scale; **RMDQ:** Roland Morris Disability Questionnaire; **SMD:** standardized mean difference; **VAS:** Visual Analog Scale; **WI:** Waddell Disability Index

Explanations

a. Comparison groups were split for trials with multiple comparisons.

b. Dalichau 2003: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 90 participants total. Mixed exercise vs no/no additional treatment: authors reported greater pain reduction in exercise group (unclear effect estimates).

c. McIlveen 1998: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 95 participants total. Mixed exercise vs no/no additional treatment: no significant difference in the number of participants who improved more than 1 point between exercise and comparison; p=0.13 (McGill Pain Questionnaire 1-5, benefit indicated by lower values).

d. Smeets 2008: not included in meta-analysis due to missing data. Rated as low overall risk of bias; 119 participants total. Mixed exercise vs no/no additional treatment. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (VAS 0-100, benefit indicated by lower values) graded activity with problem solving training alone vs combination treatment = 5.35, 95% CI -3.73 to 14.42.

e. Sokhanguei 2017: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 34 participants total. Pilates exercise vs no/no additional treatment. Authors reported greater pain reduction in Pilates group; mean difference (SEM): -2.3 (0.72); p=0.003.

f. Kanwal 2021: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 24 participants total. Core strengthening vs no/no additional treatment. Authors reported no significant difference in pain between groups; p=0.317.

g. Raza 2020: not included in meta-analysis due to missing data. 40 participants, rated as overall high risk of bias, stretching, or flexibility/mobilizing exercise. Authors reported no significant difference in median pain between groups; p=0.112.
h. Rathi 2013: not included in meta-analysis due to missing data. 30 participants, rated as overall high risk of bias, core strengthening. Authors reported significantly greater mean pain reduction in exercise group (3.8, SD 1.0) than in no treatment group (2.9, SD 0.8); $p < 0.05$ (VAS 0-10, benefit indicated by lower values).
i. Risk of bias: We downgraded twice. Most or all trials were rated as overall high risk of bias.
j. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 97%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
k. Indirectness: We did not downgrade. Trials conducted in different countries both high and low income.
l. Imprecision: We did not downgrade. OIS would have been reached. The point estimate reached the pre-specified threshold for what may be considered appreciable benefit (MD = -1 or SMD = -0.2); the confidence interval does not cross the null.
m. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 97%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
n. Imprecision: We downgraded once due to low sample size (OIS would not have been reached).
o. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 65%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
p. Indirectness: We did not downgrade. Trials conducted in different high-income countries.
q. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 98%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
r. Indirectness: We did not downgrade. Trials conducted in different low-income countries.
s. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 99%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
t. Imprecision: We downgraded once. OIS would have been reached. The point estimate reached the pre-specified threshold for what may be considered appreciable benefit (MD = -1 or SMD = -0.2); the confidence interval crosses the null.
u. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 73%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
v. Risk of bias: We downgraded once. Some of the weight (>50%) comes from trials with unclear risk of bias.
w. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 95%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
x. Imprecision: We downgraded twice due to low sample size (OIS would not have been reached).
y. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 82%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
z. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 90%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
aa. Inconsistency: We did not downgrade; however, there are no additional trials with which to compare findings.
ab. Indirectness: We downgraded once. Trial(s) conducted in one country (high income).
ac. Inconsistency: We did not downgrade. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 43%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ad. Inconsistency: We did not downgrade. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 32%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ae. Risk of bias: We did not downgrade. Trial(s) rated as overall low risk of bias.
af. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
ag. Smeets 2008 was not included in the meta-analysis (provided within-group mean changes; no follow-up scores). 119 participants, rated as overall low risk of bias. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (VAS 0-100, benefit indicated by lower values) graded activity with problem solving training alone vs combination treatment = 6.25, 95% CI -2.94 to 15.44.

ah. Smeets 2008 was not included in the meta-analysis (provided within-group mean changes; no follow-up scores). 119 participants, rated as overall low risk of bias. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (RMDQ 0-24, benefit indicated by lower values) graded activity with problem solving training alone vs combination treatment = 0.58, 95% CI -1.08 to 2.24.
ai. Dalichau 2003: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 90 participants total. Mixed exercise vs no/no additional treatment: authors reported greater disability improvement in exercise group (unclear effect estimates).
aj. McIlveen 1998: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 95 participants total. Mixed exercise vs no/no additional treatment: authors reported significantly greater number of participants improved more than 10 points in the exercise group (27%) than in the no treatment group (8%); p=0.04 (ODI 0-100).
ak. Kanwal 2021: not included in meta-analysis due to missing data. Rated as high overall risk of bias; 24 participants total. Core strengthening vs no/no additional treatment. Authors reported no significant difference in disability between groups; p=0.692.
al. Raza 2020: not included in meta-analysis due to missing data. 40 participants, rated as overall high risk of bias, stretching, or flexibility/mobilizing exercise. Authors reported significantly lower median item scores in the exercise group for personal care (exercise: median 1, IQR 0; no treatment: median 1, IQR 1; p=0.041) and travelling (exercise: median 1, IQR 0; no treatment: median 1, IQR 0; p=0.027); no significant difference for other items (ODI individual items; 0-5).
am. Da Silva 2014: not included in meta-analysis due to missing data. 18 participants total, rated as overall high risk of bias, mixed exercise. Authors reported significantly greater mean % improvement from baseline in exercise group (45% improvement) vs no exercise (2% worsening); p=0.008 (RMDQ 0-24, benefit indicated by lower values).
an. Wattamwar 2012: not included in meta-analysis due to missing data. 24 participants total, rated as overall high risk of bias, yoga exercise. Authors reported no significant difference in change scores between groups; p=0.146.
ao. Sedaghati 2017: not included in meta-analysis due to missing data. 34 participants total, rated as overall high risk of bias, mixed exercise (in and out of water) and stretching or flexibility/mobilizing exercise. Authors reported a significant difference in follow-up scores between mixed exercise (mean 23.0, SD 3.0) and no treatment (mean 27.5, SD 3.0) (Quebec Back Pain Disability Scale 0-100, benefit indicated by lower values). No significant difference in follow-up scores between stretching or flexibility/mobilizing group and no treatment.
ap. Liu 2018: not included in meta-analysis due to missing data. 43 participants total, rated as overall high risk of bias, Tai Chi and core strengthening. Authors reported the average ODI score in each domain of both exercise groups decreased significantly compared to comparison group (overall scores not reported) (ODI 0-50, benefit indicated by lower values).
aq. Rathi 2013: not included in meta-analysis due to missing data. 30 participants total, rated as overall high risk of bias, core strengthening. Authors reported significantly greater mean disability improvement in exercise group (24.1, SD 3.2) than in no treatment group (19.73, SD 3.58); $p < 0.05$ (ODI 0-100; benefit indicated by lower values).
ar. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 87%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
as. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 89%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
at. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 92%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
au. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 84%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
av. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 63%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
aw. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 97%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
ax. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 83%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
ay. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 88%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
az. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 70%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
ba. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 82%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
bb. Indirectness: We downgraded once. Trial(s) conducted in one country (low income).

bc. Smeets 2008 was not included in the meta-analysis (provided within-group mean changes; no follow-up scores). 119 participants, rated as overall low risk of bias. Participants performed combination treatment (active physical treatment [aerobic and core strengthening exercises] + graded activity with problem solving training) vs graded activity with problem solving training alone. Between-group MD (RMDQ 0-24, benefit indicated by lower values) graded activity with problem solving training alone vs combination treatment = 1.11, 95% CI -0.56 to 2.79.

References

1.Abadi. The effect of aquatic exercise program on low-back pain disability in obese women. 2019.
2.Abdel-Aziem. The Effects of Stabilization Exercises Combined With Pelvic Floor Exercise in Women With Nonspecific Low Back Pain: A Randomized Clinical Study. 2021.
3.Afzal. Effects of virtual reality exercises and routine physical therapy on pain intensity and functional disability in patients with chronic low back pain. 2022.
4.Ahmadizadeh. The Effect of Stabilization Exercises Along With Self-care Training on Transverse Abdominal Activity, Pain, and Disability in Mothers With Low Back Pain Having Children With Cerebral Palsy: A Randomized Controlled Trial. 2020.
5.Ain. Effects of Sustained Natural Apophyseal Glides with and without thoracic posture correction techniques on mechanical back pain: a randomized control trial. 2019.
6.Alikhajeh. A Comparison of 6 Weeks of Aquatic Exercise and Kinesio Taping in Patients With Chronic Nonspecific Low Back Pain. 2021.
7.Alvani. Neuromuscular exercises on pain intensity, functional disability, proprioception, and balance of military personnel with chronic low back pain. 2021.
8.Barni. Does an Aerobic Exercise Improve Outcomes in Older Sedentary Nonspecific Low Back Pain Subjects? A Randomized Controlled Study. 2018.
9.Bruehl. Are endogenous opioid mechanisms involved in the effects of aerobic exercise training on chronic low back pain? A randomized controlled trial. 2020.
10.da Luz. Neuromuscular electrical stimulation associated with core stability exercises in nonspecific postural low back pain: a randomized clinical trial. 2019.
11.da Silva. Back school program for back pain: education or physical exercise? 2014.
12.Dalichau. Quantification of spinal configuration and postural capacity by ultrasound topometry for evaluation of different muscle strengthening programs in the therapy of back pain. 2005.
13.Dineshkumar. Effect of abdominal drawing-in maneuver along with resisted ankle dorsi-flexion to activate transverse abdominis muscle in chronic non-specific low back pain. 2015.
14.Fukuda. Does adding hip strengthening exercises to manual therapy and segmental stabilization improve outcomes in patients with nonspecific low back pain? A randomized controlled trial. 2021.
15.Gao. Effect observation of Baduanjin training on chronic nonspecific low back pain. 2018.
16.Ge. Effects of core stability training on older women with low back pain: a randomized controlled trial. 2022.
17.Hatefi. The effect of static stretching exercises on hip range of motion, pain, and disability in patients with non-specific low back pain. 2021.
18.Hemmati. Effects of consecutive supervised core stability training on pain and disability in women with nonspecific chronic low back pain. 2011.
19.Idowu. Efficacy of Graded Activity with and without Daily-Monitored-Walking on Pain and Back Endurance among Patients with Concomitant Low-Back Pain and Type-2 Diabetes: A Randomized Trial. 2020.
20.Joseph. Effects of Massage as a Combination Therapy with Lumbopelvic Stability Exercises as Compared to Standard Massage Therapy in Low Back Pain: a Randomized Cross-Over Study. 2018.
21.Karimzadeh. The effect of 8 weeks core stabilization exercises on pain and functional disability induced by low back pain in the mothers of children with cerebral palsy. 2016.
22.Kim. Effects of 3D Moving Platform Exercise on Physiological Parameters and Pain in Patients with Chronic Low Back Pain. 2020.
23.Lang. A randomized controlled trial investigating effects of an individualized pedometer driven walking program on chronic low back pain. 2021.
24.Li. Clinical observation of treating chronic nonspecific low back pain with Baduanjin combining sling exercise therapy. 2015.
25.Li. The Effect of Virtual Reality Training on Anticipatory Postural Adjustments in Patients with Chronic Nonspecific Low Back Pain: A Preliminary Study. 2021.
26.Liu. Chen-Style Tai Chi for Individuals (Aged 50 Years Old or Above) with Chronic Non-Specific Low Back Pain: A Randomized Controlled Trial. 2019.
27.Madadi-Shad. Effect of a corrective exercise program on gait kinetics and muscle activities in older adults with both low back pain and pronated feet: A double-blind, randomized controlled trial. 2020.
28.Miyamoto. Efficacy of the addition of modified Pilates exercises to a minimal intervention in patients with chronic low back pain: a randomized controlled trial. 2013.
29.Nardin. Effects of photobiomodulation and deep water running in patients with chronic non-specific low back pain: a randomized controlled trial. 2022.
30.Noormohammadpour. The Role of a Multi-Step Core Stability Exercise Program in the Treatment of Nurses with Chronic Low Back Pain: A Single-Blinded Randomized Controlled Trial. 2018.
31.Prado. Influence of isostretching on patients with chronic low back pain. A randomized controlled trial. 2019.
32.Rahbar. Effect of hippotherapy simulator on pain, disability and range of motion of the spinal column in subjects with mechanical low back pain: A randomized single-blind clinical trial. 2018.
33.Rotter. Mindful Walking in Patients with Chronic Low Back Pain: A Randomized Controlled Trial. 2022.
34.Shamsi. Effects of static stretching and strengthening exercises on flexion relaxation ratio in patients with LBP: A randomized clinical trial. 2022.
35.Srivastav. Comparison between Effectiveness of Lumbar Stabilization Exercises and Conventional Physical Therapy in the Management of Mechanical Low Back Pain . 2018.
36.Suh. The effect of lumbar stabilization and walking exercises on chronic low back pain: A randomized controlled trial. 2019.
37.Takinaci. Clinical efficacy of spa therapy (balneotherapy) for chronic low back pain: A randomized single-blind trial. 2019.
38.Weiner. Efficacy of percutaneous electrical nerve stimulation and therapeutic exercise for older adults with chronic low back pain: a randomized controlled trial. 2008.
39.Yalfani. Effects of an 8-Week Virtual Reality Training Program on Pain, Fall Risk, and Quality of Life in Elderly Women with Chronic Low Back Pain: Double-Blind Randomized Clinical Trial. 2022.

40.Zakari. Comparison of the Effects of Positional Release Therapy and Lumbar Stabilization Exercises in the Management of Chronic Mechanical Low Back Pain—Randomized Controlled Trial. 2019.
41.Karimi. The effects of consecutive supervised functional lumbar stabilizing exercises on the postural balance and functional disability in low back pain. 2014.
42.Smeets. Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. 2008.
43.Shu. Observation on the Curative Effect of Massage Manipulation Combined with Core Strength Training in Patients with Chronic Nonspecific Low Back Pain. 2021.
44.Ansari. The effect of aquatic exercise on functional disability, flexibility and function of trunk muscles in postmenopausal women with chronic non-specific low back pain: Randomized controlled trial. 2020.
45.Fouda. Effects of proprioceptive neuromuscular facilitation techniques in treating chronic nonspecific low back pain patients. 2021.
46.Heidari. Comparison of the Effects of 8 Weeks of Core Stability Exercise on Ball and Sling Exercise on The Quality of Life and Pain in The Female With Non-Specific Chronic Low Back Pain (NSLBP). 2018.
47.Liu. Effects of Tai Chi on the event-related potential of patients with chronic non-specific low back pain. 2018.
48.Shariat. The impact of modified exercise and relaxation therapy on chronic lower back pain in office workers: a randomized clinical trial. 2019.

GRADE Table 3: ***What are the benefits and harms of exercise in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
5[1,2,3,4,5,a,b,c]	randomized trials	very serious[d]	not serious[e]	not serious[f]	serious[g]	none	288	166	-	MD **0.89 lower** (1.27 lower to 0.5 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults (excluding those aged 60+ years) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
3[1,3,4,a,b,c]	randomized trials	very serious[d]	not serious[h]	not serious[f]	serious[g]	none	232	115	-	MD **0.93 lower** (1.4 lower to 0.45 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
2[2,5]	randomized trials	very serious[d]	not serious[i]	not serious[j]	very serious[k]	none	56	51	-	MD **0.65 lower** (1.5 lower to 0.19 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (high or upper-middle income countries) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
4[2,3,4,5,a,b,c]	randomized trials	very serious[d]	not serious[l]	not serious[j]	serious[g]	none	243	118	-	MD **1.01 lower** (1.32 lower to 0.7 lower)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
Pain (low- or lower middle-income countries) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[1]	randomized trials	serious[m]	not serious[n]	serious[o]	very serious[k]	none	45	48	-	MD **0.1 higher** (0.81 lower to 1.01 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (core strengthening) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[4,c]	randomized trials	very serious[d]	not serious[n]	serious[p]	very serious[k]	none	7	7	-	MD **2.3 lower** (3.96 lower to 0.64 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[3,a]	randomized trials	very serious[d]	not serious[n]	serious[p]	serious[g]	none	180	60	-	MD **1.01 lower** (1.36 lower to 0.65 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (mixed exercise) (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
3[1,2,5]	randomized trials	very serious[d]	not serious[i]	not serious[f]	serious[g]	none	101	99	-	MD **0.31 lower** (0.93 lower to 0.31 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (yoga) (follow-up: closest to 2 weeks; assessed with: Aberdeen Back Pain Scale, 0-100; benefit indicated by lower values)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
16,q	randomized trials	seriousm	not seriousn	seriousp	seriousg	none	Yoga vs usual care: difference in mean change -2.42, 95% CI -4.97 to 0.12 (313 participants total).				⨁◯◯◯ Very low	CRITICAL
Pain (low ROB trials) (follow-up: closest to 2 weeks)												
0									-		-	
Pain (older adults aged 60+ years, mixed exercise, unclear ROB trial) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
12	randomized trials	seriousm	not seritableousn	seriousp	very seriousk	none	26	22	-	MD **0.3 lower** (1.66 lower to 1.06 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (low- or lower middle-income countries) (follow-up: closest to 3 months)												
0												
Pain (yoga exercise) (follow-up: closest to 12 months; assessed with: Aberdeen Back Pain Scale, 0-100; benefit indicated by lower values)												
16,q	randomized trials	seriousm	not seriousn	seriousp	seriousg	none	Yoga vs usual care: difference in mean change -0.73, 95% CI -3.30 to 1.84 (313 participants total).				⨁◯◯◯ Very low	CRITICAL
Low ROB trial on pain or trials on pain in older adults or adults in low or lower middle-income countries not identified												
0												
Function (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI; benefit indicated by lower values; scale: 0 to 100)												
61,2,3,4,5,7,a,r	randomized trials	very seriousd	not serious s	not seriousf	not serious t	none	303	181	-	MD **9.72 lower** (13.72 lower to 5.72 lower)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
Function in adults (excluding those aged 60+ years) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, modified ODI; benefit indicated by lower values; scale: 0 to 100)												
4[1,3,4,7,a,r]	randomized trials	very serious[d]	not serious[u]	not serious[f]	serious[g]	none	247	130	-	MD **9.72 lower** (14.37 lower to 5.07 lower)	⨁◯◯◯ Very low	CRITICAL
Function in older adults (aged 60+ years) (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 100)												
2[2,5]	randomized trials	very serious[d]	not serious[i]	not serious[j]	very serious[k]	none	56	51	-	MD **9.81 lower** (16.11 lower to 3.52 lower)	⨁◯◯◯ Very low	CRITICAL
Function (high or upper-middle income countries) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values; scale: 0 to 100)												
4[2,3,4,5,a,r]	randomized trials	very serious[d]	not serious[v]	not serious[j]	serious[g]	none	243	118	-	MD **8.13 lower** (10.69 lower to 5.58 lower)	⨁◯◯◯ Very low	CRITICAL
Function (low or lower middle-income countries) (follow-up: closest to 2 weeks; assessed with: ODI, modified ODI; benefit indicated by lower values; scale: 0 to 100)												
2[1,7]	randomized trials	very serious[d]	not serious[w]	serious[o]	very serious[k]	none	60	63	-	MD **14.02 lower** (19.75 lower to 8.3 lower)	⨁◯◯◯ Very low	CRITICAL
Function (aerobic exercise) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values; scale: 0 to 100)												

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	very serious[d]	not serious[n]	serious[o]	very serious[k]	none	15	15	-	MD **16 lower** (17.59 lower to 14.41 lower)	⨁◯◯◯ Very low	CRITICAL
Function (core strengthening) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values; scale: 0 to 100)												
1[4]	randomized trials	very serious[d]	not serious[n]	serious[p]	very serious[k]	none	7	7	-	MD **4.3 lower** (9.64 lower to 1.04 higher)	⨁◯◯◯ Very low	CRITICAL
Function (general/muscle strength training) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values; scale: 0 to 100)												
1[3,a]	randomized trials	very serious[d]	not serious[n]	serious[p]	serious[g]	none	180	60	-	MD **8.95 lower** (11.96 lower to 5.93 lower)	⨁◯◯◯ Very low	CRITICAL
Function (mixed exercise) (follow-up: closest to 2 weeks; assessed with: RMDQ, modified ODI; benefit indicated by lower values; scale: 0 to 100)												
3[1,2,5]	randomized trials	very serious[d]	not serious[i]	not serious[f]	serious[g]	none	101	99	-	MD **9.77 lower** (14.64 lower to 4.89 lower)	⨁◯◯◯ Very low	CRITICAL
Function (yoga) (follow-up: closest to 2 weeks; assessed with: RMDQ, 0-24; benefit indicated by lower values)												
1[6]	randomized trials	serious[m]	not serious[n]	serious[p]	serious[g]	none	Yoga vs usual care: difference in mean change -2.17, 95% CI -3.31 to -1.03 (313 participants total).				⨁◯◯◯ Very low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	exercise	usual care	Relative (95% CI)	Absolute (95% CI)		
Function (low ROB trials) (follow-up: closest to 2 weeks)												
0									-		-	0
Function (older adults aged 60+ years, mixed exercise, unclear ROB trial) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1[2]	randomized trials	serious[m]	not serious[n]	serious[p]	very serious[k]	none	26	22	-	MD **2.3 lower** (4.92 lower to 0.32 higher)	⨁◯◯◯ Very low	CRITICAL
Function (low or lower middle-income countries) (follow-up: closest to 3 months)												
Function (yoga) (follow-up: closest to 12 months; assessed with: RMDQ 0 to 24; benefit indicated by lower values)												
1[6]	randomized trials	serious[m]	not serious[n]	serious[p]	serious[g]	none	Yoga vs usual care: difference in mean change -1.57, 95% CI -2.71 to -0.42 (313 participants total).				⨁◯◯◯ Very low	CRITICAL
Low ROB trial on function or trials of function in older adults or in adults in low or lower middle countries not identified												
0												CRITICAL
Harms												
2[5,6]	randomized trials	serious[m]	not serious	not serious[j]	serious[g]	none	Tilbrook 2011: yoga vs usual care; 313 participants total: Minor adverse events: 11 of 156 (7.1%) yoga participants events were classified as nonserious and mostly related to increased pain. Major adverse events. 1 yoga participant experienced severe pain (possibly associated with yoga). In usual care group, 1 participant died; 1 had severe accident/injury. Zadro 2019: mixed exercise vs usual care; 60 older participants total: no adverse events reported.				⨁⨁◯◯ Low	CRITICAL

CI: confidence interval; **MD:** mean difference; **NRS:** Numerical Rating Scale; **ODI:** Oswestry Disability Index; **RMDQ:** Roland Morris Disability Questionnaire; **VAS:** Visual Analog Scale

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Explanations

a. Comparison groups were split for trials with multiple comparisons.
b. Tilbrook 2011: not included in meta-analysis (only reported within-group changes; follow-up scores not provided). Rated as unclear overall risk of bias. Yoga vs usual care: difference in mean change -2.42, 95% CI -4.97 to 0.12 (313 participants total; Aberdeen Back Pain Scale 0-100, benefit indicated by lower values).
c. Raoul 2019: not included in meta-analysis due to missing data. Rated as high overall risk of bias. Core strengthening vs usual care: greater mean pain reduction in exercise group (3.91, SD 2.88) than in comparison group (1.83, SD 2.80), p<0.01(67 participants total; NRS 0-10, benefit indicated by lower values).
d. Risk of bias: We downgraded twice. Most or all trials were rated as overall high risk of bias.
e. Inconsistency: We did not downgrade. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 50%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
f. Indirectness: We did not downgrade. Trials conducted in different countries both high and low income.
g. Imprecision: We downgraded once due to low sample size (OIS would not have been reached).
h. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 65%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
i. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
j. Indirectness: We did not downgrade. Trials conducted in different high-income countries.
k. Imprecision: We downgraded twice due to low sample size (OIS would not have been reached).
l. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 26%).
m. Risk of bias: We downgraded once. Some (>50%) or all weight comes from trials with unclear risk of bias.
n. Inconsistency: We did not downgrade; however, there are no additional trials with which to compare findings.
o. Indirectness: We downgraded once. Trial(s) conducted in one country (low income).
p. Indirectness: We downgraded once. Trial(s) conducted in one country (high income).
q. Tillbrook 2011: not included in meta-analysis (only reported within-group changes; follow-up scores not provided).
r. Tillbrook 2011: not included in meta-analysis (only reported within-group changes; follow-up scores not provided). Rated as unclear overall risk of bias. Yoga vs usual care: difference in mean change -2.17, 95% CI -3.31 to -1.03 (313 participants total; RMDQ 0-24, benefit indicated by lower values).
s. Inconsistency: We did not downgrade. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 80%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
t. Imprecision: We did not downgrade. OIS would have been reached. The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -10 or SMD = -0.2); the confidence interval does not cross the null.
u. Inconsistency: We did not downgrade. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 85%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
v. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 9%).
w. Inconsistency: We did not downgrade. There is similarity in the majority of point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 59%); this could not be explained due to small subgroups and may represent substantial heterogeneity.

References

1.Chhabra. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. 2018.
2.Jinnouchi. Effects of brief self-exercise education on the management of chronic low back pain: A community-based, randomized, parallel-group pragmatic trial. 2020.
3.Kell. The response of persons with chronic nonspecific low back pain to three different volumes of periodized musculoskeletal rehabilitation. 2011.
4.Mendes. Core stabilisation exercises reduce chronic low back pain in Air Force fighter pilots: a randomized controlled trial. 2022.
5.Zadro. Video-Game-Based Exercises for Older People With Chronic Low Back Pain: A Randomized Controlledtable Trial (GAMEBACK). 2019.
6.Tillbrook. Yoga for chronic low back pain: a randomized trial. 2011.
7.Gupta. The Effectiveness of Aerobic Exercise Program for Improving Functional Performance and Quality of Life in Chronic Low Back Pain. 2019.

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 4: ***What are the benefits and harms of exercise compared with a combined comparator of placebo, no intervention or usual care for adults with chronic primary low back pain?***

This GRADE Evidence Profile Table presents data from the Cochrane review by Hayden et al. (2021) with certainty assessments conducted by an independent methodologist. The certainty assessments highlighted in green illustrate where changes have been proposed compared with the original review.

Setting: Community and health facility-based

Bibliography: Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews* 2021, Issue 9. Art. No.: CD009790. DOI: https://doi.org/10.1002/14651858.CD009790.pub2. Independent ROBIS evaluation on Hayden 2021 review and re-created GRADE table below.

Certainty assessment							№ of patients		Effect		Certainty assessment for GDG by independent methodologist	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise	Placebo, No intervention or Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (0 - 100; 0 = no pain): Earliest follow-up (time point closest to 3 months) (scale: 0 to 100)												
35[a]	randomized trials	not serious[b]	serious[c]	serious[d]	not serious	none	1531	1215	-	MD **15.22 / 100 lower** (18.26 lower to 12.18 lower)	⨁⨁◯◯ Low	CRITICAL
Functional limitations ((0 - 100; 0 = no functional limitations): Earliest follow-up (time point closest to 3 months) (scale: 0 to 100)												
38[e]	randomized trials	not serious[f]	not serious[g]	serious[d]	not serious	publication bias strongly suspected[h]	1664	1278	-	MD **6.82 /100 lower** (8.32 lower to 5.32 lower)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

CI: confidence interval; **MD:** mean difference

Explanations

a. 35 trials with 47 study groups

b. Risk of bias: *From Hayden review*: Seven studies (10 groups; 526 participants) were judged to have high risk of bias (19% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions.

c. Inconsistency: *From Hayden review*: Serious unexplained inconsistency (substantial heterogeneity I^2 = 75%, point estimates and confidence intervals varied considerably).

d. Indirectness: *From Independent ROBIS evaluation:* No trials were conducted in low-income countries and no trials were conducted on the African continent, potentially limiting the applicability to all global regions. The comparator combined usual care, placebo/sham and no intervention unlike the WHO PICO which separated these comparators; however, this was not considered a reason to further downgrade. Most trials were conducted in health facilities and few in the community, limiting generalizability to settings outside health facilities. However, this was not considered sufficient to further downgrade.

e. 38 studies with 50 study groups

f. Risk of Bias: *From Hayden review*: Nine studies (13 groups; 495 participants) were judged to have high risk of bias (17% of participant data). Exclusion of these studies in sensitivity analysis did not change conclusions.

g. Inconsistency: *From Hayden review*: Some unexplained inconsistency (moderate heterogeneity I^2 = 38%, point estimates and confidence intervals varied).

h. Other considerations: *From Hayden review:* Some evidence of publication bias (Egger's test, P = 0.005).

Reference

1. Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F et al. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine (Phila Pa 1976). 2000;25:1s-33s. doi: 10.1097/00007632-200002151-00001.

B.2 Needling therapies (traditional Chinese medicine acupuncture and other dry needling modalities)

Overview of the PICO structure

Definition of the intervention	
Needling therapies considered in the guideline included traditional Chinese medicine (TCM) acupuncture and other dry needling modalities (myofascial trigger point needling, neuroreflexotherapy and Western medical acupuncture). These modalities are defined as any intervention where needles are inserted into classical meridian points (TCM acupuncture) or soft tissue trigger points (other dry needling modalities). Manual stimulation, heating by moxa, heat lamps, cupping or electrical current stimulation could be further administered.	
PICO question	
Population and subgroups	Community-dwelling adults (age 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- or middle-income countries
Comparators	a) Placebo/sham b) No or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

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** 11 Acupuncture was valued as effective by the few participants who talked about it. However, it was viewed as providing temporary relief and was expensive. LOW

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; uncertain	Small; trivial; uncertain

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Probably favours acupuncture; probably does not favour acupuncture; uncertain	Probably favours acupuncture; probably does not favour acupuncture; Uncertain
Overall certainty	Low; very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Large costs; moderate costs; varies	Large costs, moderate costs; varies
Equity and human rights	Probably reduced; uncertain	Probably reduced; uncertain
Acceptability	Probably yes; varies	Probably yes; varies
Feasibility	Uncertain; varies	Uncertain; varies

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1: ***What are the benefits and harms of acupuncture in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to sham?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
7[1,2,3,4,5,6,7,a,b]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	581	582	-	MD **0.41 lower** (0.72 lower to 0.1 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
3[1,3,5,g]	randomized trials	very serious[c]	very serious[h]	not serious[e]	serious[i]	none	138	138	-	MD **0.41 lower** (1.31 lower to 0.49 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: VAS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
4[2,4,6,7,a]	randomized trials	serious[j]	not serious[k]	not serious[e]	not serious[f]	none	443	444	-	MD **0.42 lower** (0.75 lower to 0.09 lower)	⨁⨁⨁◯ Moderate	CRITICAL
Pain in adults with radicular leg pain (follow-up: closest to 2 weeks; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -6.85 (-16.82 to 3.11) (46 participants total).				⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Pain in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: VAS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
5[1,2,3,6,7,a,b]	randomized trials	serious[i]	not serious[s]	not serious[e]	not serious[f]	none	528	529	-	MD **0.46 lower** (0.87 lower to 0.06 lower)	⨁⨁⨁◯ Moderate	CRITICAL
Pain in adults treated with acupuncture type myofascial (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
2[4,5]	randomized trials	very serious[t]	not serious[u]	not serious[e]	very serious[r]	none	53	53	-	MD **0.3 lower** (1.06 lower to 0.45 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
5[1,2,3,5,6,a,v]	randomized trials	very serious[t]	not serious[w]	not serious[e]	serious[i]	none	188	184	-	MD **0.43 lower** (1.01 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture without stimulation (follow-up: closest to 2 weeks; assessed with: VAS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
2[4,7]	randomized trials	not serious[x]	not serious[k]	not serious[e]	not serious[f]	none	393	398	-	MD **0.4 lower** (0.75 lower to 0.06 lower)	⨁⨁⨁⨁ High	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain after removing high risk of bias studies (follow-up: closest to 2 weeks; assessed with: VAS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
3[1,4,7,v]	randomized trials	not serious[x]	serious[y]	not serious[e]	serious[z]	none	443	448	-	MD **0.68 lower** (1.26 lower to 0.1 lower)	⨁⨁◯◯ Low	CRITICAL
Pain (follow-up: closest to 3 months; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
9[1,3,4,7,9,10,11,12,13,aa,ab,ac]	randomized trials	very serious[c]	very serious[ad]	not serious[e]	not serious[ae]	none	1044	847	-	MD **0.42 lower** (0.88 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 3 months; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
4[1,3,9,13,ab,af]	randomized trials	very serious[t]	not serious[k]	not serious[e]	not serious[ae]	none	255	194	-	MD **0.38 lower** (0.86 lower to 0.1 higher)	⨁⨁◯◯ Low	CRITICAL
Pain in adults with radicular leg pain (follow-up: closest to 3 months; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -6.06 (-18.50 to 6.38) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Pain in adults with and without leg pain (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	299	159	-	MD **0.35 higher** (0.13 lower to 0.83 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: VAS, BPI, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
4[4,7,11,12]	randomized trials	very serious[c]	very serious[ag]	not serious[e]	serious[ah]	none	490	494	-	MD **0.96 lower** (1.81 lower to 0.12 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Pain in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: NRS, VAS, BPI, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
7[1,3,7,10,11,12,13,ai]	randomized trials	very serious[c]	serious[aj]	not serious[e]	not serious[ae]	none	881	754	-	MD **0.17 lower** (0.57 lower to 0.22 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture type myofascial (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[4]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	23	23	-	MD **1.96 lower** (2.79 lower to 1.13 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9,af]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.92 lower** (1.76 lower to 0.08 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with electrical stimulation (follow-up: closest to 3 months; assessed with: PROMIS, 0-100; benefit indicated by lower values)												
1[14]	randomized trials	very serious[t]	not serious[p]	serious	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -2.09 (-4.27 to 0.09) (121 participants total)				⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
5[1,3,9,11,13,ab,ai]	randomized trials	very serious[t]	not serious[ak]	not serious[e]	serious[ah]	none	312	253	-	MD **0.57 lower** (1.08 lower to 0.06 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture without stimulation (follow-up: closest to 3 months; assessed with: VAS, BPI, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
3[4,7,12]	randomized trials	very serious[c]	very serious[al]	not serious[e]	serious[z]	none	433	435	-	MD **0.83 lower** (2.01 lower to 0.34 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture (stimulation not reported) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	not serious	not serious[ae]	none	299	159	-	MD **0.35 higher** (0.13 lower to 0.83 higher)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
5[1,4,7,10,11,aa,ai]	randomized trials	very serious[am]	very serious[an]	not serious[e]	serious[z]	none	802	667	-	MD **0.55 lower** (1.21 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 6 months; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
4[7,9,10,11,aa,ao]	randomized trials	very serious[c]	not serious[ap]	not serious[e]	not serious[ae]	none	859	658	-	MD **0.21 lower** (0.58 lower to 0.16 higher)	⨁⨁◯◯ Low	CRITICAL
Pain in adults with radicular leg pain (follow-up: closest to 6 months; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -7.01 (-17.50 to 3.48) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 6 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.37 lower** (1.23 lower to 0.49 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with and without leg pain (follow-up: closest to 6 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	285	153	-	MD **0.25 higher** (0.27 lower to 0.77 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with unclassified presence of leg pain (follow-up: closest to 6 months; assessed with: VAS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
2[7,11]	randomized trials	not serious[x]	not serious[k]	not serious[e]	not serious[f]	none	434	435	-	MD **0.51 lower** (0.92 lower to 0.1 lower)	⨁⨁⨁⨁ High	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Pain in adults treated with acupuncture type TCM (follow-up: closest to 6 months; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
3[7,10,11,aa,ao]	randomized trials	very serious[c]	serious[aq]	not serious[e]	not serious[ae]	none	719	588	-	MD **0.18 lower** (0.63 lower to 0.28 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture mixed type (TCM, myofascial) (follow-up: closest to 6 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.37 lower** (1.23 lower to 0.49 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 6 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[9,11,ao]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[i]	none	197	129	-	MD **0.54 lower** (1.17 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture without stimulation (follow-up: closest to 6 months; assessed with: Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[ae]	none	377	376	-	MD **0.45 lower** (0.91 lower to 0.01 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Pain in adults treated with acupuncture (stimulation not reported) (follow-up: closest to 6 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	285	153	-	MD **0.25 higher** (0.27 lower to 0.77 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults after removing high risk of bias studies (follow-up: closest to 6 months; assessed with: VAS, NRS, Von Korff Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
3[7,10,11,aa,ao]	randomized trials	very serious[am]	not serious[aq]	not serious[e]	not serious[ae]	none	719	588	-	MD **0.18 lower** (0.63 lower to 0.28 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (follow-up: closest to 12 months; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[9,10,aa]	randomized trials	very serious[t]	not serious[ar]	not serious[e]	not serious[ae]	none	428	222	-	MD **0.02 lower** (0.51 lower to 0.47 higher)	⨁⨁◯◯ Low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 12 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.57 lower** (1.43 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with and without leg pain (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	288	152	-	MD **0.2 higher** (0.33 lower to 0.73 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Pain in adults treated with acupuncture type TCM (follow-up: closest to 12 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	288	152	-	MD **0.2 higher** (0.33 lower to 0.73 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 12 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.57 lower** (1.43 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 12 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **0.57 lower** (1.43 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture (stimulation not reported) (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	288	152	-	MD **0.2 higher** (0.33 lower to 0.73 higher)	⨁◯◯◯ Very low	CRITICAL
Pain after removing high risk of bias studies (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	not serious[ae]	none	288	152	-	MD **0.2 higher** (0.33 lower to 0.73 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
4[1,4,5,7,as]	randomized trials	very serious[c]	serious[at]	not serious[e]	serious[au]	none	478	473	-	SMD **0.22 lower** (0.54 lower to 0.11 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults without leg pain (follow-up: closest to 2 weeks; assessed with: ODI, Hannover; benefit indicated by lower values)												
2[1,5,af]	randomized trials	very serious[t]	serious[aj]	not serious[e]	very serious[r]	none	80	80	-	SMD **0.48 lower** (0.92 lower to 0.05 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: RMDQ, Hannover; benefit indicated by lower values)												
2[4,7]	randomized trials	not serious[x]	serious[av]	not serious[e]	very serious[aw]	none	398	393	-	SMD **0.03 lower** (0.37 lower to 0.31 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with radicular leg pain (follow-up: closest to 2 weeks; assessed with: ODI, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -4.52 (-13.05 to 4.01) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: ODI, Hannover)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[1,7,ax]	randomized trials	not serious[x]	very serious[ay]	not serious[e]	serious[au]	none	425	429	-	SMD **0.37 lower** (0.91 lower to 0.17 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture type myofascial (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI)												
2[4,5]	randomized trials	very serious[t]	not serious[az]	not serious[e]	very serious[r]	none	53	53	-	SMD **0** (0.5 lower to 0.5 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 2 weeks; assessed with: RMDQ, 0-24; benefit indicated by lower values)												
1[14]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: −2.11 (−3.75 to −0.47) (121 participants total)				⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with electrical stimulation (follow-up: closest to 2 weeks; assessed with: RMDQ, 0-24; benefit indicated by lower values)												
1[14]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: −2.11 (−3.75 to −0.47) (121 participants total)				⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
2[1,5,ax]	randomized trials	very serious[t]	serious[aj]	not serious[e]	very serious[r]	none	80	80	-	SMD **0.48 lower** (0.92 lower to 0.05 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture without stimulation (follow-up: closest to 2 weeks; assessed with: RMDQ, Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[4,7]	randomized trials	not serious[x]	serious[av]	not serious[e]	very serious[aw]	none	398	393	-	SMD **0.03 lower** (0.37 lower to 0.31 higher)	⨁◯◯◯ Very low	CRITICAL
Function after removing high risk of bias studies (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												
3[1,4,7,ax]	randomized trials	not serious[x]	very serious[ba]	not serious[e]	very serious[bb]	none	448	443	-	SMD **0.21 lower** (0.64 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 3 months; assessed with: RMDQ, ODI, BPI, Hannover; benefit indicated by lower values)												
7[1,4,7,9,10,11,12,aa,ax]	randomized trials	very serious[c]	not serious[bc]	not serious[e]	not serious[bd]	none	911	841	-	SMD **0.03 lower** (0.17 lower to 0.11 higher)	⨁⨁◯◯ Low	CRITICAL
Function in adults with radicular leg pain (follow-up: closest to 3 months; assessed with: ODI, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -3.04 (-12.34 to 6.25) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Function in adults without leg pain (follow-up: closest to 3 months; assessed with: ODI, Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[1,9]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[i]	none	120	190	-	SMD **0.19 lower** (0.42 lower to 0.04 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults either with or without leg pain (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[be]	none	299	159	-	SMD **0.18 higher** (0.01 lower to 0.37 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: RMDQ, ODI, BPI, Hannover; benefit indicated by lower values)												
4[4,7,11,12]	randomized trials	serious[i]	not serious[k]	not serious[e]	serious[bf]	none	492	492	-	SMD **0.13 lower** (0.26 lower to 0.01 lower)	⨁⨁◯◯ Low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: RMDQ, ODI, BPI, Hannover; benefit indicated by lower values)												
5[1,7,10,11,12,aa,ax]	randomized trials	very serious[c]	serious[bg]	not serious[e]	not serious[bh]	none	818	678	-	SMD **0** (0.17 lower to 0.17 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Function in adults treated with acupuncture type myofascial (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
1[4]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	23	23	-	SMD **0.09 higher** (0.49 lower to 0.66 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: Hannover; benefit indicated by lower values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	70	140	-	SMD **0.2 lower** (0.49 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: ODI, Hannover; benefit indicated by lower values)												
3[1,9,11,ax]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[bf]	none	177	249	-	SMD **0.17 lower** (0.37 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture without stimulation (follow-up: closest to 3 months; assessed with: RMDQ, BPI, Hannover; benefit indicated by lower values)												
3[4,7,12]	randomized trials	not serious[x]	not serious[bi]	not serious[e]	serious[bf]	none	435	433	-	SMD **0.07 lower** (0.3 lower to 0.17 higher)	⨁⨁⨁◯ Moderate	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Function in adults treated with acupuncture (stimulation not reported) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[be]	none	299	159	-	SMD **0.18 higher** (0.01 lower to 0.37 higher)	⨁◯◯◯ Very low	CRITICAL
Function after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												
5[1,4,7,10,11,aa,ax]	randomized trials	very serious[am]	serious[bj]	not serious[e]	not serious[bd]	none	805	664	-	SMD **0.02 lower** (0.18 lower to 0.15 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 6 months; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												
4[7,9,10,11,aa,ax]	randomized trials	very serious[c]	not serious[s]	not serious[e]	serious[bf]	none	788	729	-	SMD **0.1 lower** (0.22 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with radicular leg pain (follow-up: closest to 6 months; assessed with: ODI, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: 0.09 (-10.80 to 10.98) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Function in adults without leg pain (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[bk]	none	70	140	-	SMD **0.09 lower** (0.38 lower to 0.2 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with and without leg pain (follow-up: closest to 6 months; assessed with: RMDQ; benefit indicated by lower values)												
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[be]	none	285	153	-	SMD **0.06 higher** (0.14 lower to 0.26 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with unclassified presence of leg pain (follow-up: closest to 6 months; assessed with: ODI, Hannover; benefit indicated by lower values)												
2[7,11]	randomized trials	not serious[x]	not serious[k]	not serious[e]	not serious[bl]	none	433	436	-	SMD **0.21 lower** (0.34 lower to 0.07 lower)	⨁⨁⨁⨁ High	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function in adults treated with acupuncture type TCM (follow-up: closest to 6 months; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
3[7,10,11,aa,ax]	randomized trials	serious[i]	not serious[bc]	not serious[e]	serious[bf]	none	718	589	-	SMD **0.09 lower** (0.25 lower to 0.06 higher)	⨁⨁◯◯ Low	CRITICAL
Function in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[bk]	none	70	140	-	SMD **0.09 lower** (0.38 lower to 0.2 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with manual stimulation (follow-up: closest to 6 months; assessed with: ODI, Hannover; benefit indicated by lower values)												
2[9,11,ax]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[i]	none	127	199	-	SMD **0.15 lower** (0.37 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture without stimulation (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bl]	none	376	377	-	SMD **0.2 lower** (0.34 lower to 0.06 lower)	⨁⨁⨁◯ Moderate	CRITICAL
Function in adults treated with acupuncture (stimulation not reported) (follow-up: closest to 6 months; assessed with: RMDQ; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[10,aa]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[be]	none	285	153	-	SMD **0.06 higher** (0.14 lower to 0.26 higher)	⨁◯◯◯ Very low	CRITICAL
Function after removing high risk of bias studies (follow-up: closest to 6 months; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												
3[7,10,11,aa,ax]	randomized trials	serious[bm]	not serious[bc]	not serious[e]	serious[bf]	none	718	589	-	SMD **0.09 lower** (0.25 lower to 0.06 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life (follow-up: closest to 2 weeks; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												
1[6,ax]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	26	20	-	MD **6.4 higher** (6.42 lower to 19.22 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults with radicular leg pain (follow-up: closest to 2 weeks; assessed with: SF-36; benefit indicated by higher values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	No significant difference between groups for mean change from baseline on any of the subscales (46 participants total).				⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[6]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	26	20	-	MD **6.4 higher** (6.42 lower to 19.22 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												
1[6,bn]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	26	20	-	MD **6.4 higher** (6.42 lower to 19.22 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												
1[6,bn]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	26	20	-	MD **6.4 higher** (6.42 lower to 19.22 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life after removing high risk of bias studies (follow-up: closest to 2 weeks)												
1[8,,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	No improvement in acupuncture versus sham group (43 participants total)				⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (follow-up: closest to 3 months; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[11,bo]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	MD **7.78 higher** (1.41 higher to 14.15 higher)	⨁◯◯◯ Very low	CRITICAL
Health -related quality of life (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
2[7,9]	randomized trials	serious[i]	very serious[bp]	serious[q]	serious[bq]	none	510	442	-	SMD **0.25 higher** (0.07 lower to 0.56 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults without leg pain (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.43 higher** (0.14 higher to 0.72 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	serious[br]	none	370	372	-	SMD **0.11 higher** (0.03 lower to 0.25 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7,bs]	randomized trials	not serious[o]	not serious[p]	serious[q]	serious[br]	none	370	372	-	SMD **0.11 higher** (0.03 lower to 0.25 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.43 higher** (0.14 higher to 0.72 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.43 higher** (0.14 higher to 0.72 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture without stimulation (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	serious[br]	none	370	372	-	SMD **0.11 higher** (0.03 lower to 0.25 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	serious[br]	none	370	372	-	SMD **0.11 higher** (0.03 lower to 0.25 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
2[7,9]	randomized trials	serious[j]	not serious[k]	serious[q]	not serious[bt]	none	510	442	-	SMD **0.01 higher** (0.12 lower to 0.14 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life in adults without leg pain (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[bu]	none	140	70	-	SMD **0.04 lower** (0.33 lower to 0.25 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	370	372	-	SMD **0.03 higher** (0.12 lower to 0.17 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	370	372	-	SMD **0.03 higher** (0.12 lower to 0.17 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[bu]	none	140	70	-	SMD **0.04 lower** (0.33 lower to 0.25 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[bu]	none	140	70	-	SMD **0.04 lower** (0.33 lower to 0.25 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture without stimulation (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	370	372	-	SMD **0.03 higher** (0.12 lower to 0.17 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	370	372	-	SMD **0.03 higher** (0.12 lower to 0.17 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life (follow-up: closest to 6 months; assessed with: SF-36; benefit indicated by higher values; scale: 0 to 100)												
1[11,bo]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	MD **3.39 higher** (2.98 lower to 9.76 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
2[7,9]	randomized trials	serious[i]	not serious[k]	serious[q]	not serious[bl]	none	513	442	-	SMD **0.2 higher** (0.07 higher to 0.32 higher)	⨁⨁◯◯ Low	CRITICAL
Health-related quality of life in adults without leg pain (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.16 higher** (0.12 lower to 0.45 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life in adults with unclassified presence of leg pain (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bl]	none	373	372	-	SMD **0.2 higher** (0.06 higher to 0.35 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bl]	none	373	372	-	SMD **0.2 higher** (0.06 higher to 0.35 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.16 higher** (0.12 lower to 0.45 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture with manual stimulation (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.16 higher** (0.12 lower to 0.45 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture (without stimulation) (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bl]	none	373	372	-	SMD **0.2 higher** (0.06 higher to 0.35 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life after removing high risk of bias studies (follow-up: closest to 6 months; assessed with: SF-36 (PCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bl]	none	373	372	-	SMD **0.2 higher** (0.06 higher to 0.35 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
2[7,9]	randomized trials	serious[j]	very serious[bv]	serious[q]	serious[br]	none	513	442	-	SMD **0.1 higher** (0.18 lower to 0.39 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults without leg pain (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[j]	none	140	70	-	SMD **0.28 higher** (0.01 lower to 0.57 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults with unclassified presence of leg pain (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	373	372	-	SMD **0.02 lower** (0.16 lower to 0.13 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 6 weeks; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	373	372	-	SMD **0.02 lower** (0.16 lower to 0.13 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.28 higher** (0.01 lower to 0.57 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture with manual stimulation (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	SMD **0.28 higher** (0.01 lower to 0.57 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults treated with acupuncture without stimulation (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	373	372	-	SMD **0.02 lower** (0.16 lower to 0.13 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Health-related quality of life after removing high risk of bias studies (follow-up: closest to 6 months; assessed with: SF-36 (MCS); benefit indicated by higher values)												
1[7]	randomized trials	not serious[o]	not serious[p]	serious[q]	not serious[bt]	none	373	372	-	SMD **0.02 lower** (0.16 lower to 0.13 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Depression (follow-up: closest to 2 weeks; assessed with: General Depression Scale; benefit indicated by lower values; scale: 0 to 60)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **2.5 lower** (5.23 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults without leg pain (follow-up: closest to 2 weeks; assessed with: General Depression Scale; benefit indicated by lower values; scale: 0 to 60)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **2.5 lower** (5.23 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Trials on depression stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Depression in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 2 weeks; assessed with: General Depression Scale; benefit indicated by lower values; scale: 0 to 60)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **2.5 lower** (5.23 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: General Depression Scale; benefit indicated by lower values; scale: 0 to 60)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[i]	none	140	70	-	MD **2.5 lower** (5.23 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Depression (follow-up: closest to 3 months; assessed with: BDI, General Depression Scale; benefit indicated by lower values)												
2[9,11]	randomized trials	very serious[t]	not serious[ak]	not serious[e]	serious[i]	none	197	129	-	SMD **0.17 lower** (0.44 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults without leg pain (follow-up: closest to 3 months; assessed with: General Depression Scale; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[aw]	none	140	70	-	SMD **0.05 lower** (0.34 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.33 lower** (0.7 lower to 0.03 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.33 lower** (0.7 lower to 0.03 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: General Depression Scale; benefit indicated by lower values)												
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[aw]	none	140	70	-	SMD **0.05 lower** (0.34 lower to 0.23 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: BDI, General Depression Scale; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
2[9,11]	randomized trials	very serious[t]	not serious[ak]	not serious[e]	serious[i]	none	197	129	-	SMD **0.17 lower** (0.44 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on depression stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Depression after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.33 lower** (0.7 lower to 0.03 higher)	⨁◯◯◯ Very low	CRITICAL
Depression (follow-up: closest to 6 months; assessed with: BDI, General Depression Scale; benefit indicated by lower values)												
2[9,11]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[i]	none	197	129	-	SMD **0.1 lower** (0.33 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults without leg pain (follow-up: closest to 6 months; assessed with: General Depression Scale; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[aw]	none	140	70	-	SMD **0.06 lower** (0.35 lower to 0.22 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults with unclassified presence of leg pain (follow-up: closest to 6 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.17 lower** (0.53 lower to 0.2 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on depression stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Depression in adults treated with acupuncture type TCM (follow-up: closest to 6 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.17 lower** (0.53 lower to 0.2 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 6 months; assessed with: General Depression Scale; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[9]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[aw]	none	140	70	-	SMD **0.06 lower** (0.35 lower to 0.22 higher)	⨁◯◯◯ Very low	CRITICAL
Depression in adults treated with acupuncture with manual stimulation (follow-up: closest to 6 months; assessed with: BDI, General Depression Scale; benefit indicated by lower values)												
2[9,11]	randomized trials	very serious[t]	not serious[k]	not serious[e]	serious[i]	none	197	129	-	SMD **0.1 lower** (0.33 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL
Depression after removing high risk of bias studies (follow-up: closest to 6 months; assessed with: BDI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	57	59	-	SMD **0.17 lower** (0.53 lower to 0.2 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on other psychological functioning (fear avoidance, catastrophizing, anxiety, self-efficacy) or social participation not identified												
0												
Adverse events/harms during intervention period												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
6[1,5,8,9,10,14,bw,bx]	randomized trials	very serious[c]	very serious[by]	not serious[e]	serious[bz]	none	66/617 (10.7%)	35/397 (8.8%)	**OR 1.62** (0.67 to 3.90)	**47 more per 1,000** (from 27 fewer to 186 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults with radicular leg pain during intervention period												
1[8,ca]	randomized trials	not serious[cb]	not serious[p]	serious[q]	very serious[r]	none	2/23 (8.7%)	0/23 (0.0%)	**OR 5.47** (0.25 to 120.37)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults with and without leg pain during intervention period												
1[10,cc]	randomized trials	very serious[t]	not serious[p]	serious[q]	serious[bz]	none	12/315 (3.8%)	0/162 (0.0%)	**OR 13.39** (0.79 to 227.53)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults without leg pain during intervention period												
4[1,5,9,14,cd,ce]	randomized trials	very serious[t]	very serious[h]	not serious[e]	serious[bz]	none	52/279 (18.6%)	35/212 (16.5%)	**OR 1.24** (0.50 to 3.04)	**32 more per 1,000** (from 75 fewer to 210 more)	⨁◯◯◯ Very low	CRITICAL
Trials on adverse events/harms stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Adverse events/harms in adults treated with acupuncture type TCM during intervention period												
3[1,8,10,bw,cf]	randomized trials	very serious[c]	serious[cg]	not serious[e]	serious[bz]	none	22/388 (5.7%)	9/235 (3.8%)	**OR 2.77** (0.39 to 19.97)	**61 more per 1,000** (from 23 fewer to 405 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture type myofascial during intervention period												
1[5,ch]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	5/30 (16.7%)	4/30 (13.3%)	**OR 1.30** (0.31 to 5.40)	**33 more per 1,000** (from 88 fewer to 320 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture type mixed (TCM, myofascial) during intervention period												
2[9,14,ci]	randomized trials	very serious[t]	very serious[cj]	not serious[e]	serious[bz]	none	39/199 (19.6%)	22/132 (16.7%)	**OR 1.43** (0.24 to 8.50)	**56 more per 1,000** (from 121 fewer to 463 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture with manual stimulation during intervention period												
3[1,5,9,ck,cl]	randomized trials	very serious[t]	not serious[k]	not serious[e]	very serious[cm]	none	28/220 (12.7%)	25/150 (16.7%)	**OR 0.76** (0.42 to 1.36)	**35 fewer per 1,000** (from 89 fewer to 47 more)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
Adverse events/harms in adults treated with acupuncture with electrical stimulation during intervention period												
1[14,cn]	randomized trials	very serious[t]	not serious[p]	serious[q]	very serious[r]	none	24/59 (40.7%)	10/62 (16.1%)	**OR 3.57** (1.52 to 8.37)	**246 more per 1,000** (from 65 more to 456 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture without stimulation during intervention period												
1[8,ca,co]	randomized trials	not serious[cb]	not serious[p]	serious[q]	very serious[r]	none	2/23 (8.7%)	0/23 (0.0%)	**OR 5.47** (0.25 to 120.37)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture (stimulation not reported) during intervention period												
1[10,cc]	randomized trials	very serious[t]	serious[p]	serious[q]	serious[bz]	none	12/315 (3.8%)	0/162 (0.0%)	**OR 13.39** (0.79 to 227.53)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms after removing high risk of bias studies during intervention period												
3[1,8,10,cf,co]	randomized trials	very serious[t]	serious[cg]	not serious[e]	serious[bz]	none	22/388 (5.7%)	9/235 (3.8%)	**OR 2.77** (0.39 to 19.97)	**61 more per 1,000** (from 23 fewer to 405 more)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
OLDER ADULTS (aged 60 years or more)												
Pain (people with radicular leg pain, high-income country) (follow-up: closest to 2 weeks; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -6.85 (-16.82 to 3.11) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Pain (people with radicular leg pain, high-income country) (follow-up: closest to 3 months; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -6.06 (-18.50 to 6.38) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Pain (people with radicular leg pain, high-income country) (follow-up: closest to 6 months; assessed with: VAS, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -7.01 (-17.50 to 3.48) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function (people with radicular leg pain, high-income country) (follow-up: closest to 2 weeks; assessed with: ODI, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -4.52 (-13.05 to 4.01) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Function (people with radicular leg pain, high-income country) (follow-up: closest to 3 months; assessed with: ODI, 0-100; benefit indicated by lower values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: -3.04 (-12.34 to 6.25) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Function (people with radicular leg pain, high-income country) (follow-up: closest to 6 months; assessed with: ODI, 0-100; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Sham	Relative (95% CI)	Absolute (95% CI)		
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	Between-group MD (95% CI) of within-group MDs: 0.09 (-10.80 to 10.98) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life (people with radicular leg pain, high-income country) (follow-up: closest to 2 weeks; assessed with: SF-36, 0-100; benefit indicated by higher values)												
1[8,l,m,n]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	No improvement in acupuncture versus sham group (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Adverse events/harms (people with radicular leg pain, high-income country)												
1[8,l,m]	randomized trials	not serious[o]	not serious[p]	serious[q]	very serious[r]	none	No serious adverse events occurred during 4-week trial; 2 of 46 participants total (4.3%) had subcutaneous hematoma after needling (both from acupuncture group) (46 participants total)				⨁◯◯◯ Very low	CRITICAL
Trials on adverse events/harms stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on psychological functioning, change in use of medications or falls not identified												
0												

BDI: Beck Depression Inventory; **BPI:** Brief Pain Inventory; **CI:** confidence interval; **MD:** mean difference; **MCS:** Mental Component Summary; **n/a:** not applicable; **OR:** odds ratio; **NRS:** numerical rating scale; **ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size; **PCS:** Physical Component Summary; **RMDQ:** Roland Morris Disability Questionnaire; **SF-36:** Short Form Health Survey – 36-item; **SMD:** standardized mean difference; **TCM:** Traditional Chinese Medicine; **VAS:** Visual Analogue Scale

The following was used to guide the ratings.
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).

Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.
Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.
Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).
Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations
a. Yu 2020 assessed two comparisons (both included in meta-analysis).
b. Two trials were not included in the meta-analysis because they reported within-group change scores. Huang 2019: 46 participants total, rated as overall low risk of bias. Acupuncture made little or no difference to back pain: between-group MD of within-group MDs: -6.85, 95% CI -16.82 to 3.11 (VAS 0-100). Ushinohama 2016: 80 participants total; rated as overall high risk of bias. Small statistically significant difference between groups for median change in pain (p=0.032; effect size=0.21) favouring acupuncture.
c. Risk of bias: We downgraded twice because most of the weight (>50%) comes from high or unclear (i.e., some concerns) risk of bias trials.
d. Inconsistency: We did not down grade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 9%).
e. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).
f. Imprecision: We did not downgrade. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The confidence interval does not cross the null or the boundary for what may be considered appreciable benefit (MD = -1).
g. One trial was not included in the meta-analysis because it only reported a within-group change score (Ushinohama 2016: 80 participants total; rated as overall high risk of bias). Small statistically significant difference between groups for median change in pain (p=0.032; effect size=0.21) favouring acupuncture.
h. Inconsistency: We downgraded twice. There is some similarity between confidence intervals and overlapping confidence intervals; statistical heterogeneity is between 50% and 90% (i.e., I2 = 69%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
i. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved).
j. Risk of bias: We downgraded once because some of the weight (<50%) comes from high or unclear (i.e., some concerns) risk of bias studies.
k. Inconsistency: We did not downgrade. There is similarity between some or all point estimates and confidence intervals overlap; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
l. Treated with acupuncture type TCM.
m. Treated with acupuncture with manual stimulation.
n. Huang 2019 did not report follow-up scores (compared within-group changes between the 2 groups).
o. Risk of bias: We did not downgrade because all of the weight comes from low risk of bias trials.
p. Inconsistency: We did not downgrade; however, there are no other trials with which to compare findings.
q. Indirectness: We downgraded once; trial(s) conducted in one country (high or upper-middle income).
r. Imprecision: We downgraded twice. The sample size is small (OIS would not have been achieved).
s. Inconsistency: We did not downgrade. Some or all of the point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 18%).
t. Risk of bias: We downgraded twice because all of the weight comes from high or unclear (i.e., some concerns) risk of bias trials.
u. Inconsistency: We did not downgrade because statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 32%).
v. One trial was not included in the meta-analysis because it reported a within-group change score (Huang 2019: 46 participants total; rated as overall low risk of bias). Acupuncture made little or no difference to back pain: between-group MD of within-group MDs: -6.85, 95% CI -16.82 to 3.11 (VAS 0-100).
w. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 31%).

x. Risk of bias: We did not downgrade because most of the weight (>50%) comes from low risk of bias trials.
y. Inconsistency: We downgraded once. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 30% and 60% (i.e., I2 = 52%). This could not be explained due to small subgroups and may represent moderate heterogeneity.
z. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The confidence interval crosses the null. The lower boundary crosses the threshold for what may be considered appreciable benefit (-1).
aa. Cherkin 2009 assessed two comparisons (both included in meta-analysis).
ab. Kim 2020 assessed two comparisons (both included in meta-analysis).
ac. Two trials were not included in the meta-analysis because they included within-group change scores. Huang 2019: 46 participants total, rated as overall low risk of bias. Acupuncture made little or no difference to back pain: between-group MD of within-group MDs: -6.06 (-18.50 to 6.38) (VAS 0-100). Kong 2020: 121 participants total, rated as overall high risk of bias. No statistically significant difference between groups for mean change from baseline.
ad. Inconsistency: We downgraded twice. The point estimates vary and have some non-overlapping confidence intervals; statistical heterogeneity is between 50% and 90% (i.e., I2 = 68%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
ae. Imprecision: We did not downgrade. The point estimate did not reach the threshold for what may be considered clinically important (MD ≥ 1). The confidence interval crosses the null but not the boundaries for appreciable benefit (MD = -1) or harm (MD = +1).
af. One trial was not included in the meta-analysis because it included a within-group change score. Kong 2020: 121 participants total, rated as high overall risk of bias. No statistically significant difference between groups for mean change from baseline.
ag. Inconsistency: We downgraded twice. The point estimates vary and have some non-overlapping confidence intervals; statistical heterogeneity is between 75% and 100% (i.e., I2 = 78%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
ah. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The confidence interval does not cross the null; the lower boundary crosses the threshold for what may be considered appreciable benefit (MD = -1).
ai. One trial was not included in the meta-analysis because it reported a within-group change score (Huang 2019: 46 participants total; rated as overall low risk of bias). Acupuncture made little or no difference to back pain: between-group MD of within-group MDs: -6.06 (-18.50 to 6.38) (VAS 0-100).
aj. Inconsistency: We downgraded once. The point estimates vary and have some overlapping confidence intervals; statistical heterogeneity is between 30% and 60% (i.e., I2 = 45%). This could not be explained due to small subgroups and may represent moderate heterogeneity.
ak. Inconsistency: We did not downgrade. There is similarity between some point estimates and overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 28%).
al. Inconsistency: We downgraded twice. The point estimates vary and have some non-overlapping confidence intervals; statistical heterogeneity is between 75% and 100% (i.e., I2 = 83%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
am. Risk of bias: We downgraded twice because most of the weight (>50%) comes from unclear (i.e., some concerns) risk of bias studies.
an. Inconsistency: We downgraded twice. The point estimates vary and have some non-overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 82%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
ao. One trial was not included in the meta-analysis because it reported a within-group change score (Huang 2019: 46 participants total; rated as overall low risk of bias). Acupuncture made little or no difference to back pain: between-group MD of within-group MDs: -7.01 (-17.50 to 3.48) (VAS 0-100).
ap. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 27%).
aq. Inconsistency: We downgraded once. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 44%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ar. Inconsistency: We did not downgrade. There is similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 16%).
as. Two trials were not included in the meta-analysis because they included within-group change scores. Huang 2019: 46 participants total, rated as overall low risk of bias. No significant difference between groups for mean change from baseline. Kong 2020: 121 participants total, rated as overall high risk of bias. No statistically significant difference between groups for mean change from baseline.
at. Inconsistency: We downgraded once. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 66%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
au. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The confidence interval crosses the null.
av. Inconsistency: We downgraded once. The point estimates differ with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 42%); this could not be explained due to small subgroups and may represent moderate heterogeneity.

aw. Imprecision: We downgraded twice. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-0.2), and the upper boundary crosses the threshold for what may be considered appreciable harm (+0.2).
ax. One trial was not included in the meta-analysis because it reported a within-group change score (Huang 2019: 46 participants total; rated as overall low risk of bias). No significant difference between groups for mean change from baseline.
ay. Inconsistency: We downgraded twice. The point estimates vary with little overlap in confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 84%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
az. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals. Statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 40%).
ba. Inconsistency: We downgraded twice. The point estimates vary with little overlap in confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 77%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
bb. Imprecision: We downgraded twice. The point estimate reached the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable harm (+0.2).
bc. Inconsistency: We did not downgrade. There is some similarity in point estimates and overlapping confidence intervals. Statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 38%).
bd. Imprecision: We did not downgrade. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper and lower boundaries of the 95% CI do not cross the threshold for what may be considered appreciable benefit (-0.2) or harm (+0.2).
be. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable harm (+0.2), but the lower boundary does not cross the threshold for what may be considered appreciable benefit (-0.2).
bf. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-0.2), but the upper boundary does not cross the threshold for what may be considered appreciable harm (+0.2).
bg. Inconsistency: We downgraded once. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 51%). This could not be explained due to small subgroups and may represent moderate heterogeneity.
bh. Imprecision: We did not downgrade. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper and lower boundaries of the 95% CI do not cross the threshold for what may be considered appreciable benefit (-0.2) or harm (+0.2).
bi. Inconsistency: We did not downgrade. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 31%).
bj. Inconsistency: We downgraded once. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 46%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
bk. Imprecision: We downgraded twice. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-0.2), and the upper boundary crosses the threshold for what may be considered appreciable harm (+0.2).
bl. Imprecision: We did not downgrade. The point estimate reached the threshold for what may be considered appreciable benefit (SMD ≥ 0.2). The confidence interval does not cross the null.
bm. Risk of bias: We downgraded once because some of the weight (<50%) comes from unclear (i.e., some concerns) risk of bias trials.
bn. One trial was not included in the meta-analysis because it reported a within-group change score (Huang 2019: 46 participants total; rated as overall low risk of bias). No significant difference between groups for mean change from baseline on any of the subscales.
bo. Cho 2013: Participants had an unknown presence of leg pain, and received acupuncture type TCM with manual stimulation. The trial did not stratify results based on gender, age, or race/ethnicity.
bp. Inconsistency: We downgraded twice. The point estimates varied with little overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 74%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
bq. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The confidence interval crosses the null.
br. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (+0.2), but the lower boundary does not cross the threshold for what may be considered appreciable harm (-0.2).
bs. One trial was not included in the meta-analysis due to missing data (Cherkin 2009: 638 participants total, rated as overall unclear risk of bias). Clinically unimportant (MD<10, scale 0-100) but statistically significant difference between groups for mean change in PCS and MCS (p<0.001) favouring acupuncture.
bt. Imprecision: We did not downgrade. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The upper and lower boundaries of the 95% CI do not cross the threshold for what may be considered appreciable benefit (+0.2) or harm (-0.2).
bu. Imprecision: We downgraded twice. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable harm (-0.2), and the upper boundary crosses the threshold for what may be considered appreciable benefit (+0.2).

bv. Inconsistency: We downgraded twice. The point estimates differed with little overlap in confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 70%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
bw. Three trials were not included in the meta-analysis due to missing data. Cho 2013 (ID#: 2002): 130 participants total, rated as overall unclear risk of bias. Authors reported no serious events; 10 minor to moderate adverse events in acupuncture group (none persisted more than 1 week): pain; bruising at acupuncture site; pain, numbness or other bothersomeness in leg; shoulder pain. Haake 2007 (ID#: 2003): 774 participants total, rated as overall low risk of bias. Authors reported 476 clinically relevant adverse effects by 257 patients (22.6%) with no significant difference between groups. Molsberger 2002 (ID#: 2007): 186 participants total, rated as overall high risk of bias. Authors reported no important adverse events or side effects were observed in any group.
bx. Minor adverse events: Brinkhaus 2006: hematoma, bleeding in both groups. Cherkin 2009: mostly short-term pain with individualized or standardized acupuncture (1 participant reported pain lasting 1 month). Huang 2019: subcutaneous hematoma after acupuncture. Kong 2020: minor pain, bruising, skin rash, and slight bleeding at needle site; mild reaction to prone position included nausea, dizziness, and mild back ache in both groups. Koppenhaver 2021: pain during treatment, dizziness, unspecified emotional change. Yuan 2016: transient worsening back pain, acupuncture point bruise, back and leg numbness and discomfort, shoulder pain (up to 1 week) in both groups.
by. Inconsistency: We downgraded twice. The point estimates vary with little overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 63%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
bz. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (OR ≥ 1.10). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (0.90).
ca. Minor adverse events: Huang 2019: subcutaneous hematoma after needling.
cb. Risk of bias: We did not downgrade because all of the weight comes from low risk of bias trials.
cc. Minor adverse events: Cherkin 2009: mostly short-term pain with individualized or standardized acupuncture (1 participant reported pain lasting 1 month).
cd. Molsberger 2002 (ID#: 2007) was not included in meta-analysis due to missing data, 186 participants total, rated as overall high risk of bias. Authors reported no important adverse events or side effects were observed in any group.
ce. Minor adverse events: Brinkhaus 2006: hematoma, bleeding in both groups. Kong 2020: minor pain, bruising, skin rash, and slight bleeding at needle site; mild reaction to prone position included nausea, dizziness, and mild back ache in both groups. Koppenhaver 2021: pain during treatment, dizziness, unspecified emotional change. Yuan 2016: transient worsening back pain, acupuncture point bruise, back and leg numbness and discomfort, shoulder pain (up to 1 week) in both groups.
cf. Minor adverse events: Cherkin 2009: mostly short-term pain with individualized or standardized acupuncture (1 participant reported pain lasting 1 month). Huang 2019: subcutaneous hematoma after acupuncture. Yuan 2016: transient worsening back pain, acupuncture point bruise, back and leg numbness and discomfort, shoulder pain (up to 1 week) in both groups.
cg. Inconsistency: We downgraded once. There is some similarity between point estimates and overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 57%). This could not be explained due to small subgroups and may represent moderate heterogeneity.
ch. Minor adverse events: Koppenhaver 2021: pain during treatment, dizziness, unspecified emotional change.
ci. Minor adverse events: Brinkhaus 2006: hematoma, bleeding in both groups. Cherkin 2009: mostly short-term pain with individualized or standardized acupuncture (1 participant reported pain lasting 1 month). Kong 2020: minor pain, bruising, skin rash, and slight bleeding at needle site; mild reaction to prone position included nausea, dizziness, and mild back ache in both groups.
cj. Inconsistency: We downgraded twice. The point estimates are in different directions with no overlap in confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 89%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
ck. Two studies were not included in the meta-analysis due to missing data. Cho 2013 (ID#: 2002): 130 participants total, rated as overall unclear risk of bias, authors reported no serious events; 10 minor to moderate adverse events in acupuncture group (none persisted more than 1 week) including pain, bruising at acupuncture site. Molsberger 2002 (ID#: 2007): 186 participant total, rated as overall high risk of bias, authors reported no important adverse events or side effects were observed in any group.
cl. Minor adverse events: Brinkhaus 2006: hematoma, bleeding in both groups. Koppenhaver 2021: pain during treatment, dizziness, unspecified emotional change. Yuan 2016: transient worsening back pain, acupuncture point bruise, back and leg numbness and discomfort, shoulder pain (up to 1 week) in both groups.
cm. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (OR ≥ 0.90). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable harm (1.10), but the lower boundary does not cross the threshold for what may be considered appreciable harm (0.90).
cn. Minor adverse events: Kong 2020: minor pain, bruising, skin rash, and slight bleeding at needle site; mild reaction to prone position included nausea, dizziness, and mild back ache in both groups.
co. One trial was not included in the meta-analysis due to missing data. Haake 2007 (ID#: 2003): 774 participants total, rated as overall low risk of bias; authors reported 476 clinically relevant adverse effects by 257 patients (22.6%) with no significant difference between groups.

References

1.Yuan Q, Liu L,Ma J, Wu W, Ye M, Zhang Y. [A clinical study of acupuncture therapy for treatment of chronic nonspecific low back pain] . 2016.
2.Yu S, Ortiz A, Gollub RL et al. Acupuncture treatment modulates the connectivity of key regions of the descending pain modulation and reward systems in patients with chronic low back pain. 2020.

3.Molsberger AF, Mau J, Pawelec DB, Winkler J. Does acupuncture improve the orthopedic management of chronic low back pain--a randomized, blinded, controlled trial with 3 months follow up. 2002.
4.Martín-Corrales C, Bautista IV, Méndez-Mera JE et al. Benefits of Adding Gluteal Dry Needling to a Four-Week Physical Exercise Program in a Chronic Low Back Pain Population. A Randomized Clinical Trial. 2020.
5.Koppenhaver SL, Weaver AM, Randall TL et al. Effect of dry needling on lumbar muscle stiffness in patients with low back pain: a double blind, randomized controlled trial using shear wave elastography. . 2021.
6.Kerr DP, Walsh DM, Baxter D. Acupuncture in the management of chronic low back pain: a blinded randomized controlled trial. 2003.
7.Haake M, Muller HH, Schade-Brittinger C,et al.. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. 2007.
8.Huang Z, Liu S, Zhou J, Yao Q, Liu Z. Efficacy and Safety of Acupuncture for Chronic Discogenic Sciatica, a Randomized Controlled Sham Acupuncture Trial. 2019.
9.Brinkhaus B, Witt CM, Jena S,et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. 2006.
10.Cherkin DC, Sherman KJ, Avins AL et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. 2009.
11.Cho YJ, Song YK, Cha YY et al. Acupuncture for chronic low back pain: a multicenter, randomized, patient-assessor blind, sham-controlled clinical trial. 2013.
12.de Castro Moura C, de Cássia Lopes Chaves E,Couto Machado Chianca T,Ruginsk SG,Alves Nogueira D,Iunes DH. Effects of auricular acupuncture on chronic pain in adults with back musculoskeletal disorders: a randomized clinical trial. 2019.
13.Kim H, Mawla I,Lee J,et al. Reduced tactile acuity in chronic low back pain is linked with structural neuroplasticity in primary somatosensory cortex and is modulated by acupuncture therapy. 2020.
14.Kong JT, Puetz C, Tian L et al. Effect of Electroacupuncture vs Sham Treatment on Change in Pain Severity Among Adults With Chronic Low Back Pain: a Randomized Clinical Trial. 2020.
15. Ushinohama A et al. Effect of a single session of ear acupuncture on pain intensity and postural control in individuals with chronic low back pain: a randomized controlled trial. 2016.

GRADE Table 2: ***What are the benefits and harms of acupuncture in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to no intervention or interventions where the effect of acupuncture could be isolated?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
21[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,a,b]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	859	858	-	MD **1.21 lower** (1.5 lower to 0.92 lower)	⨁⨁◯◯ Low	CRITICAL
Pain (mixed females and males) (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
19[1,2,3,4,6,7,8,9,10,11,12,13,14,15,17,18,19,20,21,b]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	800	799	-	MD **1.22 lower** (1.48 lower to 0.97 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in males (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[16,a]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **1.99 lower** (2.86 lower to 1.12 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults (gender not reported) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[5]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	19	19	-	MD **0.3 higher** (0.1 higher to 0.5 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 2 weeks; assessed with: VAS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
8[1,2,3,4,10,16,20,21,a]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	272	271	-	MD **1.83 lower** (2.76 lower to 0.91 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults with radicular leg pain (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
6[6,12,13,15,17,18]	randomized trials	very serious[c]	not serious[d]	not serious[i]	not serious[k]	none	257	257	-	MD **0.75 lower** (0.95 lower to 0.55 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults either with or without leg pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
3[7,11,14]	randomized trials	very serious[c]	not serious[d]	not serious[e]	serious[l]	none	181	181	-	MD **1.32 lower** (1.49 lower to 1.16 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
4[5,8,9,19,b]	randomized trials	very serious[c]	serious[m]	not serious[e]	serious[l]	none	149	149	-	MD **0.68 lower** (1.44 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
18[1,2,3,4,6,7,8,9,10,12,13,14,15,17,18,19,20,21,b]	randomized trials	very serious[c]	not serious[d]	not serious[j]	not serious[f]	none	785	784	-	MD **1.2 lower** (1.46 lower to 0.94 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
3[5,11,16,a]	randomized trials	very serious[c]	serious[n]	not serious[o]	very serious[i]	none	74	74	-	MD **1.38 lower** (3.02 lower to 0.26 higher)	⨁◯◯◯ Very low	CRITICAL
Pain stratified by race/ethnicity (follow-up: closest to 2 weeks)												
0												
Pain in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
19[1,2,3,4,6,7,8,9,10,12,13,14,15,16,17,18,19,20,21,a,b]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	825	824	-	MD **1.24 lower** (1.49 lower to 0.99 lower)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Pain in adults treated with acupuncture type myofascial (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[11]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	15	15	-	MD **2.17 lower** (3.49 lower to 0.85 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture (type not reported) (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[5]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	19	19	-	MD **0.3 higher** (0.1 higher to 0.5 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
8[2,6,8,9,13,17,20,21]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	362	363	-	MD **1.38 lower** (1.84 lower to 0.92 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults treated with acupuncture with electrical stimulation (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
5[1,4,5,14,16,a]	randomized trials	very serious[c]	not serious[d]	not serious[e]	serious[l]	none	125	124	-	MD **1.21 lower** (2.22 lower to 0.21 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with heat stimulation (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[12]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	46	45	-	MD **1.23 lower** (1.6 lower to 0.86 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with mixed stimulation methods (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
4[7,15,18,19]	randomized trials	very serious[c]	not serious[d]	not serious[i]	not serious[f]	none	257	257	-	MD **1.11 lower** (1.43 lower to 0.79 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults treated with acupuncture without stimulation (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
2[3,11,q]	randomized trials	very serious[c]	not serious[d]	not serious[e]	very serious[i]	none	50	50	-	MD **1.28 lower** (2.69 lower to 0.13 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with threading stimulation (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[10,r]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	MD **0.78 lower** (2.16 lower to 0.6 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults after removing high risk of bias studies (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[10,20]	randomized trials	very serious[c]	not serious[d]	not serious[i]	very serious[i]	none	69	69	-	MD **1.79 lower** (3.59 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 3 months; assessed with: VAS, NRS, BPI, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
9[1,4,13,14,16,20,21,22,23,a,s]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	420	342	-	MD **1.56 lower** (2.18 lower to 0.95 lower)	⨁⨁◯◯ Low	CRITICAL
Pain (mixed females and males) (follow-up: closest to 3 months; assessed with: VAS, NRS, BPI, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
8[1,4,13,14,20,21,22,23,s]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	380	302	-	MD **1.57 lower** (2.28 lower to 0.86 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in males (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[16,a]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **1.54 lower** (2.48 lower to 0.61 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in females (follow-up: closest to 3 months)												
0												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Pain stratified by race/ethnicity (follow-up: closest to 3 months)												
0												
Pain in adults with radicular leg pain (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[13]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	40	40	-	MD **0.61 lower** (0.91 lower to 0.31 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults without leg pain (follow-up: closest to 3 months; assessed with: VAS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
6[1,4,16,20,21,23,a]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	317	239	-	MD **1.89 lower** (2.55 lower to 1.22 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults either with or without leg pain (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1[14]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	26	26	-	MD **1.81 lower** (3.03 lower to 0.59 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: BPI; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[22,s]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	37	37	-	MD **0.05 higher** (1.4 lower to 1.5 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in adults in high to upper-middle income countries (follow-up: closest to 3 months; assessed with: VAS, NRS, BPI, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
8[1,4,13,14,20,21,22,23,s]	randomized trials	very serious[c]	not serious[d]	not serious[j]	not serious[f]	none	380	302	-	MD **1.57 lower** (2.28 lower to 0.86 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults in low- or lower middle-income countries (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[16]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	MD **1.54 lower** (2.48 lower to 0.61 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: VAS, NRS, BPI, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
8[1,4,13,14,16,20,21,22,a,s]	randomized trials	very serious[c]	not serious[d]	not serious[e]	not serious[f]	none	280	268	-	MD **1.45 lower** (2.07 lower to 0.83 lower)	⨁⨁◯◯ Low	CRITICAL
Pain in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[23]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	140	74	-	MD **2.41 lower** (3.15 lower to 1.67 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
4[13,20,21,23]	randomized trials	very serious[c]	not serious[d]	not serious[e]	serious[t]	none	277	200	-	MD **1.69 lower** (2.9 lower to 0.48 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture with electrical stimulation (follow-up: closest to 3 months; assessed with: VAS, NRS, Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
4[1,4,14,16,a]	randomized trials	very serious[c]	not serious[d]	not serious[e]	serious[l]	none	106	105	-	MD **1.65 lower** (2.29 lower to 1.02 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in adults treated with acupuncture (no stimulation) (follow-up: closest to 3 months; assessed with: BPI; benefit indicated by lower values; scale: 0 to 10)												
1[22,s,u]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	37	37	-	MD **0.05 higher** (1.4 lower to 1.5 higher)	⨁◯◯◯ Very low	CRITICAL
Pain after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[20]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	50	50	-	MD **0.92 lower** (1.89 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA, Aberdeen; benefit indicated by lower values)												
19[1,2,3,4,5,6,7,8,9,10,11,12,13,14,16,17,18,19,20,a,v]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	770	771	-	SMD **1.39 lower** (2 lower to 0.77 lower)	⨁⨁◯◯ Low	CRITICAL
Function (mixed females and males) (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA, Aberdeen; benefit indicated by lower values)												
17[1,2,3,4,6,7,8,9,10,11,12,13,14,17,18,19,20,v]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	711	712	-	SMD **1.66 lower** (2.29 lower to 1.04 lower)	⨁⨁◯◯ Low	CRITICAL
Function in males (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
1[16,a]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	40	40	-	SMD **1.01 lower** (1.48 lower to 0.55 lower)	⨁◯◯◯ Very low	CRITICAL
Function (gender not reported) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[5]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	19	19	-	SMD **2.93 higher** (1.98 higher to 3.87 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults with radicular leg pain (follow-up: closest to 2 weeks; assessed with: ODI, JOA; benefit indicated by lower values)												
5[6,12,13,17,18]	randomized trials	very serious[c]	not serious[w]	not serious[i]	not serious[f]	none	226	228	-	SMD **2.03 lower** (3.05 lower to 1 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults either with or without leg pain (follow-up: closest to 2 weeks; assessed with: ODI, Aberdeen; benefit indicated by lower values)												
3[7,11,14]	randomized trials	very serious[c]	serious[x]	not serious[e]	very serious[y]	none	181	181	-	SMD **1.99 lower** (4.9 lower to 0.92 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults without leg pain (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA; benefit indicated by lower values)												
7[1,2,3,4,10,16,20,a]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	214	213	-	SMD **1.02 lower** (1.42 lower to 0.61 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
4[5,8,9,18,v]	randomized trials	very serious[c]	serious[z]	not serious[e]	very serious[y]	none	149	149	-	SMD **0.8 lower** (2.74 lower to 1.15 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA, Aberdeen; benefit indicated by lower values)												
16[1,2,3,4,6,7,8,9,10,12,13,14,17,18,19,20,v]	randomized trials	very serious[c]	not serious[w]	not serious[i]	not serious[f]	none	696	697	-	SMD **1.75 lower** (2.39 lower to 1.1 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												
3[5,11,16,a]	randomized trials	very serious[c]	serious[aa]	not serious[o]	very serious[i]	none	74	74	-	SMD **0.11 higher** (1.44 lower to 1.67 higher)	⨁◯◯◯ Very low	CRITICAL
Function stratified by race/ethnicity (follow-up: closest to 2 weeks)												
0												CRITICAL
Function in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA, Aberdeen; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
17[1,2,3,4,6,7,8,9,10,12,13,14,16,17,18,19,20,a,v]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	736	737	-	SMD **1.67 lower** (2.26 lower to 1.08 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults treated with acupuncture type myofascial (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
1[11]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	15	15	-	SMD **0.32 lower** (1.04 lower to 0.4 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture (type not reported) (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
1[5]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	19	19	-	SMD **2.93 higher** (1.98 higher to 3.87 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with manual stimulation (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, JOA; benefit indicated by lower values)												
7[2,6,8,9,13,17,20]	randomized trials	very serious[c]	not serious[w]	not serious[j]	not serious[f]	none	304	305	-	SMD **1.14 lower** (1.57 lower to 0.71 lower)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Function in adults treated with acupuncture with electrical stimulation (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI, Aberdeen; benefit indicated by lower values)												
5[1,4,5,14,16]	randomized trials	very serious[c]	serious[ab]	not serious[e]	very serious[y]	none	125	124	-	SMD **0.38 lower** (1.35 lower to 0.59 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with heat stimulation (follow-up: closest to 2 weeks; assessed with: JOA; benefit indicated by lower values)												
1[12]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	45	46	-	SMD **3.44 lower** (4.1 lower to 2.79 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with mixed stimulation methods (follow-up: closest to 2 weeks; assessed with: ODI, JOA; benefit indicated by lower values)												
3[7,18,19]	randomized trials	very serious[c]	not serious[w]	not serious[j]	not serious[f]	none	227	227	-	SMD **3.73 lower** (4.84 lower to 2.62 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults treated with acupuncture without stimulation (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[3,11,v]	randomized trials	very serious[c]	serious[ac]	not serious[e]	very serious[i]	none	50	50	-	SMD **1.32 lower** (3.27 lower to 0.62 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with threading stimulation (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
1[10]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	SMD **0.15 lower** (0.79 lower to 0.49 higher)	⨁◯◯◯ Very low	CRITICAL
Function after removing high risk of bias studies (follow-up: closest to 2 weeks; assessed with: RMDQ, ODI; benefit indicated by lower values)												
2[10,20]	randomized trials	very serious[c]	serious[ad]	not serious[i]	very serious[i]	none	69	69	-	SMD **0.59 lower** (1.36 lower to 0.19 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 3 months; assessed with: RMDQ, ODI, JOA, BPI, Hannover, Aberdeen; benefit indicated by lower values)												
8[1,4,13,14,16,20,22,23,ae,af]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	287	352	-	SMD **0.57 lower** (0.92 lower to 0.22 lower)	⨁⨁◯◯ Low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Function (mixed females and males) (follow-up: closest to 3 months; assessed with: RMDQ, ODI, JOA, BPI, Hannover, Aberdeen; benefit indicated by lower values)												
7[1,4,13,14,20,22,23,ae,af]	randomized trials	very serious[c]	not serious[w]	not serious[i]	not serious[f]	none	267	332	-	SMD **0.56 lower** (0.95 lower to 0.17 lower)	⨁⨁◯◯ Low	CRITICAL
Function in males (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
1[16]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	20	20	-	SMD **0.67 lower** (1.31 lower to 0.04 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults with radicular leg pain (follow-up: closest to 3 months; assessed with: JOA; benefit indicated by lower values)												
1[13]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	40	40	-	SMD **1.05 lower** (1.52 lower to 0.58 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults either with or without leg pain (follow-up: closest to 3 months; assessed with: Aberdeen; benefit indicated by lower values)												
1[14]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	26	26	-	SMD **0.5 lower** (1.05 lower to 0.05 higher)	⨁◯◯◯ Very low	CRITICAL CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Function in adults without leg pain (follow-up: closest to 3 months; assessed with: RMDQ, ODI, Hannover; benefit indicated by lower values)												
5[1,4,16,20,23,af]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	184	249	-	SMD **0.65 lower** (0.95 lower to 0.34 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults with unclassified presence of leg pain (follow-up: closest to 3 months; assessed with: BPI; benefit indicated by lower values)												
1[22,ae]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	37	37	-	SMD **0.43 higher** (0.03 lower to 0.89 higher)	⨁◯◯◯ Very low	CRITICAL
Function in adults in high to upper-middle income countries (follow-up: closest to 3 months; assessed with: RMDQ, ODI, JOA, BPI, Hannover, Aberdeen; benefit indicated by lower values)												
7[1,4,13,14,20,22,23,ae,af]	randomized trials	very serious[c]	not serious[w]	not serious[i]	not serious[f]	none	267	332	-	SMD **0.56 lower** (0.95 lower to 0.17 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults in low- or lower middle-income countries (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[16]	randomized trials	very serious[c]	not serious[g]	serious[h]	very serious[i]	none	20	20	-	SMD **0.67 lower** (1.31 lower to 0.04 lower)	⨁◯◯◯ Very low	CRITICAL
Function stratified by race/ethnicity (follow-up: closest to 3 months)												
0												
Function in adults treated with acupuncture type TCM (follow-up: closest to 3 months; assessed with: RMDQ, ODI, JOA, BPI, Hannover, Aberdeen; benefit indicated by lower values)												
7[1,4,13,14,16,20,22,ae,af]	randomized trials	very serious[c]	not serious[w]	not serious[e]	not serious[f]	none	213	212	-	SMD **0.6 lower** (1.04 lower to 0.15 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults treated with acupuncture type mixed (TCM, myofascial) (follow-up: closest to 3 months; assessed with: Hannover; benefit indicated by lower values)												
1[23]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	74	140	-	SMD **0.48 lower** (0.77 lower to 0.2 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture with manual stimulation (follow-up: closest to 3 months; assessed with: ODI, JOA, Hannover; benefit indicated by lower values)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
3[13,20,23]	randomized trials	very serious[c]	not serious[w]	not serious[i]	not serious[f]	none	164	230	-	SMD **0.58 lower** (0.97 lower to 0.2 lower)	⨁⨁◯◯ Low	CRITICAL
Function in adults treated with acupuncture with electrical stimulation (follow-up: closest to 3 months; assessed with: RMDQ, Aberdeen; benefit indicated by lower values)												
4[1,4,14,16]	randomized trials	very serious[c]	not serious[w]	not serious[e]	very serious[i]	none	86	85	-	SMD **0.82 lower** (1.15 lower to 0.49 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults treated with acupuncture without stimulation (follow-up: closest to 3 months; assessed with: BPI; benefit indicated by lower values)												
1[22,ae,ag]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	37	37	-	SMD **0.43 higher** (0.03 lower to 0.89 higher)	⨁◯◯◯ Very low	CRITICAL
Function after removing high risk of bias studies (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values)												
1[20]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	50	50	-	SMD **0.3 lower** (0.69 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values; scale: 0 to 100)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[23,ah]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	74	140	-	MD **8.3 lower** (13.93 lower to 2.67 lower)	⨁◯◯◯ Very low	CRITICAL
Function stratified by gender (follow-up: closest to 6 months)												
0												
Function in adults without leg pain (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values; scale: 0 to 100)												
1[23]	randomized trials	very serious[c]	not serious[g]	not serious	serious[l]	none	74	140	-	MD **8.3 lower** (13.93 lower to 2.67 lower)	⨁◯◯◯ Very low	CRITICAL
Function in adults in high to upper-middle income countries (follow-up: closest to 6 months; assessed with: Hannover; benefit indicated by lower values; scale: 0 to 100)												
1[23]	randomized trials	very serious[c]	not serious[g]	not serious	serious[l]	none	74	140	-	MD **8.3 lower** (13.93 lower to 2.67 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by race/ethnicity, after removing high risk of bias studied or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life (follow-up: closest to 2 weeks; assessed with: EQ-5D; benefit indicated by higher values; scale: 0 to 1)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
1[10]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	MD **0.02 higher** (0.09 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults without leg pain (follow-up: closest to 2 weeks; assessed with: EQ-5D; benefit indicated by higher values; scale: 0 to 1)												
1[10]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	MD **0.02 higher** (0.09 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in adults in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: EQ-5D; benefit indicated by higher values; scale: 0 to 1)												
1[10]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	MD **0.02 higher** (0.09 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Health-related quality of life in adults treated with acupuncture type TCM (follow-up: closest to 2 weeks; assessed with: EQ-5D; benefit indicated by higher values; scale: 0 to 1)												
1[10,ai]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	19	19	-	MD **0.02 higher** (0.09 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life (follow-up: closest to 3 months; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1[23,ah,aj]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	140	74	-	MD **6.6 higher** (3.9 higher to 9.3 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (follow-up: closest to 3 months; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
1[23,ah,ak]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	140	74	-	MD **1.2 higher** (1.86 lower to 4.26 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity, in adults in low- or lower middle-income countries or after removing high risk of bias studies not identified												
0												
Depression (follow-up: closest to 3 months; assessed with: General Depression Scale; benefit indicated by lower values; scale: 0 to 61)												
1[23,ah]	randomized trials	very serious[c]	not serious[g]	serious[p]	serious[l]	none	140	74	-	MD **0.8 lower** (3.6 lower to 2 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on depression stratified by gender, race/ethnicity, in adults in low- or lower middle-income countries, after removing high risk of bias studies and in adults with leg pain not identified												
0												
Trial on other psychological functioning or social participation not identified												
0												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
Adverse events/harms during intervention period (acupuncture type TCM)												
3[20,24,25,al,am]	randomized trials	very serious[c]	serious[an]	not serious[i]	very serious[ao]	none	11/113 (9.7%)	2/110 (1.8%)	**OR 3.12** (0.42 to 23.44)	**36 more per 1,000** (from 10 fewer to 285 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults without leg pain during intervention period												
2[20,24,al,ap]	randomized trials	very serious[c]	not serious[aq]	not serious[i]	very serious[ao]	none	9/90 (10.0%)	0/90 (0.0%)	**OR 8.77** (1.02 to 75.35)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults with unclassified presence of leg pain during intervention period												
1[25,ar]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[ao]	none	2/23 (8.7%)	2/20 (10.0%)	**OR 0.86** (0.11 to 6.72)	**13 fewer per 1,000** (from 88 fewer to 327 more)	⨁◯◯◯ Very low	CRITICAL
Trials on adverse events/harms stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries during intervention period not identified												
0												
Adverse events/harms in adults treated with acupuncture with manual stimulation during intervention period												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[20,25,al,as]	randomized trials	very serious[c]	serious[at]	not serious[i]	very serious[ao]	none	10/73 (13.7%)	2/70 (2.9%)	**OR 3.59** (0.14 to 94.80)	**67 more per 1,000** (from 24 fewer to 707 more)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms in adults treated with acupuncture (stimulation not reported) during intervention period												
1[24,au]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[ao]	none	1/40 (2.5%)	0/40 (0.0%)	**OR 3.08** (0.12 to 77.80)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
Adverse events/harms after removing high risk of bias studies during intervention period												
1[20,av]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[ao]	none	8/50 (16.0%)	0/50 (0.0%)	**OR 20.20** (1.13 to 360.28)	**0 fewer per 1,000** (from 0 fewer to 0 fewer)	⨁◯◯◯ Very low	CRITICAL
OLDER ADULTS (aged 60 years or more)												
Pain (follow-up: closest to 2 weeks; assessed with: Pain Scale; benefit indicated by lower values; scale: 0 to 10)												
1[4,aw,ax]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	24	23	-	MD **0.9 lower** (1.53 lower to 0.27 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 3 months; assessed with: Pain Scale; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	No treatment	Relative (95% CI)	Absolute (95% CI)		
14[,aw,ax]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	24	23	-	MD **1.1 lower** (1.62 lower to 0.58 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Function (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
14[,ax]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	24	23	-	SMD **1.1 lower** (1.71 lower to 0.48 lower)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values)												
14[,ax]	randomized trials	very serious[c]	not serious[g]	serious[p]	very serious[i]	none	24	23	-	SMD **1.04 lower** (1.66 lower to 0.43 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on health-related quality of life, adverse events/harms, psychological functioning, change in use of medications or falls not identified												
0												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

BPI: Brief Pain Inventory; **CI:** confidence interval; **EQ-5D:** EuroQol 5 Dimensions; **JOA:** Japanese Orthopedic Association; **MD:** mean difference; **MCS:** Mental Component Summary; **OIS:** Optimal Information Size; **OR:** odds ratio; **NRS:** numerical rating scale; **ODI:** Oswestry Disability Index; **PCS:** Physical Component Summary; **RMDQ:** Roland Morris Disability Questionnaire; **SF-36:** Short Form Health Survey – 36-item; **SMD:** standardized mean difference; **TCM:** Traditional Chinese Medicine; **VAS:** Visual Analogue Scale

The following was used to guide the ratings.

Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).

Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.

Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.

Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).

Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations

a. Zaringhalam 2010 assessed two comparisons (there were 2 comparison groups). Both comparisons included in meta-analysis.
b. Two trials were not included in the meta-analysis because they reported within-group change scores. De Castro Moura 2019 (ID#: 32): 111 participants total, rated as overall high risk of bias. Clinically important (MD≥1, scale 0 to 10) and statistically significant within group mean difference for Chinese auricular acupuncture group: 1.38 (95% CI 0.43; 2.33); no significant within group changes for French auricular acupuncture or comparison group; no statistical comparison between groups. Weiß 2013 (ID#: 1153): 160 participants total, rated as overall high risk of bias. No significant difference between groups in the proportion of participants experiencing improvement in pain while sitting/standing or walking.
c. Risk of bias: We downgraded twice because all of the weight comes from high or unclear (i.e., some concerns) overall risk of bias trials.
d. Inconsistency: We did not downgrade. All or most trials are in the same direction, showing a reduction in pain.
e. Indirectness: We did not downgrade because the trials were conducted in different countries (high to low-income).
f. Imprecision: We did not downgrade. The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 1 or SMD ≥ 0.2). The confidence interval does not cross the null.
g. Inconsistency: We did not downgrade; however, there are no other trials with which to compare findings.
h. Indirectness: We downgraded once; trial(s) conducted in one country (low or lower-middle income).
i. Imprecision: We downgraded twice. The sample size is small (OIS would not have been achieved).
j. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).
k. Imprecision: We did not downgrade. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The confidence interval does not cross the null.
l. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved).
m. Inconsistency: We downgraded once. Most trials are in the same direction with similar point estimates. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 97%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
n. Inconsistency: We downgraded once. Most of the trials are in the same direction showing a reduction in pain. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 92%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
o. Indirectness: We did not downgrade because the trials were conducted in different countries (low or lower-middle income).

p. Indirectness: We downgraded once; trial(s) conducted in one country (high or upper-middle income).
q. One trial was not included in the meta-analysis because it reported within-group change scores. De Castro Moura 2019 (ID#: 32): 111 participants total; rated as overall high risk of bias. Clinically important (MD≥1, scale 0 to 10) and statistically significant within group mean difference for Chinese auricular acupuncture group: 1.38 (95% CI 0.43; 2.33); no significant within group changes for French auricular acupuncture or comparison group; no statistical comparison between groups.
r. One trial was not included in the meta-analysis because it reported within-group change scores. Weiß 2013 (ID#: 1153): 160 participants total, rated as overall high risk of bias. No significant difference between groups in the proportion of participants experiencing improvement in pain while sitting/standing or walking.
s. Two trials were not included in the meta-analysis because they reported within-group change scores. De Castro Moura 2019 (ID#: 32): 111 participants total, rated as overall high risk of bias. No significant within group changes acupuncture groups or comparison group; no statistical comparison between groups. Weiß 2013 (ID#: 1153): 160 participants total, rated as overall high risk of bias. Statistically significant difference between proportion of participants experiencing improvement in pain while sitting/standing ($p<0.01$) but not in pain while walking.
t. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-1).
u. Use of stimulation was not reported in Weiß 2013 (ID#: 1153).
v. One trial was not included in the meta-analysis because it reported within-group change scores. De Castro Moura 2019 (ID#: 32): 111 participants total; rated as overall high risk of bias. Clinically unimportant (MD<2.4, scale 0 to 24) but statistically significant within group mean difference for Chinese auricular acupuncture group: 1.56 (95% CI 0.10; 3.02); no significant within group changes for French auricular acupuncture or comparison group; no statistical comparison between groups.
w. Inconsistency: We did not downgrade. All or most trials are in the same direction, showing a reduction in functional limitation.
x. Inconsistency: We downgraded once. The results are in the same direction. One point estimate is much larger in magnitude; confidence intervals of the other studies do not overlap with it. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 99%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
y. Imprecision: We downgraded twice. The point estimate reached the pre-specified threshold for what may be considered clinically important (SMD ≥ 0.2). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-0.2), and the upper boundary crosses the threshold for what may be considered appreciable harm (+0.2).
z. Inconsistency: We downgraded once. The point estimates differ with little overlap in confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 98%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
aa. Inconsistency: We downgraded once. Most of the point estimates are in the same direction. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 94%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
ab. Inconsistency: We downgraded once. Most of the trials are in the same direction showing a reduction in functional limitation. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 92%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
ac. Inconsistency: We downgraded once. The point estimates are in the same direction. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 94%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
ad. Inconsistency: We downgraded once. The point estimates are in the same direction with little overlap between confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 76%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
ae. One trial was not included in the meta-analysis because it reported within-group change scores. De Castro Moura 2019 (ID#: 32): 111 participants total; rated as overall high risk of bias. No significant within group changes for acupuncture groups or comparison group; no statistical comparison between groups.
af. One trial was not included in the meta-analysis because it reported within-group change scores. Witt 2006 (ID#: 2010): 3093 participants total; rated as overall high risk of bias. Statistically significant difference between groups for mean percent disability reduction (scale 0 to 100) (22.0; 95% CI 19.3, 24.7; $p<0.001$) favouring acupuncture.
ag. Use of stimulation was not reported in Witt 2006 (ID#: 2010).
ah. Brinkhaus 2006: participants had no leg pain; in high to upper-middle income country; were treated with mixed acupuncture type (TCM, dry needling) with manual stimulation.
ai. Sung 2020: acupuncture with threading stimulation; rated as overall unclear risk of bias.
aj. One trial was not included in the meta-analysis because it reported within-group change scores. Witt 2006 (ID#: 2010): 3093 participants total; rated as overall high risk of bias. clinically unimportant (PCS: MD <10, scale 0-100) but statistically significant difference between groups for mean point increase in quality of life (4.7; 95% CI 4.0, 5.4; $p<0.001$) favouring acupuncture.
ak. One trial was not included in the meta-analysis because it reported within-group change scores. Witt 2006 (ID#: 2010): 3093 participants total; rated as overall high risk of bias. Clinically unimportant (MCS: MD<10, scale 0-100) but statistically significant different between groups for mean point increase in quality of life (2.1; 95% CI 1.4, 2.8; $p<0.001$) favouring acupuncture.
al. One trial was not included in meta-analysis due to missing data. Molsberger 2002 (ID#: 2007): 186 participants total, rated as overall high risk of bias. Authors reported no important adverse events or side effects were observed in any group.
am. Minor adverse events: Kerr 2003: increased tenderness, leg pain for a few days following treatment. Ushinohama 2016: dizziness in one participant (unknown treatment group allocation). Yuan 2016: transient (up to 1 week) worsening back pain, acupuncture point pain and bruising, back and leg numbness and discomfort, shoulder pain, foot pain.

an. Inconsistency: We downgraded once. The point estimates vary and have overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 41%). This could not be explained due to small subgroups and may represent moderate heterogeneity.
ao. Imprecision: We downgraded twice due to small sample size and number of events.
ap. Minor adverse events: Ushinohama 2016: dizziness in one participant (unknown treatment group allocation). Yuan 2016: transient (up to 1 week) worsening back pain, acupuncture point pain and bruising, back and leg numbness and discomfort, shoulder pain, foot pain.
aq. Inconsistency: We did not downgrade. The point estimates are in the same direction with overlapping confidence intervals. Statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
ar. Minor adverse events: Kerr 2003: increased tenderness, leg pain for a few days following treatment.
as. Minor adverse events: Kerr 2003: increased tenderness, leg pain for a few days following treatment. Yuan 2016: transient (up to 1 week) worsening back pain, acupuncture point pain and bruising, back and leg numbness and discomfort, shoulder pain, foot pain.
at. Inconsistency: We downgraded once. The point estimates go in different directions; there is some overlap in confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 71%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
au. Minor adverse events: Ushinohama 2016: dizziness in one participant (unknown treatment group allocation).
av. Minor adverse events: Yuan 2016: transient (up to 1 week) worsening back pain, acupuncture point pain and bruising, back and leg numbness and discomfort, shoulder pain, foot pain.
aw. Meng 2003: Pain Scale range not specified (assumed 0-10).
ax. Meng 2003: Participants had no leg pain, were in a high to upper-middle income country, and were treated with acupuncture type TCM with electrical stimulation.

References

1.Depaoli Lemos VJ, Selau RC, Blos C, Baptista Dohnert M, Boff Daitx R, de Almeida Brito V. Electroacupuncture and Transcutaneous Electrical Nerve Stimulation in Chronic Nonspecific Low Back Pain: a Blind Randomized Clinical Trial. 2021.
2.Li YJ, Zhuang WS, Cai XG, Yang Y, Han MM, Zhang DW. [Effect of acupuncture at " three points of ilioumbar" on lumbar function and pain in patients with iliopsoas muscle strain]. 2019.
3.Li SK, Zhao J, Cao XW, Zhu S, Liu ZQ, Fan YX. [Core stability training combined with acupuncture in treatment of chronic nonspecific low back pain: a prospective randomized controlled trial]. 2022.
4.Meng CF, Wang D, Ngeow J, Lao L, Peterson M, Paget S. Acupuncture for chronic low back pain in older patients: a randomized, controlled trial. 2003.
5.Moslemi F, Farokhi ZS. Effects of electroacupuncture on pain, functional disability and ultrasonographic changes of gluteus maximus muscle in non-specific chronic low back pain patients with gluteus maximus muscle trigger points. 2020.
6.Pan J, Yang L, Qiu Y et al. [Clinical trial on effect of acupuncture combined with sling exercise training on lumbar disc herniation] . 2019.
7.P, Ran. [Effect of acupuncture on pain and lumbar function in patients with discogenic low back pain]. 2021.
8.Ren B, Feng X, Zhang C. [Clinical trial on the analgesic effect of acupuncture based on meridian-tendon dialectical theory in chronic low back pain]. 2021.
9.Shi F, Dong B, Lin X, Fu Y. [Clinical observation on treatment of non-specific low back pain based on meridian theory and acupuncture]. 2021.
10.Sung WS, Hong Y, Jeon SR et al. Efficacy and safety of thread embedding acupuncture combined with acupuncture for chronic low back pain: A randomized, controlled, assessor-blinded, multicenter clinical trial. 2020.
11.Tabatabaiee A, Takamjani IE, Sarrafzadeh J, Salehi R, Ahmadi M. Ultrasound-guided dry needling decreases pain in patients with piriformis syndrome. 2019.
12.Wang X, Yu X, Huang F, Luo X, Gong Y. [Effect of acupuncture combined with fire acupuncture on pain and serum inflammatory factor in patients with lumbar disc herniation]. 2020a.
13.Wang L, Wang Z, Wang Y, Guo Y. Observations on the efficacy of acupuncture for lumbar intervertebral disc herniation. 2020b.
14.Yeung CK, Leung MC, Chow DH. The use of electro-acupuncture in conjunction with exercise for the treatment of chronic low-back pain. 2003.
15.Yu XJ, Zhang L, Lu WY, Gao Q, Liu L, Wang Y. Effect of electroacupuncture combined with caudal epidural injection on functional rehabilitation of patients with lumbar hernia. 2021.
16.Zaringhalam J, Manaheji H, Rastqar A, Zaringhalam M. Reduction of chronic non-specific low back pain: a randomized controlled clinical trial on acupuncture and baclofen. 2010.
17.Zhai B, Wang C. [Therapeutic effect of drug cupping combined with deep puncture on Jiaji point on lumbar disc herniation and its effect on M-JOA score, ODI index and quality of life]. 2019.
18.Zhu X, Shen X, Kang L et al.. Clinical trial on acupuncture combined with bone setting manipulation in the treatment of lumbar disc herniation. 2020.
19.Z, Zhu. Clinical trial on warming-needle moxibustion combined with joint loosening for chronic low back pain. 2021.
20.Yuan Q, Liu L, Ma J, Wu W, Ye M, Zhang Y. [A clinical trial of acupuncture therapy for treatment of chronic nonspecific low back pain] . 2016.
21.Molsberger AF, Mau J, Pawelec DB, Winkler J. Does acupuncture improve the orthopedic management of chronic low back pain--a randomized, blinded, controlled trial with 3 months follow up. 2002.
22.de Castro Moura C, de Cássia Lopes Chaves E,Couto Machado Chianca T,Ruginsk SG,Alves Nogueira D,Iunes DH. Effects of auricular acupuncture on chronic pain in adults with back musculoskeletal disorders: a randomized clinical trial. 2019.
23.Brinkhaus B, Witt CM, Jena S et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. 2006.

24.Ushinohama A, Cunha BP, Costa LO, Barela AM, Freitas PB. Effect of a single session of ear acupuncture on pain intensity and postural control in individuals with chronic low back pain: a randomized controlled trial. 2016.
25.Kerr DP, Walsh DM, Baxter D. Acupuncture in the management of chronic low back pain: a blinded randomized controlled trial. 2003.

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 3: ***What are the benefits and harms of acupuncture in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Usual care	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 3 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
11[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[e]	none	299	148	-	MD **1.35 lower** (1.86 lower to 0.84 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 6 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
11[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[f]	none	285	145	-	MD **0.65 lower** (1.17 lower to 0.13 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 12 months; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
11[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[g]	none	288	143	-	MD **0.5 lower** (1.02 lower to 0.02 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Usual care	Relative (95% CI)	Absolute (95% CI)		
Function (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 3 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
11[1],[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[h]	none	299	148	-	MD **2.55 lower** (3.7 lower to 1.4 lower)	⨁◯◯◯ Very low	CRITICAL
Function (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 6 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
11[1],[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[i]	none	285	145	-	MD **1.65 lower** (2.83 lower to 0.47 lower)	⨁◯◯◯ Very low	CRITICAL
Function (in adults with and without leg pain, in high-income country, treated with acupuncture type TCM) (follow-up: closest to 12 months; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
11[1],[a]	randomized trials	very serious[b]	not serious[c]	not serious[d]	serious[i]	none	288	143	-	MD **1.9 lower** (3.15 lower to 0.65 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
Trials on health-related quality of life, adverse events/harms, psychological functioning and social participation not identified												
0												

OLDER ADULTS (aged 60 years or more)

Web Annex D.B2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture	Usual care	Relative (95% CI)	Absolute (95% CI)		
Trials on pain, function, health-related quality of life, adverse events/harms, psychological functioning, change in use of medications and falls not identified												
0												

CI: confidence interval; **MD:** mean difference; **NRS:** numerical rating scale; **RMDQ:** Roland Morris Disability Questionnaire; **TCM:** Traditional Chinese Medicine

The following was used to guide the ratings.

Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).

Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.

Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.

Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).

Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations

a. Cherkin 2009 had 2 comparisons (both included in meta-analysis); acupuncture stimulation not reported; rated as overall unclear risk of bias.
b. Risk of bias: We downgraded twice because all of the weight comes from high or unclear (i.e., some concerns) risk of bias studies.
c. Inconsistency: We did not downgrade; however, there are no other studies with which to compare findings.
d. Indirectness: We downgraded once because the trial was conducted in one country (high-income).
e. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-1).
f. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-1).
g. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 1). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-1), but the upper boundary does not cross the threshold for what may be considered appreciable harm (+1).
h. Imprecision: We downgraded once. The point estimate reached the pre-specified threshold for what may be considered clinically important (MD ≥ 2.4). The upper boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-2.4).

i. Imprecision: We downgraded once. The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD ≥ 2.4). The lower boundary of the 95% CI crosses the threshold for what may be considered appreciable benefit (-2.4).

References

1.Cherkin DC, Sherman KJ,Avins AL,et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. 2009.

B.3 Spinal manipulative therapy (SMT)

Overview of the PICO structure

Definition of the intervention	
Spinal manipulative therapy (SMT) is considered to be any "hands-on" treatment that involves movement of the spinal joints, including both high-velocity, low-amplitude manipulation and low-velocity, low-amplitude mobilization. Mobilization uses low-grade velocity (relative to manipulation) and small- or large-amplitude passive movement techniques within the person's spinal joint range of motion and control, while manipulation uses a high-velocity impulse or thrust applied to a synovial joint over a short amplitude at, or close to, the end of the passive or physiological range of motion, which is often accompanied by an audible "crack".	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial) d) Adjuvant therapy, i.e. where the additional effect of the intervention could be isolated

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified.

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Moderate; small; trivial; uncertain; varies	Moderate; small; trivial; uncertain
Harms	Small; trivial; uncertain	Small; trivial; uncertain

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Balance benefits to harms	Probably favours SMT; probably does not favour SMT; uncertain	Probably favours SMT; probably does not favour SMT; uncertain
Overall certainty	Very low; low	Very low
Values and preferences	Probably important uncertainty or variability; possibly important uncertainty or variability	Probably important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate costs; varies	Moderate costs; varies
Equity and human rights	No impact; probably reduced (traction especially); uncertain; varies	No impact; probably reduced (traction especially); uncertain; varies
Acceptability	Yes; probably yes; probably no; uncertain; varies	Yes; probably yes; probably no; uncertain; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

GRADE Table 1. ***What are the benefits and harms of SMT in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham SMT/placebo treatment?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain - Pain at 1 month												
15	randomized trials	serious[a]	serious[b]	not serious[c]	serious[d]	none	719	683	-	MD **6.07 lower** (13.09 lower to 0.95 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
11	randomized trials	serious[a]	serious[b]	serious[c]	serious[d]	none	670	614	-	MD **4.9 lower** (14.57 lower to 4.77 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development – low- or lower middle-income countries												
4	randomized trials	serious[e]	not serious[f]	serious[g]	very serious[h]	none	88	122	-	MD **8.25 lower** (14.62 lower to 1.88 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[r]	serious[i]	very serious[h]	none	69	67	-	MD **2.48 lower** (9.87 lower to 4.91 higher)	⨁◯◯◯ Very low	
Pain - Pain at 3 months												
8	randomized trials	serious[a]	serious[j]	not serious[c]	serious[m]	none	514	449	-	MD **0.9 lower** (4.68 lower to 2.87 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
6	randomized trials	serious[a]	serious[j]	not serious[c]	serious[s]	none	494	412	-	MD **0.78 lower** (6.00 lower to 4.43 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
2	randomized trials	serious[e]	not serious[f]	serious[g]	very serious[h]	none	58	69	-	MD 0.**49 lower** (3.83 lower to 2.84 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[r]	serious[i]	very serious[h]	none	69	66	-	MD **2.22 lower** (9.96 lower to 5.52 higher)	⨁◯◯◯ Very low	
Pain - Pain at 6 months												
2	randomized trials	serious[k]	serious[l]	serious[g]	very serious[h]	none	58	56	-	MD **0.96 higher** (6.34 lower to 8.26 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - high-income countries												
1	randomized trials	very serious[e]	serious[r]	serious[g]	very serious[h]	none	32	19	-	MD **7.1 higher** (5.16 lower to 19.36 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[m]	serious[r]	serious[g]	very serious[h]	none	26	37	-	MD **1.3 lower** (6.31 lower to 3.71 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age - not reported (no subgroup analysis performed; no trial reporting outcomes at this follow-up **Pain - Pain at 12 months**												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[m]	Serious[r]	serious[g]	very serious[h]	none	26	37	-	MD **0.2 higher** (5.33 lower to 5.73 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries - not reported (no subgroup analysis performed; no trial reporting outcomes at this follow-up)												
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[m]	Serious[r]	serious[g]	very serious[h]	none	26	37	-	MD **0.2 higher** (5.33 lower to 5.73 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age - not reported												
Back-specific functional status (higher scores mean more disability)												
Back-specific functional status - back-specific functional status at 1 month												
12	randomized trials	serious[n]	serious[b]	not serious[c]	serious[o]	none	678	642	-	SMD **0.43 lower** (0.74 lower to 0.12 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
9	randomized trials	serious[n]	serious[b]	not serious[c]	serious[o]	none	622	572	-	SMD **0.34 SD lower** (0.68 lower to 0)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
3	randomized trials	serious[e]	serious[p]	serious[g]	very serious[h]	none	56	70	-	SMD **0.79 SD lower** (1.36 lower to 0.21 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age												
1	randomized trials	serious[a]	Serious[r]	serious[i]	very serious[h]	none	69	67	-	SMD **0.07 SD lower** (0.4 lower to 0.27 higher)	⨁◯◯◯ Very low	
Population subgroup 6: ODI												
8	randomized trials	serious[n]	serious[b]	serious[c]	very serious[h]	none	214	250	-	SMD **0.65 SD lower** (1.2 lower to 0.11 lower)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
4	randomized trials	serious[a]	serious[b]	not serious[c]	very serious[h]	none	398	325	-	SMD **0.71 SD lower** (1.48 lower to 0.06 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 3 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
7	randomized trials	serious[n]	not serious[f]	not serious[c]	serious[s]	none	512	449	-	SMD **0.14 SD lower** (0.27 lower to 0.01 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
5	randomized trials	serious[n]	not serious[f]	not serious[c]	very serious[d]	none	454	380	-	SMD **0.14 SD lower** (0.28 lower to 0)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
2	randomized trials	serious[e]	not serious[f]	serious[g]	very serious[h]	none	58	69	-	SMD **0.13 SD lower** (0.18 lower to 0.22 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age												
1	randomized trials	serious[a]	serious[i]	serious[i]	very serious[h]	none	67	67	-	SMD **0.29 SD lower** (0.63 lower to 0.05 higher)	⨁◯◯◯ Very low	
Population subgroup 6: ODI												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
3	randomized trials	serious[n]	not serious[f]	serious[g]	very serious[h]	none	125	136	-	SMD 0.**26 SD lower** (0.48 lower to 0.03 lower)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
3	randomized trials	serious[a]	not serious[f]	not serious[c]	very serious[h]	none	367	295	-	SMD **0.09 SD lower** (0.24 lower to 0.07 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 6 months												
2	randomized trials	serious[m]	not serious	serious[g]	very serious[h]	none	58	56	-	SMD **0.12 lower** (0.5 lower to 0.25 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
1	randomized trials	very serious[e]	serious[i]	serious[i]	very serious[h]	none	32	19	-	SMD **0.04 SD higher** (0.52 lower to 0.61 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[m]	serious[i]	serious[i]	very serious[h]	none	26	37	-	SMD **0.25 SD lower** (0.76 lower to 0.25 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age - not reported												
Population subgroup 6: ODI												
1	randomized trials	serious[m]	serious[i]	serious[i]	very serious[h]	none	26	27	-	SMD **0.25 SD lower** (0.76 lower to 0.25 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
1	randomized trials	very serious[e]	serious[i]	serious[i]	very serious[h]	none	32	19	-	SMD **0.04 SD higher** (0.52 lower to 0.61 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status 12 months												
1	randomized trials	serious[m]	serious[i]	serious[i]	very serious[h]	none	26	37	-	SMD **0.19 lower** (0.69 lower to 0.31 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries - not reported (no subgroup analysis performed; one trial performed in high-income countries)												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[m]	serious[i]	serious[i]	very serious[h]	none	26	37	-	SMD **0.19 SD lower** (0.69 lower to 0.31 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age - not reported												
Population subgroup 6: ODI												
1	randomized trials	serious[m]	serious[i]	serious[i]	very serious[h]	none	26	37	-	SMD **0.19 SD lower** (0.69 lower to 0.31 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ - not reported												
Health-related quality of life (higher scores mean better health)												
Health-related quality of life – Health-related quality of life at 1 month												
1	randomized trials	very serious[e]	serious[r]	serious[i]	very serious[h]	none	26	37	-	MD **4.5 SD higher** (0.46 higher to 8.54 higher)	⨁◯◯◯ Very low	
Health-related quality of life – Health-related quality of life at 3 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[e]	serious[ir]	serious[i]	very serious[h]	none	26	37	-	MD **2.8 SD higher** (1.24 lower to 6.84 higher)	⨁◯◯◯ Very low	
Health-related quality of life – Health-related quality of life at 6 months												
1	randomized trials	very serious[e]	serious[r]	serious[i]	very serious[h]	none	26	37	-	MD **1.7 SD higher** (2.34 lower to 5.74 higher)	⨁◯◯◯ Very low	
Health-related quality of life – Health-related quality of life at 12 months												
1	randomized trials	very serious[e]	serious[r]	serious[i]	very serious[h]	none	26	37	-	MD **1.7 SD higher** (2.34 lower to 5.74 higher)	⨁◯◯◯ Very low	
Return to work - Return to work at 1 month												
1	randomized trials	very serious[e]	serious[ir]	serious[i]	very serious[h]	none	1/2 (50.0%)	7/17 (41.2%)	**RR 1.21** (0.27 to 5.43)	**86 more per 1.000** (from 301 fewer to 1.000 more)	⨁◯◯◯ Very low	
Return to work - Return to work at 3 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[e]	serious[ir]	serious[i]	very serious[h]	none	2/3 (66.7%)	11/17 (64.7%)	**RR 1.03** (0.43 to 2.47)	**19 more per 1.000** (from 369 fewer to 951 more)	⨁◯◯◯ Very low	
General functional status (higher scores mean less disability)												
General functional status - General functional status at 1 month												
2	randomized trials	serious[m]	serious[b]	not serious[c]	very serious[h]	none	111	90	-	SMD 0.**57 higher** (0.55 lower to 1.69 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed)												
Population subgroup 5: participants over 60 years of age												
1	randomized trials	serious[a]	serious[r]	serious[i]	very serious[h]	none	69	67	-	SMD **0.02 SD higher** (0.32 lower to 0.36 higher)	⨁◯◯◯ Very low	
General functional status - General functional status at 3 months												
2	randomized trials	serious[m]	not serious[f]	not serious[c]	very serious[h]	none	103	85	-	SMD **0.07 lower** (0.36 lower to 0.22 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed)												
Population subgroup 5: participants over 60 years of age												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[r]	serious[i]	very serious[h]	none	67	67	-	SMD **0.02 SD lower** (0.36 lower to 0.32 higher)	⨁◯◯◯ Very low	
General functional status - General functional status at 6 months												
1	randomized trials	very serious[q]	serious[ir]	serious[i]	very serious[h]	none	32	19	-	SMD **0** (0.57 lower to 0.57 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3, 4 and 5 - not reported (no subgroup analysis performed)												
General functional status - Functional status at 12 months - not reported												
Psychological functioning - at 1 month												
2	randomized trials	very serious[t]	Serious[j]	Serious[g]	very serious[h]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	
Psychological functioning - at 3 months												
1	randomized trials	very serious[t]	Serious[j]	Serious[i]	very serious[h]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	
Psychological functioning - at 6 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	WHO SMT	sham/placebo SMT	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[t]	serious[i]	serious[i]	very serious[h]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	

Psychological functioning - at **12 months - not reported** (subgroup analysis of psychological functioning not conducted as data could not be pooled)

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Downgrade due to the presence of performance bias (lack of patient blinding) in all trials. We did not downgrade for the other risk of bias domains because most of the weight (>50%) comes from trials with a low risk of bias.
b. Downgrade because $I^2 > 75\%$, and treatment effects were in different directions, and were not able to be explained. Poor overlap of 95% CIs
c. We did not downgrade because trials were included from different countries, from different settings and populations.
d. Downgraded for the following: 1) sample <2000 participants; anTd 2) the lower 95% CI crosses the barrier of a potentially clinically-relevant threshold and the upper border is in favour of the control group.
e. Downgraded due to selection bias (unclear treatment allocation), performance bias (unclear risk due to co-interventions and compliance), and high risk of attrition bias.
f. Not downgraded due to treatment effect are similar, $I^2<50\%$ and CIs overlap
g. Downgraded because all trials that provided data were small for this outcome; single-center trials and not from different settings or countries .
h. Downgraded because <2,000 participants were included.
i. Downgraded because just one (small) trial provided data for this outcome; single-center trial and therefore not from different settings or countries.
j. Downgrade because treatment effects were in different directions. Poor overlap of 95% CIs. $I^2 > 50\%$
k. Downgrade due to attrition bias.
l. Downgraded although the $I^2 < 50\%$, the treatment effects were in different directions.
m. Downgraded due to selection bias (unclear treatment allocation), and high risk of attrition bias.
n. Downgraded because of a high risk of performance bias (patients and clinicians were not blinded in a majority of the trials) and unclear risk of selection bias (e.g. treatment allocation).
o. Downgraded one level because there were <2,000 participants but more than 1000 and the 95% CI was relatively broad (including a strong, clinically-relevant effect and no effect).
p. Not downgraded due to treatment effect are similar, $I^2<75\%$ and CIs overlap
q. Downgraded due to selection bias (unclear treatment allocation), performance bias (unclear risk of blinding patients and clinicians), and high risk of attrition bias and selective outcome reporting bias.
r. Downgraded because data comes one trial, small in size.
s. Downgraded one level as almost 1000 participants were included
t. Downgraded due to presence of performance bias and high risk of attrition bias.

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***What are the benefits and harms of SMT in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT	no intervention	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain - Pain at 1 month												
4	randomized trials	serious[a]	serious[b]	serious[c]	very serious[e]	none	218	107	-	MD **14 lower** (27.35 lower to 0.64 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - High-income countries												
3	randomized trials	serious[a]	serious[b]	serious[c]	very serious[e]	none	198	87	-	MD **8.8 lower** (18.17 lower to 0.57 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - Low- or lower middle-income Income countries												
1	randomized trials	serious[a]	serious[i]	serious[c]	very serious[e]	none	20	20	-	MD **36 lower** (43.9 lower to 28.1 higher)	⨁◯◯◯ Very low	
Pain - Pain at 3 months												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none	36	16	-	MD **14.2 lower** (26.89 lower to 1.51 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed) **Population subgroup 4: regional economic development - High-income countries**												
1	randomized trials	serious[f]	serious[i]	serious[c]	very serious[e]	none	36	16	-	MD **14.2 lower** (26.89 lower to 1.51 lower)	⨁◯◯◯ Very low	

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT	no intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 4: regional economic development - Low- or lower middle-income Income countries - not reported												
Pain - Pain at 6 months												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none	32	15	-	MD **4.9 lower** (18.68 lower to 8.88 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - High-income countries												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none	32	15	-	MD **4.9 higher** (18.68 higher to 8.88 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - Low- or lower middle-income Income countries - not reported												
Pain - Pain at 12 months - not reported												
Back-specific functional status (higher scores mean more disability)												
Back-specific functional status - back-specific functional status at 1 month –												
4	randomized trials	serious[a]	not serious[g]	serious[c]	very serious[e]	none	205	107	-	SMD **0.57 lower** (0.82 lower to 0.32 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, and 3 - not reported **Population subgroup 4: regional economic development - High-income countries**												
3	randomized trials	serious[a]	not serious[g]	serious[c]	very serious[e]	none	185	87	-	SMD **0.6 SD lower** (0.89 lower to 0.31 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - Low- or lower middle-income countries												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT	no intervention	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[i]	serious[c]	very serious[e]	none	20	20	-	SMD **0.38 SD lower** (1.01 lower to 0.24 higher)	⨁◯◯◯ Very low	
Population subgroup 5: ODI												
2	randomized trials	serious[a]	not serious[g]	serious[c]	very serious[e]	none	34	48	-	SMD **0.36 SD lower** (0.81 lower to 0.09 higher)	⨁◯◯◯ Very low	
Population subgroup 5: RMDQ												
2	randomized trials	serious[a]	not serious[g]	serious[c]	very serious[e]	none	171	59	-	SMD **0.66 SD lower** (1 lower to 0.33 lower)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 3 months												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none	36	17	-	SMD **0.03 higher** (0.54 lower to 0.61 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed) **Population subgroup 4: regional economic development - High-income countries**												
1	randomized trials	very serious[f]	serious[e]	serious[e]	very serious[e]	none	36	17	-	SMD **0.03 higher** (0.54 lower to 0.61 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 6 months												
1	randomized trials	very serious[f]	serious[e]	serious[e]	very serious[e]	none	32	15	-	SMD **0.18 lower** (0.8 lower to 0.43 higher)	⨁◯◯◯ Very low	

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT	no intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed) **Population subgroup 4: regional economic development - High-income countries**												
1	randomized trials	serious[f]	serious[e]	serious[e]	very serious[e]	none	32	15	-	SMD **0.18 lower** (0.8 lower to 0.43 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 12 months - not reported												
Health-related quality of life (higher scores mean better health) **Health-related quality of life - Health-related quality of life at 1 month**												
1	randomized trials	serious[i]	serious[e]	serious[e]	very serious[e]	none	129	42	-	MD **4.95 higher** (3.2 higher to 6.71 higher)	⨁◯◯◯ Very low	
Health-related quality of life - Health-related quality of life at 3 months, 6 months or 12 months - not reported												
General functional status (higher scores mean less disability) **General functional status - functional status at 1 month**												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none	42	17	-	MD **5.5 higher** (1.99 lower to 12.99 higher)	⨁◯◯◯ Very low	
General functional status - functional status at 3 months												
1	randomized trials	very serious[f]	serious[e]	serious[e]	very serious[e]	none	36	17	-	MD **10.4 higher** (2.79 higher to 18.01 higher)	⨁◯◯◯ Very low	
General functional status - functional status at 6 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT	no intervention	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[f]	serious[e]	serious[e]	very serious[e]	none	32	15	-	MD **8.5 higher** (0.12 higher to 16.88 higher)	⨁◯◯◯ Very low	
General functional status - Functional status at 12 months - not reported												
Psychological functioning - at 1 month												
2	randomized trials	very serious[f]	Serious[i]	Serious[c]	very serious[e]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	
Psychological functioning - at 3 months												
1	randomized trials	very serious[f]	Serious[i]	Serious[c]	very serious[e]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	
Psychological functioning - at 6 months												
1	randomized trials	very serious[f]	serious[i]	serious[c]	very serious[e]	none			-	Data was not pooled, because they used different measurements	⨁◯◯◯ Very low	
Psychological functioning - at **12 months - not reported**												
Subgroups for psychological functioning were not conducted as data could not be pooled												

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded due to the presence of performance bias (lack of patient blinding) in all trials. We did not downgrade for the other risk of bias domains because most of the weight (>50%) comes from trials with a low risk of bias.

b. Downgraded due to the presence of statistical heterogeneity (I2 = 68%) which could not be explained by subgroup analysis. In addition, the treatment effects and corresponding 95% CI varied in direction.
c. Downgraded because data comes from only single-centre trials and data does not come from different settings or countries.
d. Downgraded because the upper 95% CI crosses the barrier of a potentially clinically-relevant threshold and the lower border is close to no effect.
e. Downgraded because less than 2000 participants provided data for this outcome.
f. Downgraded due to the presence of high risk of performance bias (lack of patient blinding), attrition bias and selective reporting.
g. Not downgraded because the $I^2 < 50\%$, and there was sufficient overlap of the 95% CI's.
h. Downgraded because relatively few participants were recruited.
i. Downgraded due to the presence of performance bias (lack of patient blinding).
j. Downgraded because data comes from one trial small in size.

GRADE Table 3. ***What are the benefits and harms of SMT in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

One trial: data could not be extracted for GRADE assessment.

GRADE Table 4. ***What are the benefits and harms of SMT as an adjuvant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain)*?**

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain - Pain at 1 month												
10	randomized trials	serious[a]	serious[b]	not serious[c]	not serious[d]	none	650	864	-	MD **5.16 lower** (9.32 lower to 1 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
6	randomized trials	serious[a]	serious[b]	not serious[c]	not serious[d]	none	479	691	-	MD **3.13 lower** (7.73 higher to 1.48 higher)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development low- or lower middle-income income countries												
4	randomized trials	serious[a]	not serious[e]	serious[f]	very serious[g]	none	171	173	-	MD **9.05 lower** (14.71 lower to 3.39 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants over 60 years of age												
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	87	79	-	MD **2.9 lower** (8.85 lower to 3.05 higher)	⨁◯◯◯ Very low	
Pain - Pain at 3 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
5	randomized trials	serious[a]	not serious[e]	not serious[c]	not serious[d]	none	739	658	-	MD **4.34 lower** (8.83 lower to 0.15 higher)	⨁⨁⨁◯ Moderate	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	96	96	-	MD **9 lower** (24.42 lower to 6.42 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - high-income countries												
4	randomized trials	serious[a]	not serious[e]	not serious[c]	not serious[d]	none	722	640	-	MD **6.4 lower** (9.053 lower to 3.76 higher)	⨁⨁⨁◯ Moderate	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[a]	serious[i]	serious[f]	very serious[g]	none	171	173	-	MD **1.20 lower** (1.32 lower to 3.72 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	80	76	-	MD **7.9 lower** (13.89 lower to 1.91 lower)	⨁◯◯◯ Very low	
Pain - Pain at 6 months												
3	randomized trials	serious[a]	serious[b]	not serious[c]	very serious[i]	none	206	204	-	MD **4.22 lower** (15.12 lower to 6.67 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	79	77	-	MD **1.2 higher** (4.82 lower to 7.22 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
2	randomized trials	serious[a]	serious[l]	serious[f]	very serious[g]	none	127	127	-	MD **10.8 lower** (13.2 lower to 8.4 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[n]	serious[f]	very serious[g]	none	79	77	-	MD **1.2 higher** (4.82 lower to 7.22 higher)	⨁◯◯◯ Very low	
Pain - Pain at 12 months												
5	randomized trials	serious[a]	not serious [e]	not serious[c]	not serious[d]	none	823	745	-	MD **3.92 higher** (8.53 lower to 0.69 higher)	⨁⨁⨁◯ Moderate	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	96	96	-	MD **4 lower** (21.45 lower to 13.45 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - high-income countries												
4	randomized trials	serious[a]	not serious[k]	not serious[c]	not serious[d]	none	713	635	-	MD **2.42 lower** (5.19 lower to 0.35 higher)	⨁⨁⨁◯ Moderate	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[b]	serious[h]	very serious[g]	none	110	110	-	MD **10.4 lower** (13.01 lower to 7.79 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	80	76	-	MD **1.30 lower** (4.69 lower to 7.29 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 1 month (higher score mean more disability)												
7	randomized trials	serious[a]	serious[b]	not serious[c]	serious[i]	none	573	792	-	SMD **0.38 lower** (0.73 lower to 0.04 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
5	randomized trials	serious[a]	serious[b]	not serious[c]	serious[i]	none	446	663	-	SMD **0.14 SD lower** (0.36 lower to 0.09 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
2	randomized trials	serious[a]	serious[k]	serious[f]	very serious[g]	none	127	129	-	SMD **1.05 SD lower** (1.39 lower to 0.71 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	81	79	-	SMD **0.08 SD higher** (0.23 lower to 0.39 higher)	⨁◯◯◯ Very low	
Population subgroup 6: ODI												
3	randomized trials	serious[a]	serious[b]	serious[f]	very serious[g]	none	75	80	-	SMD **0.73 SD lower** (1.48 lower to 0.02 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
6	randomized trials	serious[a]	serious[b]	not serious[c]	serious[i]	none	523	742	-	SMD **0.4 SD lower** (0.8 lower to 0.01 lower)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 3 months												
5	randomized trials	serious[a]	not serious[k]	not serious[c]	serious[i]	none	763	696	-	SMD **0.13 lower** (0.29 lower to 0.03 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	serious[h]	serious[h]	very serious[g]	none	96	96	-	SMD **0.19 SD lower** (0.47 lower to 0.1 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - high-income countries												
4	randomized trials	serious[a]	not serious[e]	not serious[c]	serious[i]	none	746	687	-	SMD **0.14 SD lower** (0.31 lower to 0.03 higher)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[a]	Serious[n]	serious[f]	very serious[g]	none	17	18	-	SMD **0.11 SD higher** (0.55 lower to 0.77 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	80	76	-	SMD **0.01 SD lower** (0.32 lower to 0.31 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
5	randomized trials	serious[a]	serious[k]	not serious[c]	serious[i]	none	763	696	-	SMD **0.13 SD lower** (0.29 lower to 0.03 higher)	⨁◯◯◯ Very low	

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
Back-specific functional status (higher scores mean more disability)												
Back-specific functional status - back-specific functional status at 6 months												
3	randomized trials	serious[a]	serious[b]	not serious[c]	very serious[i]	none	206	204	-	SMD **0.4 lower** (0.91 lower to 0.11 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - high-income countries												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	79	77	-	SMD **0.28 SD lower** (0.6 lower to 0.04 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
2	randomized trials	serious[a]	serious[b]	serious[f]	very serious[g]	none	127	127	-	SMD **0.43 SD lower** (1.34 lower to 0.49 higher)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	79	77	-	SMD **0.28 SD lower** (0.6 lower to 0.04 lower)	⨁◯◯◯ Very low	
Population subgroup 6: ODI												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	17	17	-	SMD **0.05 SD higher** (0.62 lower to 0.73 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
3	randomized trials	serious[a]	serious[b]	not serious[c]	very serious[g]	none	206	204	-	SMD **0.4 SD lower** (0.91 lower to 0.11 higher)	⨁◯◯◯ Very low	
Back-specific functional status - back-specific functional status at 12 months												
4	randomized trials	serious[a]	not serious[e]	not serious[c]	serious[i]	none	816	746	-	SMD **0.23 lower** (0.43 lower to 0.03 lower)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	Serious[n]	serious[f]	very serious[g]	none	96	96	-	SMD **0.1 SD lower** (0.38 lower to 0.19 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - high-income countries												
3	randomized trials	serious[a]	not serious[k]	not serious[c]	serious[i]	none	706	636	-	SMD **0.16 SD lower** (0.27 lower to 0.05 lower)	⨁⨁◯◯ Low	

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 4: regional economic development - low- or lower middle-income income countries												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	110	110	-	SMD **0.67 SD lower** (0.94 lower to 0.4 lower)	⨁◯◯◯ Very low	
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	80	76	-	SMD **0.08 SD higher** (0.23 lower to 0.4 higher)	⨁◯◯◯ Very low	
Population subgroup 6: RMDQ												
4	randomized trials	serious[a]	serious[b]	not serious[c]	serious[i]	none	816	746	-	SMD **0.23 SD lower** (0.43 lower to 0.03 lower)	⨁◯◯◯ Very low	
Health-related quality of life - Health-related quality of life at 1 month (higher scores mean better health)												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[h]	none	81	79	-	MD **0.6 SD higher** (1.25 lower to 2.45 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis performed)												
Population subgroup 4: regional economic development - not reported (No subgroup analysis performed; only one trial)												
Population subgroup 5: participants 60 years and older												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	81	79	-	MD **0.6 higher** (1.25 lower to 2.45 higher)	⨁◯◯◯ Very low	
Health-related quality of life (higher scores mean better health)												
Health-related quality of life - Health-related quality of life at 3 months												
3	randomized trials	serious[a]	not serious[k]	not serious[c]	very serious[l]	none	435	399	-	MD **1.78 SD higher** (0.19 higher to 3.36 higher)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	serious[n]	serious[h]	very serious[g]	none	96	96	-	MD **3.4 higher** (3.2 lower to 10 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - not reported												
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	serious	serious[h]	very serious[g]	none	80	76	-	MD **0.5 higher** (1.38 lower to 2.38 higher)	⨁◯◯◯ Very low	
Health-related quality of life - Health-related quality of life at 6 months												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[h]	none	79	77	-	SMD **0.3 SD lower** (2.21 lower to 1.61 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed)												
Population subgroup 5: participants 60 years and older												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	79	77	-	MD **0.3 lower** (2.21 lower to 1.61 higher)	⨁◯◯◯ Very low	
Health-related quality of life - Health-related quality of life at 12 months												
4	randomized trials	serious[a]	serious[b]	not serious[c]	serious[l]	none	428	393	-	MD **0.31 higher** (2.29 lower to 2.91 higher)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular pain												
1	randomized trials	serious[a]	Seriousn	serious[h]	very serious[g]	none	96	96	-	MD **1.5 higher** (4.96 lower to 7.96 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development – not reported (No subgroup analysis performed; only one trial)												
Population subgroup 5: participants 60 years and older												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMT with an intervention	Same intervention alone	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[a]	Serious	serious[h]	very serious[g]	none	80	76	-	MD **1.5 lower** (3.38 lower to 0.38 higher)	⨁◯◯◯ Very low	
Psychological functioning - Psychological functioning at 1 month												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[g]	none	81	79	-	MD **0.4 SD higher** (1.38 lower to 2.18 higher)	⨁◯◯◯ Very low	
Psychological functioning - Psychological functioning at 3 months												
3	randomized trials	serious[a]	not serious[k]	not serious[c]	very serious[g]	none	435	399	-	MD **1.33 SD higher** (0.91 lower to 3.58 higher)	⨁◯◯◯ Very low	
Psychological functioning - Psychological functioning at 6 months												
1	randomized trials	serious[a]	Serious[n]	serious[h]	very serious[h]	none	79	77	-	MD **1.7 SD higher** (0.18 lower to 3.58 higher)	⨁◯◯◯ Very low	
Psychological functioning - Psychological functioning at 12 months												
3	randomized trials	serious[a]	serious[b]	not serious[c]	very serious[g]	none	428	393	-	MD **0.42 SD higher** (1.42 lower to 2.27 higher)	⨁◯◯◯ Very low	
Subgroup analysis of psychological functioning not conducted.												

Web Annex D.B3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgrade due to the presence of performance bias (lack of patient blinding) in all trials. We did not downgrade for the other risk of bias domains because most of the weight (>50%) comes from trials with a low risk of bias.
b. Downgraded suggesting substantial statistical heterogeneity (I2 >50%). In addition, the treatment effects and corresponding 95% CI varied in direction and could not be explained.
c. We did not downgrade because trials were included from different countries, from different settings and populations.
d. Not downgraded. The 95% CI's are sufficiently narrow and do not across the line of no effect nor the clinically-relevant threshold.
e. Not downgraded because although the I^2 is high, all treatment effects were in the same direction, except one small trial, and there was sufficient overlap of the 95% CI's.
f. Downgraded because only single centered (small) trials and data does not come from different settings or countries.
g. Downgraded because < 2000 participants, very few participants were recruited.
h. Downgraded because just one (small) trial provided data for this outcome, therefore data does not come from different settings or countries..
i. Downgraded for the following: the lower 95% CI crosses the barrier of a potentially clinically-relevant threshold, and the upper border is close to no effect.
j. Downgraded for the following: 1) 410 participants; and 2) the lower 95% CI crosses the barrier of a potentially clinically-relevant threshold and the upper border is in favour of the control group.
k. Not downgraded because the I^2 < 50%, and there was sufficient overlap of the 95% CI's.
l. Downgraded because the upper 95% CI crosses the barrier of a potentially clinically-relevant threshold, and the lower border is close to no effect.
m. Downgraded because data is provided from almost 1000 participants.
n. Downgraded because data comes from one trial, small in size.

B.4 Massage

Overview of the PICO structure

Definition of the intervention	
Massage is the manual manipulation of soft body tissues to enhance health and well-being. Practised globally, there are more than 80 different forms of massage, many developed in the last 30 years. While massage may be used for a variety of specific indications (e.g., relaxation, comfort at the end of life, relieving pain, enhancing athletic performance), it is undertaken with the general goal of helping the body achieve or increase health and well-being. In the evidence review for this guideline, massage was broadly defined and included any soft-tissue manipulation using hands or another mechanical device and traditional, complementary and integrative (TCI) medicine massage. Massage could be applied to any body part, to the lumbar region only, or to the whole body.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial) d) Adjuvant therapy, i.e. where the additional effect of an intervention could be isolated

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; trivial; uncertain; varies	Small; trivial; uncertain
Harms	Uncertain	Uncertain

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Balance benefits to harms	Probably favours massage; probably does not favour massage; uncertain	Probably favours massage; probably does not favour massage; uncertain
Overall certainty	Low; very low	Low; very low
Values and preferences	Probably important uncertainty or variability; possibly important uncertainty or variability	Probably important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate costs; uncertain; varies	Moderate costs; varies
Equity and human rights	No impact; probably reduced (traction especially); varies	No impact; probably reduced (traction especially); uncertain; varies
Acceptability	Yes; probably yes; probably no; uncertain; varies	Yes; probably yes; probably no; uncertain; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of massage in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham massage?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain intensity (higher scores mean more pain) - Pain in immediate term (1 month)												
5[1]	randomized trials	serious[a]	not serious	serious[b]	very serious[c]	none	102	103	-	MD **3.07 lower** (7.34 lower to 1.21 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low income 1[2]	randomized trials	serious[a]	not serious	serious[b]	very serious[c]	none	26	25	-	MD 0.7 **higher** (4.20 lower to 5.60 higher)	⨁◯◯◯ Very low	
High income 4[3]	randomized trials	serious[a]	not serious	serious[b]	very serious[c]	none	76	78	-	MD 7.6 **lower** (13.76 lower to 1.48 lower)	⨁◯◯◯ Very low	
Population subgroup 4: Older adults (over 60 years of age)												
Older adults[2]	randomized trials	serious[a]	serious	serious[b]	very serious[c]	none	26	25	-	MD **0.70 lower** (4.20 lower to 5.60 higher)	⨁◯◯◯ Very low	

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain) - Pain in short term (1-3 months)												
3[4]	randomized trials	serious[a]	not serious	serious[b]	very serious[d]	none	60	60	-	MD **14.25 lower** (20.28 lower to 8.22 lower)	⨁◯◯◯ Very low	
Population subgroup 1: gender and/or sex												
Women[5]	randomized trials	serious[a]	not serious	serious[b]	very serious[d]	none	26	25	-	MD **13.30 lower** (20.91 lower to 5.69 lower)	⨁◯◯◯ Very low	
Men[6]	randomized trials	serious[a]	not serious	serious[b]	very serious[d]	none	34	35	-	MD **15.85 lower** (25.71 lower to 5.98 lower)	⨁◯◯◯ Very low	
Population subgroups 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low income[7]	randomized trials	serious[a]	not serious	serious[b]	very serious[d]	none	26	25	-	MD **13.30 lower** (20.91 lower to 5.69 lower)	⨁◯◯◯ Very low	
High income[8]	randomized trials	serious[a]	not serious	serious[b]	very serious[d]	none	34	35	-	MD **15.85 lower** (25.71 lower to 5.96 lower)	⨁◯◯◯ Very low	
Population subgroup 5: Older adults												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
Older adults[2]	randomized trials	serious[a]	serious	serious[b]	very serious[c]	none	26	25	-	MD 13.30 lower (20.91 lower to 5.69 higher)	⨁◯◯◯ Very low	
Pain intensity (higher scores mean more pain) - Pain in intermediate term (3-6 months)												
1[9]	randomized trials	serious[e]	serious[f]	serious[g]	very serious[f]	none	7	8	-	MD **10 lower** (16.58 lower to 3.42 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development - not reported (no subgroup analysis was performed, only 1 study included)												
Pain intensity (higher scores mean more pain) - Pain in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Functioning (higher scores mean more disability) - Functioning in immediate term (1 month)												
4[10]	randomized trials	serious[a]	not serious	serious[h]	very serious[c]	none	76	78	-	SMD **0.5 lower** (0.96 lower to 0.04 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
'Functioning (higher scores mean more disability) - Functioning in short term (1-3 months)												
4[11]	randomized trials	serious[e]	not serious	serious[i]	very serious[c]	none	98	96	-	SMD **0.4 lower** (0.68 lower to 0.11 lower)	⨁◯◯◯ Very low	

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 1: gender and/or sex												
Only women[12]	randomized trials	serious[e]	not serious	serious[i]	very serious[c]	none	26	25	-	SMD **1.33 lower** (4.90 lower to 2.24 higher)	⨁◯◯◯ Very low	
Men & Women[13]	randomized trials	serious[e]	not serious	serious[i]	very serious[c]	none	72	71	-	SMD **2.44 lower** (4.57 lower to 0.31 lower)	⨁◯◯◯ Very low	
Population subgroup 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low income[14]	randomized trials	serious[e]	not serious	serious[i]	very serious[c]	none	38	36	-	SMD **0.49 lower** (0.95 lower to 0.03 lower)	⨁◯◯◯ Very low	
High income[15]	randomized trials	serious[e]	not serious	serious[i]	very serious[c]	none	60	60	-	SMD **0.34 lower** (0.70 lower to 0.02 higher)	⨁◯◯◯ Very low	
Population subgroup 5: Older adults (over 60 years of age)												
Older adults[2]	randomized trials	serious[a]	serious	serious[b]	very serious[c]	none	26	25	-	MD 0.20 lower (0.75 lower to 0.35 higher)	⨁◯◯◯ Very low	
Functioning (higher scores mean more disability) - Functioning in intermediate term (3-6 months)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
2[16]	randomized trials	serious[e]	not serious	serious[g]	very serious[d]	none	45	44	-	SMD **0.35 lower** (0.76 lower to 0.07 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of Life (higher scores mean better QoL)												
Quality of Life (higher scores mean better QoL) - QoL in immediate term (1 month)												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of Life (higher scores mean better QoL) - QoL in short term (1-3 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of Life (higher scores mean better QoL) - QoL in intermediate term (3-6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of Life (higher scores mean better QoL) - QoL in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Fear avoidance belief (higher scores mean more fear avoidance) - Fear avoidance in immediate term (1 month)												
2[17]	randomized trials	not serious	not serious	not serious	very serious[d]	none	45	45	-	MD **14 lower** (22.84 lower to 5.15 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Fear avoidance belief (higher scores mean more fear avoidance) - Fear avoidance in short term (1-3 months)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Sham	Relative (95% CI)	Absolute (95% CI)		
2[18]	randomized trials	not serious	not serious	not serious	very serious[d]	none	45	45	-	MD **13.5 lower** (22.86 lower to 4.14 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Fear avoidance belief (higher scores mean more fear avoidance) - Fear avoidance in intermediate term (3-6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	
Fear avoidance belief (higher scores mean more fear avoidance) - Fear avoidance in long term (> 6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; **MD:** mean difference; **SMD:** standardized mean difference

Explanations

a. Downgraded for selection bias (unclear treatment allocation), performance bias (unclear co-interventions and compliance), and selective outcome reporting bias.
b. Downgraded because Kim 2021 only included participants >65 years of age and only women (and responsible for >50% of the weight in the meta-analysis); in 4 out of 5 studies (80% of the weight) massage of the spine was used, while Quinn 2008 (17% of the weight) used a different form of massage (reflexology - foot massage representative of the points in the spine).
c. Downgraded by one level because there were very few participants (ca. 200), and downgraded by one level based on a relatively broad 95% CI.
d. Downgraded by one level because there were very few participants (ca. 100), and downgraded by one level based on a relatively broad 95% CI.
e. Downgraded by for selection bias (unclear treatment allocation) and performance bias (unclear co-interventions).
f. Downgraded by because just one small study examined this treatment comparison.
g. Downgraded by because Quinn 2008 used a different form of massage (reflexology - foot massage representative of the points in the spine).
h. Downgraded by because all the studies were single-centre; high income; and intervention is different for one study (Quinn 2008 (15% of the weight)).
i. Downgraded by because all the studies were single-centre; some low, some high income; and the intervention was different across the studies (myofascial release, foot reflexology, acupressure).

References

1. Arguisuela 2017, Arguisuela 2019, Geisser 2015, Kim 2021, Quinn 2008
2. Kim 2021
3. Arguisuela 2017, Arguisuela 2019, Geisser 2015, Quinn 2008
4. Arguisela 2017, Kim 2021, Quinn 2008
5. Kim 2021

6. Arguisela 2017, Quinn 2008
7. Kim 2021
8. Arguisuela 2017, Quinn 2008
9. Quinn 2008
10. Arguisuela 2017, Arguisuela 2019, Geisser 2015, Quinn 2008
11. Ajimsha 2014, Arguisuela 2017, Kim 2021, Quinn 2008
12. Kim 2021
13. Ajimsha 2014, Arguisuela 2017, Quinn 2008
14. Ajimsha 2014
15. Arguisuela 2017, Kim 2021, Quinn 2008
16. Arguisuela 2017, Quinn 2008
17. Arguisuela 2017, Arguisuela 2019
18. Arguisuela 2017, Arguisuela 2019

GRADE Table 2. ***What are the benefits and harms of massage in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

No trials

GRADE Table 3. ***What are the benefits and harms of massage in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain intensity (higher scores mean more pain) - Pain in immediate term (1 month)												
1[1]	randomized trials	serious[a]	serious[b]	serious[c]	very serious[b]	none	30	24	-	MD **5 lower** (16.44 lower to 6.44 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain intensity (higher scores mean more pain) - Pain in short term (1-3 months)												
2[2]	randomized trials	serious[d]	not serious	serious[c]	very serious[e]	none	95	69	-	MD **12.19 lower** (20.16 lower to 4.22 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain intensity (higher scores mean more pain) - Pain in intermediate term (3-6 months)												
1[3]	randomized trials	serious[d]	serious[b]	serious[c]	very serious[b]	none	57	45	-	MD **2.9 lower** (14.16 lower to 8.36 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain intensity (higher scores mean more pain) - Pain in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-
Functioning (higher scores mean more disability)												
Functioning (higher scores mean more disability) - Functioning in immediate term (1 month)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Usual care	Relative (95% CI)	Absolute (95% CI)		
1[4]	randomized trials	serious[a]	serious[b]	serious[c]	very serious[b]	none	30	24	-	SMD **0.06 lower** (0.6 lower to 0.48 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in short term (1-3 months)												
3[5]	randomized trials	serious[f]	not serious	not serious	very serious[g]	none	363	202	-	SMD **0.51 lower** (0.72 lower to 0.3 lower)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Radicular pain[6]	randomized trials	serious[f]	not serious	not serious	very serious[g]	none	363	202	-	SMD **0.59 lower** (0.80 lower to 0.37 lower)	⨁◯◯◯ Very low	
Radicular pain not presented[7]	randomized trials	serious[f]	not serious	not serious	very serious[g]	none	363	202	-	SMD **0.37 lower** (0.69 lower to 0.06 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in intermediate term (3-6 months)												
2[8]	randomized trials	serious[f]	not serious	not serious	very serious[g]	none	325	178	-	SMD **0.34 lower** (0.52 lower to 0.15 lower)	⨁◯◯◯ Very low	

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Usual care	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in long term (>6 months)												
1[9]	randomized trials	serious[b]	serious[b]	not serious	very serious[b]	none	268	132	-	SMD **0.18 lower** (0.46 lower to 0.09 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Quality of Life (higher scores mean better QoL)												
Quality of Life (higher scores mean better QoL) - QoL in immediate term (1 month)												
1[10]	randomized trials	serious[f]	serious[b]	serious[c]	very serious[b]	none	30	24	-	SMD **0.99 lower** (1.56 lower to 0.42 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Quality of Life (higher scores mean better QoL) - QoL in short term (1-3 months)												
1[11]	randomized trials	serious[f]	serious[b]	serious[c]	very serious[c]	none	57	45	-	SMD **0.33 lower** (0.72 lower to 0.07 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Quality of Life (higher scores mean better QoL) - QoL in intermediate term (3-6 months)												
1[12]	randomized trials	serious[f]	serious[b]	serious[c]	very serious[c]	none	57	45	-	SMD **0.12 lower** (0.51 lower to 0.27 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage	Usual care	Relative (95% CI)	Absolute (95% CI)		
Quality of Life (higher scores mean better QoL) - QoL in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-
Depression (higher scores mean more depression)												
Depression (higher scores mean more depression) - Depression in immediate term (1 month)												
-	-	-	-	-	-	-	-	-	-	-	-	
Depression (higher scores mean more depression) - Depression in short term (1-3 months)												
1[13]	randomized trials	serious[f]	serious[b]	serious[c]	very serious[c]	none	57	45	-	MD **3.4 lower** (7.45 lower to 0.65 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Depression (higher scores mean more depression) - Depression in intermediate term (3-6 months)												
1[14]	randomized trials	serious[f]	serious[b]	serious[c]	very serious[c]	none	57	45	-	MD **1.2 lower** (5.1 lower to 2.7 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Depression (higher scores mean more depression) - Depression in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-

CI: confidence interval; **MD:** mean difference; **SMD:** standardized mean difference

Explanations

a. Downgraded due to high risk of performance bias (patients and clinicians were not blinded to the intervention).
b. Downgraded because just one study examined this comparison.
c. Downgraded because single-center study with few participants.
d. Downgraded by two levels due to high risk of selection bias (treatment allocation), performance bias (patients and clinicians were not blinded to the intervention), and unclear risk for selective outcome reporting bias.
e. Downgraded because relatively few participants were included (ca. 200).

f. Downgraded due to high risk of selection bias (treatment allocation), and high risk of performance bias (patients and clinicians were not blinded to the intervention),
g. Downgraded because few participants were included (ca. 550).

References

1. Kobayashi 2019
2. Kobayashi 2019, Poole 2017
3. Poole 2017
4. Kobayashi 2019
5. Cherkin 2011, Kobayashi 2019, Poole 2007
6. Cherkin 2011
7. Kobayashi 2019, Poole 2007
8. Cherkin 2011, Poole 2007
9. Cherkin 2011
10. Kobayashi, 2019
11. Poole 2007
12. Poole 2007
13. Poole 2007
14. Poole 2007

GRADE Table 4. ***What are the benefits and harms of massage as an adjuvant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain)?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage as Adjuvant therapy	placebo	Relative (95% CI)	Absolute (95% CI)		
Pain intensity (higher scores mean more pain)												
Pain intensity (higher scores mean more pain) - Pain in immediate term (1 month)												
4[1]	randomized trials	serious[a]	serious[b]	not serious	very serious[c]	none	123	123	-	MD **2.35 lower** (10.54 lower to 5.83 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain intensity (higher scores mean more pain) - Pain in short term (1-3 months)												
4[2]	randomized trials	serious[d]	serious[b]	not serious	very serious[c]	none	108	109	-	MD **8.13 lower** (13.93 lower to 2.33 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Population subgroup 5: Older adults (over 60 years of age)												
Older adults[7]	randomized trials	serious[a]	serious[b]	serious[b]	very serious[c]	none	22	23		MD 13.40 lower (21.84 lower to 4.96 lower)	⨁◯◯◯ Very low	
Pain intensity (higher scores mean more pain) - Pain in intermediate term (3-6 months)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage as Adjuvant therapy	placebo	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	-
Pain intensity (higher scores mean more pain) - Pain in long term (> 6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-
Functioning (higher scores mean more disability)												
Functioning (higher scores mean more disability) - Functioning in immediate term (1 month)												
4[3]	randomized trials	serious[a]	not serious	not serious	very serious[c]	none	123	123	-	SMD **0.38 lower** (0.63 lower to 0.13 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in short term (1-3 months)												
2[4]	randomized trials	serious[a]	serious[e]	not serious	very serious[e]	none	56	56	-	SMD **0.86 lower** (1.90 lower to 0.17 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Functioning (higher scores mean more disability) - Functioning in intermediate term (3-6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-
Functioning (higher scores mean more disability) - Functioning in long term (>6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage as Adjuvant therapy	placebo	Relative (95% CI)	Absolute (95% CI)		
Quality of Life (higher scores mean better QoL)												
Quality of Life (higher scores mean better QoL) - QoL in immediate term (1 month)												
1[5]	randomized trials	serious[a]	serious[e]	not serious	very serious[e]	none	56	56	-	MD **1.00 higher** (-8.24 lower to 10.24 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Quality of Life (higher scores mean better QoL) - QoL in short term (1-3 months)												
2[6]	randomized trials	serious[a]	serious[e]	not serious	very serious[e]	none	56	56	-	MD **1.48 lower** (-7.12 lower to 4.26 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Population subgroup 5: Older adults (over 60 years of age)												
Older adults[7]	randomized trials	serious[a]	serious[b]	serious[b]	very serious[c]	none		22	23	MD 3.52 lower (10.74 lower to 3.7 higher)	⨁◯◯◯ Very low	
Quality of Life (higher scores mean better QoL) - QoL in intermediate term (3-6 months)												
-	-	-	-	-	-	-	-	-	-	-	-	-
Quality of Life (higher scores mean better QoL) - QoL in long term (>6 months)												

Web Annex D.B4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage as Adjuvant therapy	placebo	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	-

CI: confidence interval; **MD:** mean difference; **SMD:** standardized mean difference

Explanations

a. Downgraded for high risk of bias (performance bias (patients and clinicians were not blinded to the intervention)).
b. Downgraded for substantial statistical heterogeneity (I-squared>75%).
c. Downgraded because there were very few participants (ca. 200).
d. Downgraded for selection bias (because the treatment allocation was unclear for >50% weight of studies), and high risk of performance bias.
e. Downgraded by one level because just one study with a small number of participants examined this comparison, and downgraded by one level based on a relatively broad 95%CI

References

1. Ali-Khorsand 2019, Bellido-Fernandez 2021, Boff 2020, Shu 2021
2. Ali-Khorsand 2019, Boff 2020, Ozsoy 2019, Zheng 2012
3. Ali-Khorsand 2019, Bellido-Fernandez 2021, Boff 2020, Shu 2021
4. Ali-Khorsand 2019, Boff 2020
5. Boff 2020
6. Boff 2020, Ozsoy 2019
7. Ozsoy 2019

B.5 Traction

Overview of the PICO structure

Definition of Intervention	
Traction is the application of a distraction force to the long axis of the spine, achieved using body weight (either of a therapist or patient), external weights, and/or pulleys. The evidence review for this guideline included all types of traction such as mechanical or motorized traction (where the traction is exerted by a motorized pulley), manual traction (in which the traction is exerted by the therapist, using their body weight to alter the force and direction of the pull), auto-traction (where the person controls the traction forces by grasping and pulling bars at the head of the traction table), and also less common forms such as underwater traction (where the person is fixed perpendicularly in a deep pool, a bar grasped under the arms and traction applied) and gravitational traction (e.g. bed rest traction, in which the person is fixed to a tilted table or bed, or inverted traction, where the participant is held in an inverted position by the ankles and another part of the lower extremities and gravity provides the force). Traction can be intermittent or continuous and applied for a few seconds to several hours.	
PICO question	
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

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Comparators	a) Placebo/sham b) No or minimal intervention, or where the effect of Intervention can be isolated c) Usual care (described as usual care in the trial) d) Adjuvant therapy, i.e. where the additional effect of an intervention could be isolated
Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

Summary of *values and preferences*	
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.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; trivial; uncertain	Small; trivial; uncertain

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Harms	Uncertain	Uncertain
Balance benefits to harms	Probably does not favour traction; uncertain	Probably does not favour traction; uncertain
Overall certainty	Very low	Very low
Values and preferences	Probably important uncertainty or variability; possibly important uncertainty or variability	Probably important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate costs; varies	Moderate costs; varies
Equity and human rights	Probably reduced; uncertain; varies	Probably reduced; uncertain; varies
Acceptability	Yes; probably yes; probably no; uncertain; varies	Yes; probably yes; probably no; uncertain; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of traction in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham traction?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction	sham	Relative (95% CI)	Absolute (95% CI)		
Pain intensity - Pain in immediate term (1 month) - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Pain intensity (higher scores mean more pain) - Pain in short term (1-3 months)												
1[a]	randomized trials	serious[b]	serious[c]	not serious	very serious[c]	none	31	29	-	MD 4.00 lower (17.65 lower to 9.65 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed, only one included study for this outcome)												
Pain intensity - Pain in intermediate (3-6 months) or long term (>6 months)- no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Function - Function in immediate (1 month), short (1-3 months), intermediate (3-6 months) or long term (> 6 months) - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of life - Quality of life in immediate (1 month), short (1-3 months), intermediate (3-6 months) or long term (> 6 months) - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events , psychological functioning (depression) or social participation - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference

Explanations

a. Schimmel 2006

b. Downgraded for selective outcome reporting bias.

c. Downgraded by one level because there were very small number of participants and downgraded by one level based on a relatively broad 95%

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***What are the benefits and harms of traction in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

No trials

GRADE Table 3. ***What are the benefits and harms of traction in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

No trials

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 4. ***What are the benefits and harms of traction as adjuvant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain)?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction used as an adjuvant treatment (e.g. traction + intervention	Intervention alone)	Relative (95% CI)	Absolute (95% CI)		
Pain intensity - Pain in immediate term (1 month, assessed with: VAS at rest; Scale from: 0 to 100)												
6[a]	randomized trials	serious[b]	not serious[c]	serious[d]	very serious[e]	none	256	203	-	MD 6.2 lower (9.67 lower to 2.74 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Middle income 5[f]	randomized trials	serious[b]	not serious[c]	serious[d]	very serious[e]	none	226	173	-	MD 5.98 lower (8.61 lower to 3.34 lower)	⨁◯◯◯ Very low	
High income 1[g]	randomized trials	serious[b]	not serious[h]	serious[i]	very serious[e]	none	30	30	-	MD 5.4 lower (8.47 lower to 2.33 lower)	⨁◯◯◯ Very low	

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction used as an adjuvant treatment (e.g. traction + intervention	Intervention alone)	Relative (95% CI)	Absolute (95% CI)		
Pain intensity - Pain in short term (1-3 months, assessed with: VAS at rest; Scale from: 0 to 100)												
3[j]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[m]	none	85	89	-	MD 4.07 lower (12.81 lower to 4.66 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain intensity - Pain in intermediate term (3-6 months, assessed with: VAS at rest; Scale from: 0 to 100)												
3[n]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[m]	none	92	93	-	MD 13.27 lower (20.71 lower to 5.83 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Middle income 2[o]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[m]	none	62	63	-	MD 15.47 lower (28.21 lower to 2.73 lower)	⨁◯◯◯ Very low	

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction used as an adjuvant treatment (e.g. traction + intervention	Intervention alone)	Relative (95% CI)	Absolute (95% CI)		
High income 1[g]	randomized trials	serious[b]	serious[h]	serious[i]	very serious[e]	none	30	30	-	MD 9.50 lower (12.43 lower to 6.57 lower)	⨁◯◯◯ Very low	
Pain intensity - Pain in long term (> 6 months) no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Function - Functioning in immediate term (1 month, assessed with: ODI; Scale from: 0 to 100)												
6[a]	randomized trials	serious[b]	serious[p]	serious[d]	very serious[e]	none	256	203	-	MD 3.8 lower (6.26 lower to 1.34 lower	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Middle income 5[f]	randomized trials	serious[b]	serious[p]	serious[d]	very serious[e]	none	226	173	-	MD 4.28 lower (7.25 lower to 1.32 lower	⨁◯◯◯ Very low	
High income 1[g]	randomized trials	serious[b]	serious[h]	serious[i]	very serious[e]	none	30	30	-	MD 1.93 lower (2.77 lower to 1.09 lower	⨁◯◯◯ Very low	

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction used as an adjuvant treatment (e.g. traction + intervention	Intervention alone)	Relative (95% CI)	Absolute (95% CI)		
Function - Functioning in short term (1-3 months, assessed with: ODI; Scale from: 0 to 100)												
3[j]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[m]	none	85	89	-	MD 1.91 lower (4.56 lower to 0.73 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Function - Functioning in intermediate term (3-6 months, assessed with: ODI; Scale from: 0 to 100)												
3[n]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[e]	none	92	93	-	MD 4.64 lower (7.75 lower to 1.54 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Middle income 2[o]	randomized trials	serious[k]	serious[l]	serious[d]	very serious[e]	none	62	63	-	MD 5.69 lower (10.40 lower to 0.99 lower)	⨁◯◯◯ Very low	

Web Annex D.B5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Traction used as an adjuvant treatment (e.g. traction + intervention	Intervention alone)	Relative (95% CI)	Absolute (95% CI)		
High income 1[g]	randomized trials	serious[b]	serious[h]	serious[i]	very serious[m]	none	30	30	-	MD 2.66 lower (3.38 lower to 1.94 lower)	⨁◯◯◯ Very low	
Function - Functioning in long term (> 6 months) no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Quality of life - Quality of life in immediate term (1 month, assessed with SF-36)												
1[q]	randomized trials	serious[k]	serious[h]	serious[r]	very serious[e]		30	30	-	MD 1.97 lower (7.29 lower to 3.35 higher)	-	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Quality of life - Quality of life in short (1-3 months), intermediate (3-6 months) or long term (> 6 months) - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events, psychological functioning or social participation - no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; **MD**: mean difference

Explanations

a. Al Amar 2019; Amjad 2022; Bilgilisoy Filiz 2018; Borman 2003; Gulsen 2018; Mohamed 2020.

b. Downgraded given high risk of bias due to performance bias (lack of patient and clinician blinding), and two other domains which were unclear (selection bias and selective outcome reporting bias)
c. We did not downgrade because the majority of the studies favored Intervention and sufficient consistency across the studies.
d. All patients were recruited in an outpatient clinic from hospitals with leg pain, and all received high load mechanical traction.
e. Downgraded because there were relatively few participants (<500)
f. Amjad 2022; Bilgilisoy Filiz 2018; Borman 2003; Gulsen 2018; Mohamed 2020.
g. Al Amar 2019
h Inconsistency not assessed because only one study included in this analysis.
i. Indirectness downgraded because only one study included in this subgroup, unclear if it is representative of all high-income countries.
j. Borman 2003; Diab 2013; Moustafa 2012.
k. Downgraded due high risk of performance bias (lack of patient and clinician blinding).
l. Downgraded due to substantial statistical heterogeneity (I-squared>75%).
m Downgraded by one level because there were very small number of participants and downgraded by one level based on a relatively broad 95% CI (the lower border is consistent with a potentially clinically relevant effect).
n. Al Amar 2019; Diab 2013; Moustafa 2012.
o. Diab 2013; Moustafa 2012
p. Downgraded by because there was substantial statistical heterogeneity.
q. Amjad 2022
r. Indirectness downgraded because only one study included in this subgroup.

B.6 Therapeutic ultrasound

Overview of the PICO structure

Definition of the intervention	
Therapeutic ultrasound is an electrophysical treatment modality postulated to deliver energy to deep tissue sites through ultrasonic waves, to increase tissue temperature and/or create non-thermal physiological changes. Physiological changes are purported to improve symptoms (pain, inflammation) and promote or accelerate tissue healing. Unlike diagnostic ultrasound for medical imaging (which transmits ultrasonic waves and transforms the returning echo into an image), therapeutic ultrasound is a one-way energy delivery system which uses a crystal sound head to transmit acoustic waves at 1 or 3 MHz and at amplitude densities of between 0.1 W/cm^2 and 3 W/cm^2, in continuous or pulsed mode.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

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

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; trivial; uncertain	Small; trivial; uncertain
Harms	Trivial; uncertain	Trivial; uncertain

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Balance benefits to harms	Does not favour ultrasound; probably does not favour ultrasound; uncertain	Does not favour ultrasound; probably does not favour ultrasound; uncertain
Overall certainty	Low; very low	Low; very low
Values and preferences	Possibly important uncertainty or variability; probably no important uncertainty or variability	Possibly important uncertainty or variability; probably no important uncertainty or variability
Resource considerations	Moderate; moderate costs; negligible; negligible costs and savings	Moderate; moderate costs; negligible; negligible costs and savings
Equity and human rights	No impact; probably reduced; uncertain	No impact; probably reduced; uncertain
Acceptability	Yes; probably yes; probably no; varies	Yes; probably yes; probably no; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

GRADE Table 1. ***What are the benefits and harms of therapeutic ultrasound in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham ultrasound?***

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	Sham ultrasound	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (assessed with: VAS at rest; Scale from: 0 to 100)[a]												
4[b,c]	randomized trials	serious[d]	very serious[e]	not serious	serious[f]	none	69	70	-	MD 10.24 lower (24.3 lower to 3.81 higher)	⨁◯◯◯ Very low	Analysis 1.1
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular leg pain												
Radicular leg pain excluded 2[g]	randomized trials	serious[h]	very serious[i]	not serious	very serious[j]	none	42	39	-	MD 8.71 lower (30.46 lower to 13.04 higher)	⨁◯◯◯ Very low	
Not specified whether participants had radicular leg pain 2[k]	randomized trials	serious[l]	very serious[m]	not serious	very serious[j,n]	none	27	31	-	MD 11.67 lower (35.87 lower to 12.53 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
High income 1[o]	randomized trials	serious[d]	not serious[p]	serious[q]	serious[r]	none	12	16	-	MD 0.9 higher (8.2 lower to 10 higher)	⨁◯◯◯ Very low	

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	Sham ultrasound	Relative (95% CI)	Absolute (95% CI)		
Low/middle income 3[s]	randomized trials	serious[l]	very serious[t]	not serious	very serious[j]	none	57	54	-	MD 13.86 lower (30.55 lower to 2.82 higher)	⨁◯◯◯ Very low	
Pain - short term (assessed with >=30% reduction)												
1	randomized trials	Serious[ac]	Not serious[p]	not serious	Serious[r]	none	128/233 (54.9%)	120/222 (54.1%)	RR 1.02 (0.86 to 1.20)	11 more per 1000 (from 76 fewer to 108 more)	⨁⨁◯◯ Low	
Pain - short term (assessed with >=50% reduction)												
1	randomized trials	Serious[ac]	Not serious[p]	not serious	Serious[r]	none	103/233 (44.2%)	90/222 (40.5%)	RR 1.09 (0.88 to 1.35)	36 more per 1000 (from 49 fewer to 142 more)	⨁⨁◯◯ Low	
Pain - intermediate term or long term – no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status - short term (assessed with: FRI, m-OSW, RMDQ)[g]												
4[v,w]	randomized trials	serious[x]	not serious[y]	not serious	serious[r]	none	280	266	-	SMD 0.23 SD lower (0.59 lower to 0.13 higher)	⨁⨁◯◯ Low	Analysis 1.7
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular leg pain												

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	Sham ultrasound	Relative (95% CI)	Absolute (95% CI)		
Radicular leg pain excluded [3z]	randomized trials	serious[aa]	not serious	not serious	serious[r]	none	47	44	-	SMD 0.46 SD lower (0.88 lower to 0.04 lower)	⨁⨁◯◯ Low	
Not specified whether participants had radicular leg pain [1ab]	randomized trials	serious[ac]	not serious[p]	not serious	serious[r]	none	233	222	-	SMD 0 SD (0.18 lower to 0.18 higher)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development												
High income [1ab]	randomized trials	serious[ac]	not serious[p]	serious[q]	serious[r]	none	233	222	-	SMD 0 SD (0.18 lower to 0.18 higher)	⨁◯◯◯ Very low	
Low/middle income [3z]	randomized trials	serious[aa]	not serious	not serious	serious[r]	none	47	44	-	SMD 0.46 SD lower (0.88 lower to 0.04 lower)	⨁⨁◯◯ Low	
Back-specific functional status - intermediate term or long term - no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term - no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health related quality of life - short term (assessed with: SF36 (general health); Scale from: 0 to 100)[l]												
2[ae]	randomized trials	serious[h]	not serious	not serious	serious[r]	none	254	243	-	MD 0.76 lower (5.1 lower to 3.59 higher)	⨁⨁◯◯ Low	Analysis 1.11

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	Sham ultrasound	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1 and 2 - not reported (no subgroup analysis performed)												
Population subgroup 3: presence of radicular leg pain												
Radicular leg pain excluded [1af]	randomized trials	serious[d]	not serious[p]	not serious	very serious[ag]	none	21	21	-	MD 3.09 higher (8.91 lower to 15.09 higher)	⨁◯◯◯ Very low	
Not specified whether participants had radicular leg pain [1ab]	randomized trials	serious[ac]	not serious[p]	not serious	serious[r]	none	233	222	-	MD 1.34 lower (6 lower to 3.32 higher)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development												
High income [1ab]	randomized trials	serious[ac]	not serious[p]	serious[q]	serious[r]	none	233	222	-	MD 1.34 higher (6 lower to 3.32 higher)	⨁◯◯◯ Very low	
Low/middle income [1af]	randomized trials	serious[d]	not serious[p]	serious[ah]	very serious[ag]	none	21	21	-	MD 3.09 higher (8.91 lower to 15.09 higher)	⨁◯◯◯ Very low	
Health-related quality of life - intermediate term or long term - no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events[n]												
[1ab]	randomized trials	serious[ac]	not serious[p]	not serious	very serious[n]	none	14/233 (6.0%)	13/222 (5.9%)	**RR 1.03** (0.49 to 2.13)	**2 more per 1.000** (from 30 fewer to 66 more)	⨁◯◯◯ Very low	Analysis 1.14

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	Sham ultrasound	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Serious adverse events[n]												
1[ab]	randomized trials	serious[ac]	not serious[p]	not serious	very serious[n]	none	3/233 (1.3%)	6/222 (2.7%)	**RR 0.48** (0.12 to 1.88)	**14 fewer per 1.000** (from 24 fewer to 24 more)	⨁◯◯◯ Very low	Analysis 1.15
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Psychological functioning (depression)- short term (assessed with: BDI; Scale from: 0 to 63)[p]												
1[af]	randomized trials	serious[d]	not serious[p]	not serious	serious[r]	none	21	21	-	MD 1.25 lower (5.71 lower to 3.21 higher)	⨁⨁◯◯ Low	Analysis 1.16
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Psychological functioning (depression)- long term - no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation -short term (assessed as lost one or more work days in past 4 weeks because of LBP)[r]												
1[ab]	randomized trials	serious[al]	not serious[p]	not serious	very serious[j]	none	14/112 (12.5%)	6/99 (6.1%)	**RR 2.06** (0.82 to 5.16)	**64 more per 1.000** (from 11 fewer to 252 more)	⨁◯◯◯ Very low	Analysis 1.17
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Social participation - intermediate term or long term - no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

CI: confidence interval; MD: mean difference; RR: risk ratio; SMD: standardized mean difference; VAS: visual analogue scale; FRI: Functional Rating Index; m-OSW: modified Oswestry scale; RMDQ: Roland Morris Disability Questionnaire; SD: standard deviation: SF36: Short Form 36; BDI: Beck Depression Inventory; LBP: Low back pain

Explanations

a. FU time between 2- 8 weeks
b. Durmus 2010a; Ebadi 2012; Grubisic 2006; Khan 2013
c. One study measured the outcome on an additional scale (Khan 2013): PRI at 4 weeks: n=30; mean difference -5.42, 95% CI (-7.40 to -3.44).
d. Risk of bias downgraded by 1 level due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of care providers, incomplete outcome data, selective reporting, co-interventions, and compliance with the intervention.
e. Inconsistency downgraded by 2 levels: considerable heterogeneity I^2>90%. Two studies showing little to no difference and two studies showing effects in favour of therapeutic ultrasound, not explained by pre-defined subgroups.
f. Imprecision downgraded by 1 level: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.
g. Durmus 2010a; Ebadi 2012
h. Risk of bias downgraded by 1 level due to unclear or high risk regarding randomisation sequence generation, allocation concealment, blinding of care providers, incomplete outcome data, selective reporting, co-interventions, and compliance.
i. Inconsistency downgraded by 2 levels: unexplained considerable heterogeneity I^2 = 91%
j. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.
k. Grubisic 2006; Khan 2013
l. Risk of bias downgraded by 1 level due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of care providers, incomplete outcome data, selective reporting, similar groups, co-interventions, and compliance.
m. Inconsistency downgraded by 2 levels: unexplained considerable heterogeneity I^2 = 95%
n. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for harm and low number of participants.
o. Grubisic 2006
p. Inconsistency not assessed because only one study included in this analysis.
q. Indirectness downgraded by 1 level: only one study included in this subgroup, unclear if it is representative of all high-income countries.
r. Imprecision downgraded by 1 level: low number of participants.
s. Durmus 2010a; Ebadi 2012; Khan 2013
t. Inconsistency downgraded by 2 levels: unexplained considerable heterogeneity I^2 = 93%
u. FU time between 3 - 12 weeks
v. Ansari 2006; Durmus 2010a; Ebadi 2012; Licciardone 2013
w. One study measured this outcome on an additional scale (Durmus 2010a): PDI at 3 weeks: n=42; mean difference 8.25, 95% CI (-0.67 to 17.17)
x. Risk of bias downgraded by 1 level due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of care providers, incomplete outcome data, selective reporting, co-interventions and compliance with the intervention.
y. Despite moderate heterogeneity (I^2 = 43%), not downgraded for inconsistency because this may be explained by subgroup analyses.
z. Ansari 2006; Durmus 2010a; Ebadi 2012
aa. Risk of bias downgraded by 1 level due to unclear or high risk regarding randomisation sequence generation, allocation concealment, blinding of care providers, dropouts, intention-to-treat, selective reporting, similar groups at baseline, co-interventions, and compliance.
ab. Licciardone 2013
ac. Risk of bias downgraded by 1 level due to high risk of bias regarding blinding of care providers.
ad. FU time 3 weeks and 12 weeks
ae. Durmus 2010a; Licciardone 2013
af. Durmus 2010a

ag. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for no effect and the possibility for harm and low number of participants.
ah. Indirectness downgraded by 1 level: only one study included in this subgroup, unclear if it is representative of all low/middle-income countries.
ai. FU time not specified
aj. FU time 3 weeks
ak. FU time 12 weeks
al. Risk of bias downgraded by 1 level due to high risk of bias regarding blinding of care providers and incomplete outcome data (no ITT analysis; outcome was assessed only in a subgroup of participants employed at baseline).

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***What are the benefits and harms of therapeutic ultrasound in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (assessed with: VAS at rest, NPRS; Scale from: 0 to 100)[a]												
5[b,c]	randomized trials	very serious[d]	serious[e]	not serious	very serious[f]	none	125	99	-	MD 18.56 lower (27.98 lower to 9.13 lower)	⨁◯◯◯ Very low	Analysis 2.1
Population subgroup 1: gender and/or sex												
Females 2[g]	randomized trials	very serious[h]	serious[i]	not serious	very serious[f]	none	70	44	-	MD 27.26 lower (48.42 lower to 6.1 lower)	⨁◯◯◯ Very low	
Mixed 3[j]	randomized trials	very serious[d]	not serious	not serious	very serious[f]	none	55	55	-	MD 12.2 lower (18.98 lower to 5.41 lower)	⨁◯◯◯ Very low	
Population subgroup 2: race/ethnicity (no subgroup analysis performed; no studies included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Radicular leg pain excluded 2[k]	randomized trials	very serious[d]	not serious	not serious	very serious[f]	none	35	35	-	MD 17.21 lower (24.7 lower to 9.7 lower)	⨁◯◯◯ Very low	

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
Not specified whether participants had radicular leg pain 3[l]	randomized trials	very serious[d]	serious[m]	not serious	very serious[f]	none	90	64	-	MD 19.7 lower (37.11 lower to 2.3 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
High income 1[n]	randomized trials	very serious[o]	not serious[p]	serious[q]	very serious[f]	none	15	15	-	MD 17.8 lower (32.55 lower to 3.05 lower)	⨁◯◯◯ Very low	
Low/middle income 4[r]	randomized trials	very serious[d]	serious[s]	not serious	very serious[f]	none	110	84	-	MD 18.81 lower (30.28 lower to 7.34 lower)	⨁◯◯◯ Very low	
Pain - intermediate term (assessed with: NPRS; Scale from: 0 to 100)[g]												
1[u]	randomized trials	very serious[v]	not serious[p]	not serious	serious[w]	none	17	17	-	MD 23.5 lower (30.68 lower to 16.32 lower)	⨁◯◯◯ Very low	Analysis 2.6
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Pain - long term - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
Population subgroups 1, 2, 3 and 4 - not reported												

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
Back-specific functional status - short term (assessed with: m-OSW, ODI, RMDQ)[a]												
6[x,y]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	144	119	-	SMD 0.48 SD lower (0.81 lower to 0.15 lower)	⨁◯◯◯ Very low	Analysis 2.7
Population subgroup 1: gender and/or sex												
Female 3[z]	randomized trials	very serious[d]	serious[aa]	not serious	serious[w]	none	89	64	-	SMD 0.39 SD lower (1.08 lower to 0.29 higher)	⨁◯◯◯ Very low	
Mixed 3[j]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	55	55	-	SMD 0.54 SD lower (0.92 lower to 0.16 lower)	⨁◯◯◯ Very low	
Population subgroup 2: race/ethnicity (no subgroup analysis performed; no studies included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Radicular leg pain excluded 3[ab]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	54	55	-	SMD 0.18 SD lower (0.55 lower to 0.2 higher)	⨁◯◯◯ Very low	

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
Not specified whether participants had radicular leg pain 3[l]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	90	64	-	SMD 0.75 SD lower (1.09 lower to 0.41 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
High income 1[n]	randomized trials	very serious[d]	not serious[p]	serious[q]	serious[w]	none	15	15	-	SMD 0.53 SD lower (1.26 lower to 0.2 higher)	⨁◯◯◯ Very low	
Low/middle income 5[ac]	randomized trials	very serious[d]	serious[ad]	not serious	serious[w]	none	129	104	-	SMD 0.46 SD lower (0.86 lower to 0.07 lower)	⨁◯◯◯ Very low	
Back-specific functional status - intermediate term (assessed with: ODI; Scale from: 0 to 100)[g]												
1[u]	randomized trials	very serious[v]	not serious[p]	not serious	very serious[f]	none	17	17	-	MD 9.12 lower (17.62 lower to 0.62 lower)	⨁◯◯◯ Very low	Analysis 2.12
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Back-specific functional status - long term - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term - not reported												

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	
Health related quality of life - short term (assessed with: SF36 (general health); Scale from: 0 to 100)[l]												
3[af]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	62	62	-	MD 0.46 lower (6.53 lower to 5.62 higher)	⨁◯◯◯ Very low	Analysis 2.13
Population subgroup 1: gender and/or sex												
Female 2[ag]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	39	39	-	MD 2.55 lower (9.61 lower to 4.52 higher)	⨁◯◯◯ Very low	
Mixed 1[ah]	randomized trials	very serious[d]	not serious[p]	not serious	very serious[ai]	none	23	23	-	MD 4.6 higher (6.47 lower to 15.67 higher)	⨁◯◯◯ Very low	
Population subgroup 2: race/ethnicity (no subgroup analysis performed; no studies included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Radicular leg pain excluded 2[ag]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	39	39	-	MD 2.55 lower (9.61 lower to 4.52 higher)	⨁◯◯◯ Very low	

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
Not specified whether participants had radicular leg pain [1ah]	randomized trials	very serious[d]	not serious[p]	not serious	very serious[ai]	none	23	23	-	MD 4.6 higher (6.47 lower to 15.67 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development (no subgroup analysis performed; all studies were carried out in low- or middle-income settings)												
Health-related quality of life - intermediate term or long term - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events												
1[aj]	randomized trials	very serious[v]	not serious[p]	not serious	very serious[ak]	none	0/20 (0.0%)	0/20 (0.0%)	not estimable		⨁◯◯◯ Very low	Analysis 2.16
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed; only one included study for this outcome)												
Serious adverse events - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term (assessed with: BDI; Scale from: 0 to 63)[r]												
2[ag]	randomized trials	very serious[d]	not serious	not serious	serious[w]	none	39	40	-	MD 0.83 lower (2.44 lower to 0.78 higher)	⨁◯◯◯ Very low	Analysis 2.17
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis performed)												
Psychological functioning (depression) - intermediate term or long term - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term - not reported												

Web Annex D.B6: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Therapeutic ultrasound	no intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-					-	

CI: confidence interval; MD: mean difference; RR: risk ratio; SMD: standardized mean difference; VAS: visual analogue scale; FRI: Functional Rating Index; m-OSW: modified Oswestry scale; RMDQ: Roland Morris Disability Questionnaire; SD: standard deviation: SF36: Short Form 36; BDI: Beck Depression Inventory; LBP: Low back pain

Explanations

a. FU time 3 - 12 weeks
b. Durmus 2013, Rubira 2019, Tantawy 2019, Tanveer 2022, Yurdakul 2019
c. One study measured the outcome on an additional scale (Rubira 2019): McGill at 4 weeks: n=74; MD -18.11, 95%CI (-27.25 to -8.97)
d. Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.
e. Inconsistency downgraded by 1 level: unexplained substantial heterogeneity I^2=71%
f. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.
g. Durmus 2013, Rubira 2019
h. Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, and compliance with the intervention.
i. Inconsistency downgraded by 1 level: unexplained considerable heterogeneity I^2 = 87%
j. Tantawy 2019, Tanveer 2022, Yurdakul 2019
k. Durmus 2013, Tantawy 2019
l. Rubira 2019, Tanveer 2022, Yurdakul 2019
m. Inconsistency downgraded by 1 level: unexplained considerable heterogeneity I^2 = 86%
n. Tantawy 2019
o. Risk of bias downgraded by 2 levels: due to unclear or high risk of bias regarding random sequence generation, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, and compliance with the intervention.
p. Inconsistency not assessed because only one study included in this analysis.
q. Indirectness downgraded by 1 level: only one study included in this subgroup, unclear if it is representative of all high-income countries.
r. Durmus 2013, Rubira 2019, Tanveer 2022, Yurdakul 2019
s. Inconsistency downgraded by 1 level: unexplained substantial heterogeneity I^2=78%
t. FU time 20 weeks
u. Tanveer 2022
v. Risk of bias downgraded by 2 levels: due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, selective reporting, co-interventions, and compliance with the intervention.
w. Imprecision downgraded by 1 level: low number of participants.
x. Durmus 2010b, Durmus 2013, Rubira 2019, Tantawy 2019, Tanveer 2022, Yurdakul 2019
y. Three studies measured the outcome on an additional scale: PDI at 6-8 weeks: Durmus 2010b (n=39): MD -0.29, 95% CI (-3.07 to 2.49); Durmus 2013 (n=40): MD -0.10, 95% CI (-2.9 to 2.7); Tantawy 2019 n=30: MD -6.4, 95% CI (-15.14 to 2.34)
z. Durmus 2010b, Durmus 2013, Rubira 2019

aa. Inconsistency downgraded by 1 level: unexplained substantial heterogeneity I^2=76%
ab. Durmus 2010b, Durmus 2013, Tantawy 2019
ac. Durmus 2010b, Durmus 2013, Rubira 2019, Tanveer 2022, Yurdakul 2019
ad. Inconsistency downgraded by 1 level: unexplained heterogeneity I^2=52%
ae. FU time 3-6 week
af. Durmus 2010b, Durmus 2013, Yurdakul 2019
ag. Durmus 2010b, Durmus 2013
ah. Yurdakul 2019
ai. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for harm and the possibility for no effect and low number of participants.
aj. Durmus 2013
ak. Imprecision downgraded by 2 levels: no events in either group
al. FU time 6 weeks.

GRADE Table 3. ***What are the benefits and harms of therapeutic ultrasound in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

No trials

B.7 Transcutaneous electrical nerve stimulation (TENS)

Overview of the PICO structure

Definition of the intervention	
TENS is a non-invasive peripheral electrical stimulation modality applied to the skin using surface electrodes. TENS uses low-voltage electrical currents between the electrodes to modify the perception of pain, acting through segmental inhibition or activation of descending nociceptive-inhibitory systems. TENS devices may be used in health facilities or may be portable for use at home. A range of stimulation parameters may be selected, based on clinical indication, including pulse intensity, frequency, duration and type (burst or continuous). Among the included trials used to inform the guideline, TENS interventions involved electrode placement over the paravertebral lumbosacral area and sometimes the affected leg in the case of associated leg pain, using conventional continuous or burst pulse parameters.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; uncertain	Small; uncertain
Harms	Small; uncertain	Small; uncertain
Balance benefits to harms	Uncertain	Uncertain
Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate costs; high costs; varies (according to country and health system)	Moderate costs; high costs; varies (according to country and health system)
Equity and human rights	No impact; probably reduced; varies	No impact; probably reduced; varies
Acceptability	Probably yes; uncertain; varies	Probably yes; uncertain; varies
Feasibility	Probably yes	Probably yes

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of TENS in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with sham?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
9[a]	randomized trials	very serious[1,2,3,4,5,6,7,8,b]	serious[c]	not serious[d]	serious[e]	none	280	187	-	MD **0.9 lower** (1.54 lower to 0.26 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in females (follow-up: closest to 2 weeks; assessed with: Borg scale; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	very serious[5,b]	not serious[g]	serious[h]	serious[i]	none	23	21	-	MD **0.1 higher** (0.2 lower to 0.4 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in females and males (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
8	randomized trials	very serious[b]	serious[k]	not serious[d]	serious[l]	none	257	187	-	MD **1.03 lower** (1.69 lower to 0.36 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in people without leg pain (follow-up: closest to 2 weeks weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
5	randomized trials	very serious[1,2,4,5,8,b]	serious[m]	not serious[d]	serious[n]	none	129	102	-	MD **0.64 lower** (1.83 lower to 0.54 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
2[o]	randomized trials	very serious[3,7,b]	not serious[p]	not serious[q]	serious[l]	none	100	47	-	MD **1.34 lower** (2.44 lower to 0.25 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in people with mixed radicular and non-radicular leg pain (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
2[r]	randomized trials	very serious[6,10,b]	very serious[s]	not serious[q]	very serious[t]	none	51	38	-	MD **0.96 lower** (4.59 lower to 2.67 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials undertaken in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
8[u]	randomized trials	very serious[1,2,3,5,6,7,8,10,b]	serious[v]	not serious[d]	serious[l]	none	219	125	-	MD **1.01 lower** (1.69 lower to 0.34 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in trials undertaken in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	serious[4,w]	not serious[g]	serious[x]	serious[j]	none	30	32	-	MD **0** (0.4 lower to 0.4 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials using a single TENS treatment session (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
4[y]	randomized trials	very serious[1,3,4,6,b]	very serious[z]	not serious[d]	serious[n]	none	135	90	-	MD **0.68 lower** (2 lower to 0.65 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials using 10-20 TENS treatment sessions (follow-up: closest to 2 weeks; assessed with: VAS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
5[aa]	randomized trials	very serious[2,5,7,8,10,b]	serious[ab]	not serious[q]	serious[l]	none	145	97	-	MD **1.06 lower** (1.94 lower to 0.18 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (after removing high risk of bias trials) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
2	randomized trials	serious[4,8,ac]	serious[ad]	not serious[d]	very serious[t]	none	80	55	-	MD **0.63 lower** (2.78 lower to 1.53 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (follow-up: closest to 3 months; assessed with: VAS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
2[ae]	randomized trials	very serious[5,8,af]	serious[ag]	not serious[q]	very serious[t]	none	73	44	-	MD **0.4 lower** (2.21 lower to 1.41 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in females (follow-up: closest to 3 months; assessed with: Borg scale; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	very serious[5,af]	not serious[g]	serious[h]	serious[i]	none	23	21	-	MD **0.1 higher** (0.23 lower to 0.43 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in females and males (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												
1[ae]	randomized trials	serious[8,w]	not serious[g]	serious[h]	very serious[ah]	none	50	23	-	MD **1.06 lower** (4.23 lower to 2.12 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (after removing high risk of bias trials) (follow-up: closest to 3 months; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
1[ae]	randomized trials	serious[8,w]	not serious[g]	serious[h]	very serious[ah]	none	50	23	-	MD **1.06 lower** (4.23 lower to 2.12 higher)	⨁◯◯◯ Very low	CRITICAL
Pain stratified by race/ethnicity												
0												
Function (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
4[ai]	randomized trials	very serious[2,5,7,10,b]	very serious[aj]	not serious[q]	very serious[ak]	none	95	74	-	SMD **0.96 SD lower** (3.2 lower to 1.28 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females and males (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
3	randomized trials	very serious[2,7,10,b]	very serious[aj]	not serious[q]	very serious[ak]	none	72	53	-	SMD **1.3 lower** (4.38 lower to 1.78 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
1	randomized trials	very serious[5,af]	not serious[g]	serious[h]	very serious[ah]	none	23	21	-	SMD **0.27 higher** (0.33 lower to 0.86 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people with no leg pain (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
2	randomized trials	very serious[2,5,b]	not serious[p]	not serious[q]	very serious[al]	none	34	32	-	SMD **0.16 higher** (1.19 lower to 1.51 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
Function in people either with or without radicular leg pain (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
1	randomized trials	very serious[10,b]	not serious[g]	serious[h]	serious[am]	none	31	30	-	SMD **1.97 lower** (2.59 lower to 1.36 lower)	⨁◯◯◯ Very low	CRITICAL
Function in people with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: ODI; benefit indicated by lower values)												
1[ai]	randomized trials	very serious[7,b]	not serious[g]	serious[h]	very serious[al]	none	30	12	-	SMD **1.67 higher** (28.66 lower to 25.33 higher)	⨁◯◯◯ Very low	CRITICAL
Function (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
2[ae]	randomized trials	very serious[5,8,af]	serious[an]	not serious[q]	serious[ao]	none	73	44	-	MD **0.24 lower** (4.3 lower to 3.81 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
1	randomized trials	very serious[5,b]	not serious[g]	serious[h]	serious[ao]	none	23	21	-	MD **0.5 higher** (1.22 lower to 2.22 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females and males (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values; scale: 0 to 50)												
1[ae]	randomized trials	serious[8,w]	not serious[g]	serious[h]	serious[ap]	none	50	23	-	MD **2.61 lower** (6.42 lower to 1.2 higher)	⨁◯◯◯ Very low	CRITICAL
Function (after removing high risk of bias trials) (follow-up: closest to 3 months; assessed with: ODI; scale: 0 to 50)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
1[ae]	randomized trials	serious[8,w]	not serious[g]	serious[h]	serious[ao]	none	50	23	-	MD **2.61 lower** (6.42 lower to 1.2 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by race/ethnicity, number of treatment sessions or in adults in low- or lower middle-income countries not identified												
Health-related quality of life (follow-up: closest to 2 weeks; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
2[ai]	randomized trials	very serious[2,7,b]	serious[an]	not serious[q]	very serious[aq]	none	41	23	-	MD **3.21 higher** (21.17 lower to 27.59 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in people with no radicular leg pain (follow-up: closest to 2 weeks; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1	randomized trials	very serious[2,b]	not serious[g]	serious[h]	very serious[ar]	none	11	11	-	MD **20.45 lower** (56.67 lower to 15.77 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in people with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1[ai]	randomized trials	very serious[7,b]	not serious[g]	serious[h]	serious[as]	none	30	12	-	MD **5.91 higher** (0.44 lower to 12.26 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (follow-up: closest to 2 weeks; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
2[ai]	randomized trials	very serious[2,7,b]	very serious[at]	serious[h]	serious[as]	none	41	23	-	MD **3.57 higher** (30.06 lower to 37.2 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in people with no radicular leg pain (follow-up: closest to 2 weeks; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[2,b]	not serious[g]	serious[h]	serious[au]	none	11	11	-	MD **11.63 lower** (20.59 lower to 2.67 lower)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life in people with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
1[ai]	randomized trials	very serious[7,b]	not serious[g]	serious[h]	serious[l]	none	30	12	-	MD **11.63 higher** (9.96 higher to 13.31 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity, number of treatment sessions or in adults in low- or lower middle-income countries not identified												
Depression (follow-up: closest to 3 months; assessed with: BDI; benefit indicated by lower values; scale: 0 to 63)												
1[ae]	randomized trials	serious[8,w]	not serious[g]	serious[h]	very serious[av]	none	50	23	-	MD **3.04 higher** (19.15 lower to 25.22 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on depression stratified by gender, race/ethnicity, number of treatment sessions, national economic development or presence of leg pain not identified												
Trials on fear avoidance, catastrophizing, anxiety or self-efficacy not identified												
Adverse events/harms (high-income country, no leg pain)												
1	randomized trials	serious[8,w]	not serious[g]	serious[h]	serious[aw]	none	Authors reported that no TENS-associated adverse events developed in any participants.				⨁◯◯◯ Very low	CRITICAL
Trials on adverse events/harms stratified by gender, race/ethnicity, number of treatment sessions, presence of leg pain or in adults in low- or lower middle-income countries not identified												
Trials on social participation not identified												
0												

<u>OLDER ADULTS (aged 60 years or more)</u>

Pain (follow-up: closest to 2 weeks; assessed with: NRS; benefit indicated by lower values; scale: 0 to 10)

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	Sham	Relative (95% CI)	Absolute (95% CI)		
1[r]	randomized trials	very serious[6,b]	not serious[g]	serious[h]	very serious[t]	none	20	8	-	MD **0.13 higher** (9.8 lower to 10.06 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by gender, race/ethnicity, number of treatment sessions, national economic development or presence of leg pain not identified												
0												
Trials on function, health-related quality of life, depression, fear avoidance, catastrophizing, anxiety, self-efficacy, social participation, change in use of medications, falls or adverse events/harms not identified												
0												

BDI: Beck Disabiltiy Index; **CI:** confidence interval; **MCS:** Mental Component Summary; **MD:** mean difference; **MPQ:** McGill Pain Questionnaire; **NRS:** numeric rating scale; **ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size; **PCS:** Physical Component Summary; **RMDQ:** Roland-Morris Disability Questionnaire; **SMD:** standardized mean difference; **VAS:** visual analogue scale

The following was used to guide the ratings:
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).
Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.
Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.
Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).
Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations
a. Four trials had 2 arms: Dias 2021 (TENS (GT100Hz) vs. sham, TENS (GT2Hz) vs. sham), Topuz 2004 (conventional TENS vs. sham, low-frequency TENS vs. sham), Yaksi 2021 (burst TENS vs. sham, conventional TENS vs. sham) and Shimoji 2007 (TENS bidirectional modulated sine wave vs. sham, TENS conventional bidirectional pulsed wave vs. sham). For each of these 4 trials we included both arms in meta-analysis and split the comparison groups in half. One trial reporting only p-values was not included in meta-analysis (Bloodworth 2004); results were reported narratively and graded. In this cross-over design, 11 participants with radiculopathy received 4 different TENS interventions and 2 placebo TENS interventions in random order in a single day. Only p-values were provided. Trial authors reported no significant differences between groups (stochastic resonance TENS on back/leg vs. sham, p=0.096; conventional TENS on back/leg vs. sham, p=0.519).
b. Risk of bias: We downgraded twice. Most or all of the trials were rated as overall high risk of bias.
c. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 77%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
d. Indirectness: We did not downgrade. Multiple trials are included from different countries both high- and lower-middle income.
e. Imprecision: We downgraded once due to small sample size (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -1); the confidence interval does not cross the null but the lower boundary crosses the threshold for what may be considered appreciable benefit (MD = -1).
f. Risk of bias: We downgraded twice due to unclear items related to selection and reporting bias.
g. Inconsistency: We did not downgrade; however, there are no additional trials with which to compare these findings.
h. Indirectness: We downgraded once. This is a single trial from a single centre (high or upper-middle income country).

i. Imprecision: We downgraded twice due to low sample size (the OIS would not have been reached).
j. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -1); the confidence interval crossed the null but not the thresholds for what may be considered appreciable benefit (-1) or harm (+1).
k. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 73%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
l. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate reached the pre-specified threshold for what may be considered appreciable benefit (MD = -1). The confidence interval did not cross the null.
m. Inconsistency: We downgraded once. There is similarity in the majority of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 74%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
n. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -1). The confidence interval crossed the null and the lower boundary crossed the threshold for what may be considered appreciable benefit (MD = -1).
o. These trials had 2 arms each: Dias 2021 (TENS (GT100Hz) vs. sham, TENS (GT2Hz) vs. sham), Topuz 2004 (conventional TENS vs. sham, low-frequency TENS vs. sham).
p. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
q. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).
r. Shimoji 2007 included 2 arms (TENS bidirectional modulated sine wave vs. sham, TENS conventional bidirectional pulsed wave vs. sham). Both were included in meta-analysis and the comparison group was split in half.
s. Inconsistency: We downgraded twice. The point estimates differ with some overlap in confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 72%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
t. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD = -1); the confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds of appreciable benefit (MD = -1) and harm (MD = +1).
u. Four trials had 2 arms: Dias 2021 (TENS (GT100Hz) vs. sham, TENS (GT2Hz) vs. sham), Topuz 2004 (conventional TENS vs. sham, low-frequency TENS vs. sham), Yaksi 2021 (burst TENS vs. sham, conventional TENS vs. sham), and Shimoji 2007. For each of these 4 trials we included both arms in meta-analysis and split the comparison groups in half.
v. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 78%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
w. Risk of bias: We downgraded once due to the potential for selection and performance bias.
x. Indirectness: We downgraded once. This is a single trial from a single centre (low or lower-middle income country).
y. Two trials included 2 arms (Dias 2021: (TENS (GT100Hz) vs. sham, TENS (GT2Hz) vs. sham); and Shimoji 2007. All arms were included in the meta-analyses by splitting the comparison groups in half.
z. Inconsistency: We downgraded twice. Some estimates differ in direction. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 64%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
aa. Two trials had 2 arms each (Topuz 2004: conventional TENS vs. sham, low-frequency TENS vs. sham; Yaksi 2021: burst TENS vs. sham, conventional TENS vs. sham). For each of these 2 trials we included both arms in meta-analysis and split the comparison groups in half.
ab. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 84%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
ac. Risk of bias: We downgraded once. Items were rated as unclear in the selection, performance and reporting domains.
ad. Inconsistency: We downgraded once. There is similarity in some of the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 70%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
ae. Yaksi 2021 had 2 arms (burst TENS vs. sham, conventional TENS vs. sham); both arms were included in the meta-analysis with the comparison group split in half.
af. Risk of bias: We downgraded twice due to the potential for selection, performance and reporting biases.
ag. Inconsistency: We downgraded once. There is similarity in the point estimates with overlapping confidence intervals. Statistical heterogeneity is between 30% and 60% (i.e., I2 = 50%); this could not be explained due to small subgroups and may represent moderate heterogeneity.
ah. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD = -1); the confidence interval crosses the null.
ai. Topuz 2004 had 2 arms(conventional TENS vs. sham, low-frequency TENS vs. sham); both were included in the meta-analysis and the comparison group was split in half.
aj. Inconsistency: We downgraded twice. The results are in different directions with some non-overlapping confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 92%); this could not be explained due to small subgroups and may represent considerable heterogeneity.
ak. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important (MD = -0.2); the confidence interval crosses the null.
al. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD = -0.2); the confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds of appreciable benefit (MD = -0.2) and harm (MD = +0.2).
am. Imprecision: We downgraded once due to low sample size (the OIS would not have been reached).
an. Inconsistency: We downgraded once. The point estimates are in different directions with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).

ao. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -5); the confidence interval crossed the null but not the thresholds for what may be considered appreciable benefit (-5) or harm (+5).
ap. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -5). The confidence interval crossed the null; the lower boundary crossed the threshold for what may be considered appreciable benefit (-5).
aq. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD = +10); the confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds of appreciable benefit (MD = +10) and harm (MD = -10).
ar. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important favouring the comparison (MD = -10); the confidence interval crossed the null.
as. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = +10); the confidence interval crossed the null.
at. Inconsistency: We downgraded twice. The point estimates differ in direction and the confidence intervals do not overlap. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 87%); this could not be explained due to small subgroups and may represent substantial heterogeneity.
au. Imprecision: We downgraded once. The sample size was small (OIS would not have been reached). The pointe estimate reached the threshold for what may be considered clinically important favouring the comparison (MD = -10); the confidence interval did not cross the null.
av. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD = -6.3). The confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds of appreciable benefit (MD = -6.3) and harm (MD = +6.3).
aw. Imprecision: We downgraded once due to low sample size (the OIS would not have been reached).

References

1.Cheing GL, Hui-Chan CW. Transcutaneous electrical nerve stimulation: nonparallel antinociceptive effects on chronic clinical pain and acute experimental pain. Arch Phys Med Rehabil; 1999.
2.de Alencar Caldas VV, Maciel DG,Cerqueira MS,et al. Effect of Pain Education, Cryotherapy, and Transcutaneous Electrical Nerve Stimulation on the Pain, Functional Capacity, and Quality of Life in Patients With Nonspecific Chronic Low Back Pain: A Single-Blind Randomized Controlled Trial. Am J Phys Med Rehabil; 2021.
3.Dias LV, Cordeiro MA,Schmidt de Sales R,et al. Immediate analgesic effect of transcutaneous electrical nerve stimulation (TENS) and interferential current (IFC) on chronic low back pain: Randomized placebo-controlled trial. Journal of Bodywork & Movement Therapies; 2021.
4.Ezema CI, Onyeso OK,Nna EO,et al. Transcutaneous electrical nerve stimulation effects on pain-intensity and endogenous opioids levels among chronic low-back pain patients: A randomized controlled trial. J Back Musculoskelet Rehabil; 2022.
5.Kofotolis ND, Vlachopoulos SP,Kellis E. Sequentially allocated clinical trial of rhythmic stabilization exercises and TENS in women with chronic low back pain. Clin Rehabil; 2008.
6.Shimoji K, Takahashi N,Nishio Y,Koyanagi M,Aida S. Pain relief by transcutaneous electric nerve stimulation with bidirectional modulated sine waves in patients with chronic back pain: a randomized, double-blind, sham-controlled trial. Neuromodulation; 2007.
7.Topuz O, Ozfidan E,Ozgen M,Ardic F. Efficacy of transcutaneous electrical nerve stimulation and percutaneous neuromodulation therapy in chronic low back pain. Journal of Back and Musculoskeletal Rehabilitation; 2004.
8.Yaksi E, Ketenci A,Baslo MB,Orhan EK. Does transcutaneous electrical nerve stimulation affect pain, neuropathic pain, and sympathetic skin responses in the treatment of chronic low back pain? A randomized, placebo-controlled trial. Korean J Pain; 2021.
9.Bloodworth DM, Nguyen BN,Garver W,et al. Comparison of stochastic vs. conventional transcutaneous electrical stimulation for pain modulation in patients with electromyographically documented radiculopathy. Am J Phys Med Rehabil; 2004.
10.Kibar S, Konak HE,Ay S,Doganay Erdogan B,Evcik D. The effectiveness of combined transcutaneous electrical nerve stimulation and interferential current therapy on chronic low back pain: a randomized, double-blind, sham-controlled trial. Fiziksel Tip ve Rehabilitasyon Bilimleri Dergisi [Journal of Physical Medicine and Rehabilitation Sciences]; 2020.

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***What are the benefits and harms of TENS in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no treatment or treatments where the effect of TENS could be isolated?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
ALL ADULTS												
Pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
8	randomized trials	very serious[1,2,3,4,5,6,7,8,a,b]	not serious[c]	not serious[d]	serious[e]	none	192	146	-	MD **0.19 lower** (0.51 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in females and males (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
7	randomized trials	very serious[1,2,3,4,5,7,8,b]	not serious[c]	not serious[d]	serious[f]	none	171	123	-	MD **0.35 lower** (0.66 lower to 0.03 lower)	⨁◯◯◯ Very low	CRITICAL
Pain in females (follow-up: closest to 2 weeks; assessed with: Borg scale; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	very serious[6,b]	not serious[g]	serious[h]	serious[e]	none	21	23	-	MD **0.2 higher** (0.07 lower to 0.47 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people without leg pain (follow-up: closest to 2 weeks; assessed with: VAS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
4	randomized trials	very serious[2,6,7,8,a,b]	not serious[i]	not serious[d]	serious[e]	none	122	79	-	MD **0** (0.42 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
Pain in people with unclassified presence of leg pain (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
2	randomized trials	very serious[1,3,b]	not serious[i]	not serious[d]	serious[j]	none	27	27	-	MD **0.18 higher** (0.12 higher to 0.24 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in people with and without leg pain (radicular or non-radicular) (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
2	randomized trials	very serious[4,5,b]	serious[k]	not serious[l]	very serious[m]	none	43	40	-	MD **0.48 lower** (5.31 lower to 4.35 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials undertaken in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
6	randomized trials	very serious[1,4,5,6,7,8,b]	not serious[n]	not serious[l]	serious[e]	none	151	120	-	MD **0.15 lower** (0.49 lower to 0.19 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials undertaken in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: VAS, NRS; benefit indicated by lower values; scale: 0 to 10)												
2	randomized trials	very serious[2,3,b,o]	not serious[p]	not serious[q]	very serious[m]	none	41	26	-	MD **0.53 lower** (3 lower to 1.95 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials using 10-20 TENS treatment sessions (follow-up: closest to 2 weeks; assessed with: VAS, NRS, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
6	randomized trials	very serious[2,3,4,5,6,7,b,o]	not serious[r]	not serious[d]	serious[e]	none	116	100	-	MD **0.21 lower** (0.72 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL
Pain in trials using <10 TENS treatment sessions (follow-up: closest to 2 weeks; assessed with: VAS; benefit indicated by lower values; scale: 0 to 10)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
2[s]	randomized trials	very serious[1,8,b]	not serious[i]	not serious[l]	serious[e]	none	76	46	-	MD **0.04 higher** (0.3 lower to 0.38 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (high-income country) (follow-up: closest to 3 months; assessed with: Brief Pain Inventory, Borg scale; benefit indicated by lower values; scale: 0 to 10)												
2	randomized trials	very serious[6,9,t,u]	very serious[v]	not serious[l]	very serious[m]	none	50	54	-	MD **0.98 lower** (16.83 lower to 14.88 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (females and males, either with or without radicular or non-radicular leg pain) (follow-up: closest to 3 months; assessed with: Brief Pain Inventory; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	serious[9,t]	not serious[g]	serious[h]	serious[w]	none	29	31	-	MD **2.3 SD lower** (3.51 lower to 1.09 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (females, no leg pain) (follow-up: closest to 3 months; assessed with: Borg Scale; benefit indicated by lower values; scale: 0 to 10)												
1	randomized trials	very serious[6,b]	not serious[g]	serious[h]	serious[f]	none	21	23	-	MD **0.2 higher** (0.01 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on pain stratified by race/ethnicity, number of treatment sessions or in adults in low- or lower middle-income countries not identified												
0												
Function (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
6	randomized trials	very serious[1,2,3,4,7,10,b,o]	not serious[x]	not serious[d]	serious[y]	none	108	91	-	SMD **0.32 lower** (0.71 lower to 0.07 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females (follow-up: closest to 2 weeks; assessed with: modified ODI; benefit indicated by lower values)												
1	randomized trials	very serious[10,b]	not serious[g]	serious[z]	very serious[aa]	none	8	8	-	SMD **0.29 lower** (1.28 lower to 0.69 higher)	⨁◯◯◯ Very low	CRITICAL
Function in females and males (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
5	randomized trials	very serious[1,2,3,4,7,b]	not serious[ab]	not serious[d]	serious[y]	none	100	83	-	SMD **0.32 lower** (0.78 lower to 0.15 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people without leg pain (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
3	randomized trials	very serious[2,7,10,b,o]	not serious[i]	not serious[d]	serious[ac]	none	49	34	-	SMD **0.15 lower** (0.37 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people with unclassified presence of leg pain (follow-up: closest to2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomized trials	very serious[b,o]	not serious[i]	not serious[d]	very serious[ad]	none	27	27	-	SMD **0.08 lower** (0.74 lower to 0.58 higher)	⨁◯◯◯ Very low	CRITICAL
Function in people either with or without radicular leg pain (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
1	randomized trials	very serious[4,b]	not serious[g]	serious[h]	serious[w]	none	32	30	-	SMD **1.03 lower** (1.56 lower to 0.49 lower)	⨁◯◯◯ Very low	CRITICAL
Function in trials undertaken in low- or lower middle-income countries (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												
3	randomized trials	very serious[2,3,10,b,o]	not serious[i]	not serious[q]	serious[ae]	none	49	34	-	SMD **0.16 lower** (0.36 lower to 0.03 higher)	⨁◯◯◯ Very low	CRITICAL
Function in trials undertaken in high to upper-middle income countries (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values)												
3	randomized trials	very serious[1,4,7,b]	serious[af]	not serious[ag]	very serious[aa]	none	59	57	-	SMD **0.47 lower** (1.94 lower to 1 higher)	⨁◯◯◯ Very low	CRITICAL
Function in trials using 10-20 TENS treatment sessions (follow-up: closest to 2 weeks; assessed with: ODI, RMDQ; benefit indicated by lower values)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
5	randomized trials	very serious[2,3,4,7,10,b,o]	not serious[ah]	not serious[d]	serious[ai]	none	92	75	-	SMD **0.35 lower** (0.82 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL
Function in trials using <10 treatment sessions (follow-up: closest to 2 weeks; assessed with: RMDQ; benefit indicated by lower values; scale: 0 to 24)												
1	randomized trials	very serious[1,aj,b]	not serious[g]	serious[h]	very serious[ad]	none	16	16	-	SMD **0.12 lower** (0.82 lower to 0.57 higher)	⨁◯◯◯ Very low	CRITICAL
Function (high-income country) (follow-up: closest to 3 months; assessed with: ODI, PDI; benefit indicated by lower values)												
2	randomized trials	very serious[6,9,b]	very serious[ak]	serious[h]	very serious[aa]	none	50	54	-	SMD **1.05 higher** (18.51 lower to 20.61 higher)	⨁◯◯◯ Very low	CRITICAL
Function (females, no leg pain) (follow-up: closest to 3 months; assessed with: ODI; benefit indicated by lower values)												
1	randomized trials	very serious[6,b]	not serious[g]	serious[h]	serious[w]	none	21	23	-	SMD **2.6 higher** (1.78 higher to 3.42 higher)	⨁◯◯◯ Very low	CRITICAL
Function (females and males, either with or without radicular or non-radicular leg pain) (follow-up: closest to 3 months; assessed with: PDI; benefit indicated by lower values)												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious[9,t]	not serious[g]	serious[h]	serious[y]	none	29	31	-	SMD **0.48 lower** (0.99 lower to 0.04 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on function stratified by race/ethnicity not identified												
0												
Health-related quality of life (no leg pain, high-income country) (follow-up: closest to 2 weeks; assessed with: SF-36 (PCS); benefit indicated by higher values; scale: 0 to 100)												
1	randomized trials	very serious[7,b]	not serious[g]	serious[h]	very serious[al]	none	11	11	-	MD **6.82 lower** (27.06 lower to 13.42 higher)	⨁◯◯◯ Very low	CRITICAL
Health-related quality of life (no leg pain, high-income country) (follow-up: closest to 2 weeks; assessed with: SF-36 (MCS); benefit indicated by higher values; scale: 0 to 100)												
1	randomized trials	very serious[7,b]	not serious[g]	serious[h]	serious[am]	none	11	11	-	MD **2.91 lower** (10.25 lower to 4.43 higher)	⨁◯◯◯ Very low	CRITICAL
Trials on health-related quality of life stratified by gender, race/ethnicity, presence of leg pain or in adults in low- or lower middle-income countries not identified												
0												
Depression (either with or without radicular or non-radicular leg pain, high-income country) (follow-up: closest to 3 months; assessed with: HADS; benefit indicated by lower values; scale: 0 to 21)												
1	randomized trials	serious[9,t]	not serious[g]	serious[h]	very serious[an]		29	31	-	MD **1.4 lower** (5.57 lower to 2.77 higher)	-	CRITICAL
Trials on depression stratified by gender, race/ethnicity, presence of leg pain or in adults in low- or lower middle-income countries not identified												

Web Annex D.B7: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of trials	Trial design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	TENS	No treatment	Relative (95% CI)	Absolute (95% CI)		
0												
Catastrophizing (either with or without radicular or non-radicular leg pain, high-income country) (follow-up: closest to 3 months; assessed with: Pain Catastrophizing Scale; benefit indicated by lower values; scale: 0 to 52)												
1	randomized trials	serious[9,t]	not serious[g]	serious[h]	serious[w]	none	29	31	-	MD **11.2 lower** (17.88 lower to 4.52 lower)	⨁◯◯◯ Very low	CRITICAL
Trials on catastrophizing stratified by gender, race/ethnicity, presence of leg pain or in adults in low- or lower middle-income countries not identified												
0												
Trials on fear avoidance, anxiety, self-efficacy or social participation not identified												
0												
Adverse events/harms (high-income country, either with or without leg pain (radicular or non-radicular)												
1	randomized trials	serious[t]	not serious[g]	serious[h]	serious[w]	none	Authors reported that none of the participants reported experiencing any long-term adverse events from using high-frequency TENS.				⨁◯◯◯ Very low	CRITICAL
Trials on adverse events/harms stratified by gender, race/ethnicity or in adults in low- or lower middle-income countries not identified												
0												
OLDER ADULTS (aged 60 years or more)												
Trials in older adults on pain, function, health-related quality of life, depression, fear avoidance, catastrophizing, anxiety, self-efficacy, social participation, adverse events, change in use of medications or falls not identified												
0												

BDI: Beck Disabiltiy Index; **CI:** confidence interval; **MCS:** Mental Component Summary; **MD:** mean difference; **MPQ:** McGill Pain Questionnaire; **NRS:** numeric rating scale; **ODI:** Oswestry Disability Index; **OIS:** Optimal Information Size; **PCS:** Physical Component Summary; **PDI:** Pain Disability Index; **RMDQ:** Roland-Morris Disability Questionnaire; **SMD:** standardized mean difference; **VAS:** visual analogue scale

The following was used to guide the ratings:
Risk of bias: *Not serious:* all or most of the weight (>50%) comes from overall low risk of bias trial(s). *Serious:* some of the weight (<50%) comes from overall low risk of bias trial(s). *Very serious:* all or most of the weight (>50%) comes from overall high or unclear risk of bias trial(s).
Inconsistency: *Not serious:* high extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 0% and 40%, which might not be important. *Serious:* some extent of similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 30% and 60%, which could not be explained due to small subgroups and may represent moderate

heterogeneity. *Very serious:* little or no similarity of point estimates and overlap of confidence intervals; statistical heterogeneity (I^2) is between 50% and 90% or 75% and 100%, which could not be explained due to small subgroups and may represent substantial or considerable heterogeneity, respectively.

Indirectness: *Not serious:* trial(s) were conducted in different countries or settings. *Serious:* trial(s) were conducted from a single country/setting. *Very serious:* evidence is not directly related to PICO question.

Imprecision: *Not serious:* Optimal Information Size (OIS) was reached (i.e., sample sizes with at least 200 participants per group may provide prognostic balance); and the entire confidence interval lies on one side of the threshold that may be considered clinically important (≥10% scale range or SMD ≥0.2 for continuous variables, ≥10% for binary variables), such that the clinical course of action would not differ if the upper versus the lower boundary of the confidence interval represented the truth. *Serious:* OIS would not have been reached (sample sizes with less than 200 participants per group); if the OIS was reached, the clinical course of action might differ if the upper versus the lower boundary of the confidence interval represented the truth. *Very serious:* similar to 'serious' but to a greater extent (e.g., very small sample sizes and confidence intervals crossing appreciable benefit and harm).

Other considerations: *Not serious:* Publication bias is undetected. *Serious/very serious:* Publication bias is strongly suspected.

Explanations

a. Elserty 2016 included 2 arms (fixed pulse TENS + exercise vs. exercise; adjusted pulse TENS + exercise vs. exercise). Both were included in meta-analysis by splitting the comparison group number in half. Petrofsky 2020 included 4 arms (Continuous TENS + spent sham heat vs. spent sham heat; continuous TENS + LLCH (low-level continuous heat) vs. LLCH; TENS last 15 min + LLCH vs. LLCH; TENS last 15 min + spent sham heat vs. spent sham heat). All were included in meta-analysis by splitting the comparison group numbers accordingly.
b. Risk of bias: We downgraded twice. Most or all of the trials were rated as overall high risk of bias.
c. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 6%).
d. Indirectness: We did not downgrade. Trials are included from different countries both high- and lower-middle income.
e. Imprecision: We downgraded once. The sample size is small (OIS would not have been achieved). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -1). The confidence interval crossed the null but not the thresholds for what may be considered appreciable benefit (-1) or harm (+1).
f. Imprecision: We downgraded once. The sample size was small (OIS would not have been reached). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD = -1); the confidence interval did not cross the null.
g. Inconsistency: We did not downgrade; however, there are no additional trials with which to compare these findings.
h. Indirectness: We downgraded once. This is a single trial from a single centre (high or upper-middle income country).
i. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 0%).
j. Imprecision: We downgraded once. The sample size was small (OIS would not have been reached). The point estimate did not reach the threshold for what may be considered clinically important (MD = -1); the confidence interval did not cross the null.
k. Inconsistency: We downgraded once. The point estimates are close with some overlap in the confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 65%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
l. Indirectness: We did not downgrade because the trials were conducted in different countries (high or upper-middle income).
m. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (MD = -1); the confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds of appreciable benefit (MD = -1) and harm (MD = +1).
n. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 10%).
o. Elserty 2016 included 2 arms (fixed pulse TENS + exercise vs. exercise; adjusted pulse TENS + exercise vs. exercise). Both were included in meta-analysis by splitting the comparison group number in half.
p. Inconsistency: We did not downgrade. The point estimates differ in direction but the confidence intervals overlap; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 4%).
q. Indirectness: We did not downgrade because the trials were conducted in different countries (low or lower-middle income).
r. Inconsistency: We did not downgrade. The point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 30% and 60%, which may represent moderate heterogeneity (i.e., I2 = 48%).
s. Depaoli Lemos 2021 used 4 TENS sessions; Petrofsky 2020 used a single TENS session.
t. Risk of bias: We downgraded once due to the potential for selection, performance and other biases.
u. Kofotolis and Jamison: Participants had 20-90 treatment sessions.
v. Inconsistency: We downgraded twice. The point estimates were in different directions with little to no overlap in confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 94%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
w. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached).
x. Inconsistency: We did not downgrade. Most of the point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 28%).
y. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate reached the pre-specified threshold for what may be considered appreciable benefit (SMD = -0.2). The confidence interval crossed the null.
z. Indirectness: We downgraded once. This is a single trial from (low or lower-middle income country).
aa. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate reached the pre-specified threshold for what may be considered clinically important (SMD = -0.2); the confidence interval crosses the null.
ab. Inconsistency: We did not downgrade. Most of the point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 39%).

ac. Imprecision: We downgraded once. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD = -0.2). The confidence interval crossed the null but the upper boundary did not cross the pre-specified threshold for what may be considered appreciable harm (SMD = +0.2).
ad. Imprecision: We downgraded twice. The sample size is small (OIS would have not been achieved). The point estimate did not reach the pre-specified threshold for what may be considered clinically important (SMD = -0.2); the confidence interval crosses the null with the lower and upper boundaries crossing the pre-specified thresholds for what may be considered appreciable benefit (SMD = -0.2) and harm (SMD = +0.2).
ae. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (SMD = -0.2). The confidence interval crossed the null.
af. Inconsistency: We downgraded once. There was some difference in magnitude and direction of the point estimates, but there was some overlap in confidence intervals. Statistical heterogeneity is between 50% and 90% (i.e., I2 = 69%). This could not be explained due to small subgroups and may represent substantial heterogeneity.
ag. We did not downgrade because the trials were conducted in different countries.
ah. Inconsistency: We did not downgrade. Most of the point estimates are similar with overlapping confidence intervals; statistical heterogeneity is between 0% and 40%, which might not be important (i.e., I2 = 39%).
ai. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate reached the pre-specified threshold for what may be considered appreciable benefit (SMD = -0.2). The confidence interval crossed the null but the upper boundary did not cross the pre-specified threshold for what may be considered appreciable harm (SMD = +0.2).
aj. Depaoli Lemos 2021 used 4 TENS sessions.
ak. Inconsistency: We downgraded twice. The point estimates were in different directions with little to no overlap in confidence intervals. Statistical heterogeneity is between 75% and 100% (i.e., I2 = 97%). This could not be explained due to small subgroups and may represent considerable heterogeneity.
al. Imprecision: We downgraded twice due to small sample size (the OIS would not have been reached). The point estimate did not reach the pre-specified threshold for what may be considered appreciable harm (-10). The confidence interval crossed the null with the boundaries crossing the pre-specified thresholds for what may be considered appreciable harm (-10) and benefit (+10).
am. Imprecision: We downgraded once due to small sample size (the OIS would not have been reached). The point estimate did not reach the pre-specified threshold for what may be considered appreciable harm (MD = -10). The confidence interval crossed the null.
an. Imprecision: We downgraded twice due to small sample size (the OIS would not have been reached). The point estimate did not reach the pre-specified threshold for what may be considered appreciable benefit (MD = -2.1). The confidence interval crossed the null with the boundaries crossing the pre-specified thresholds for what may be considered appreciable benefit (-2.1) or harm (+2.1).

References

1.Depaoli Lemos VJ, Selau RC,Blos C,Baptista Dohnert M,Boff Daitx R,de Almeida Brito V.. Electroacupuncture and Transcutaneous Electrical Nerve Stimulation in Chronic Nonspecific Low Back Pain: a Blind Randomized Clinical Trial.. Muscles, ligaments & tendons journal; 2021.
2.Elserty N, Kattabei O,Elhafez H.. Effect of Fixed Versus Adjusted Transcutaneous Electrical Nerve Stimulation Amplitude on Chronic Mechanical Low Back Pain. J Altern Complement Med; 2016.
3.Tella BA, Oghumu SN,Gbiri CAO. Efficacy of transcutaneous electrical nerve stimulation and interferential current on tactile acuity of individuals with nonspecific chronic low back pain. Neuromodulation; 2021.
4.Kibar S, Konak HE,Ay S,Doganay Erdogan B,Evcik D. The effectiveness of combined transcutaneous electrical nerve stimulation and interferential current therapy on chronic low back pain: a randomized, double-blind, sham-controlled trial. Fiziksel Tip ve Rehabilitasyon Bilimleri Dergisi [Journal of Physical Medicine and Rehabilitation Sciences]; 2020.
5.Shimoji K, Takahashi N,Nishio Y,Koyanagi M,Aida S. Pain relief by transcutaneous electric nerve stimulation with bidirectional modulated sine waves in patients with chronic back pain: a randomized, double-blind, sham-controlled trial. Neuromodulation; 2007.
6.Kofotolis ND, Vlachopoulos SP,Kellis E. Sequentially allocated clinical trial of rhythmic stabilization exercises and TENS in women with chronic low back pain. Clin Rehabil; 2008.
7.de Alencar Caldas VV, Maciel DG,Cerqueira MS,et al. Effect of Pain Education, Cryotherapy, and Transcutaneous Electrical Nerve Stimulation on the Pain, Functional Capacity, and Quality of Life in Patients With Nonspecific Chronic Low Back Pain: A Single-Blind Randomized Controlled Trial. Am J Phys Med Rehabil; 2021.
8.Petrofsky J, Laymon M,Lee H. The effect of transcutaneous electrical nerve stimulation and low-level continuous heat on non-specific low back pain: a randomized controlled trial. Gazzetta Medica Italiana; 2020.
9.Jamison RN, Wan L,Edwards RR,Mei A,Ross EL. Outcome of a High-Frequency Transcutaneous Electrical Nerve Stimulator (hfTENS) Device for Low Back Pain: A Randomized Controlled Trial. Pain Pract; 2019.
10.Alizadeh MH, Ahmadizad S. A Comparison of Exercise Therapy and Transcutaneous Electrical Nerve Stimulation for the Treatment of Chronic Low Back Pain. World Journal of Sport Sciences; 2009.

GRADE Table 3. ***What are the benefits and harms of TENS in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

No trials

B.8 Assistive products: lumbar braces, belts and/or supports and mobility assistive products

Overview of the PICO structure

Definition of the intervention	
The WHO defines assistive products as any external product (including devices, equipment, instruments or software), specially produced or generally available, the primary purpose of which is to maintain or improve an individual's functioning and independence, and thereby promote well-being. ***Non-rigid and rigid lumbar braces, belts and/or supports*** include plastic (rigid) or flexible (elastic or non-elastic) material with or without rigid inserts wrapping the lumbar/thoracolumbar trunk to block/limit mobility and/or reduce strains and physical demands on the lower back. These products are commonly used for CPLBP either as a treatment or to reduce recurrences of pain. They are accessible in most countries, with limitations due to costs (they are usually out of pocket expense) and climate (they are difficult to wear in high temperatures).	
PICO question	
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 or minimal intervention c) Usual care (described as usual care in the trial)

Web Annex D.B8: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • 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

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Moderate; trivial; uncertain: no evidence	Trivial; uncertain: no evidence
Harms	Moderate; trivial; uncertain: no evidence	Moderate; uncertain: no evidence

Web Annex D.B8: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Balance benefits to harms	Probably favours lumbar braces, belts and/or supports; probably does not favour lumbar braces, belts and/or supports; uncertain: no evidence	Probably favours lumbar braces, belts and/or supports; probably does not favour lumbar braces, belts and/or supports; uncertain: no evidence
Overall certainty	Very low: no evidence	Very low: no evidence
Values and preferences	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty; no important uncertainty or variability	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty; no important uncertainty or variability
Resource considerations	Moderate; moderate costs; negligible; varies	Moderate; moderate costs; negligible; varies
Equity and human rights	No impact; reduced; uncertain	No impact; reduced; uncertain
Acceptability	Yes, probably yes; probably no	Yes; probably yes; probably no
Feasibility	Yes; probably yes; uncertain	Yes; probably yes; uncertain

Web Annex D.B8: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of lumbar braces, belts and/or supports in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo/sham?***

No trials

GRADE Table 2. ***What are the benefits and harms of lumbar braces, belts and/or supports in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no or minimal intervention?***

No trials

GRADE Table 3. ***What are the benefits and harms of lumbar braces, belts and/or supports in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care or where the effect of the intervention could be isolated?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lumbar support plus usual care	usual care	Relative (95% CI)	Absolute (95% CI)		
Pain (follow-up: 4 weeks; assessed with: VAS and NRS) - better outcomes indicated by lower SMD												
2	randomized trials	serious[a]	not serious	not serious	serious[b]	none	98	51	-	SMD **1.19 lower** (2.38 lower to 0.01 lower)	⨁⨁◯◯ Low	
Disability (follow-up: 4 weeks; assessed with: RMDQ and ODI) - better outcomes indicated by lower SMD												

Web Annex D.B8: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lumbar support plus usual care	usual care	Relative (95% CI)	Absolute (95% CI)		
2	randomized trials	serious[a]	serious[c]	not serious	serious[b]	none	98	51	-	SMD **0.63 lower** (1.43 lower to 0.17 higher)	⨁◯◯◯ Very low	

Explanations

a. Risk of Bias: Downgraded one level for high risk of performance and detection bias in all RCTs
b. Imprecision: Downgraded one level for imprecision (less than 400 participants)
c. Inconsistency: Downgraded one level for inconsistency (I^2>75%)

Web Annex D.B8: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Narrative synthesis

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lumbar support plus usual care	usual care	Relative (95% CI)	Absolute (95% CI)		
Pain narrative												
2	randomized trials	serious[a]	not serious	not serious	very serious[b]	none	No significant differences in pain changes over the study period in all four studies				⨁◯◯◯ Very low	
Disability narrative												
3	randomized trials	serious[a]	not serious	not serious	very serious[b]	none	No significant differences in disability in two studies and significant changes (p<0.01) in one study over the study period				⨁◯◯◯ Very low	
Quality of life narrative												
1	randomized trials	serious[a]	not serious[c]	not serious	very serious[b]	none	Significant differences in quality of life changes (p<0.05)				⨁◯◯◯ Very low	

Explanations
a. Risk of bias: Downgraded one level for high of performance, detection and attrition biases for all RCTs.
b. Imprecision: Downgraded two levels for imprecision (less than 100 participants)
c. Inconsistency: It could not be judged due to a single trial.

C.1 Operant therapy

Overview of the PICO structure

Definition of the intervention	
Operant therapy aims to replace pain-related behaviours with helpful, healthy behaviours (e.g. exercise, work). Time-contingent exercises (i.e. quotas) and encouraging people with CPLBP to increase their activity levels are its main principles. This type of therapy is aligned with behavioural activation therapy.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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)

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Moderate; uncertain	Moderate; uncertain
Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Probably favours operant therapy; uncertain	Probably favours operant therapy; uncertain

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate; large; varies	Moderate; large; varies
Equity and human rights	Possibly reduced; no impact; uncertain; varies	Possibly reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; varies	Probably yes; probably no; varies
Feasibility	Varies	Varies

GRADE Table 1. ***What are the benefits and harms of operant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo***?

No trials.

GRADE Table 2. ***What are the benefits and harms of operant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Operant therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
4	randomized trials	very serious[a]	Not serious[b]	not serious	serious[c]	none	89	77	-	SMD **0.66 lower** (1.14 lower to 0.17 lower)	⨁◯◯◯ Very low	
Population subgroup 1: gender and/or sex												
Females 1	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	36	30	-	SMD **1.04 lower** (1.55 lower to 0.52 lower)	⨁◯◯◯ Very low	
Mixed 3	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	53	47	-	SMD **0.45 lower** (0.94 lower to 0.04 higher)	⨁◯◯◯ Very low	
Population subgroups 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term												
2	randomized trials	very serious[a]	not serious	not serious	very serious[d]	none	40	36	-	SMD **0.76 lower** (1.24 lower to 0.29 lower)	⨁◯◯◯ Very low	
Population subgroup 1: gender and/or sex												
Females 1	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	36	30	-	SMD **0.69 lower** (1.19 lower to 0.19 lower)	⨁◯◯◯ Very low	
Mixed 1	randomized trials	very serious[a]	not serious	not serious	very serious[g]	none	4	6	-	SMD **1.37 lower** (2.85 lower to 0.11 higher)	⨁◯◯◯ Very low	

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Operant therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - long term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	5	5	-	MD **0.66 lower** (1.7 lower to 0.38 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Back-specific functional status – short term												
3	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	55	47	-	MD **1.38 lower** (3.65 lower to 0.9 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	6	-	MD **5.36 lower** (17.11 lower to 6.39 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Back-specific functional status - long term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	5	-	MD **1.33 lower** (13.59 lower to 10.93 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies identified that reported on this outcome												

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Operant therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
3	randomized trials	very serious[a]	serious[f]	not serious	very serious[d]	none	62	56	-	SMD **0.29 lower** (1.27 lower to 0.69 higher)	⨁◯◯◯ Very low	
Population subgroup 1: gender and/or sex												
Females 1	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	36	30	-	SMD **1.13 lower** (1.65 lower to 0.60 lower)	⨁◯◯◯ Very low	
Mixed 2	randomized trials	very serious[a]	not serious	not serious	very serious[g]	none	4	6	-	SMD **0.2 higher** (0.35 lower to 0.74 higher)	⨁◯◯◯ Very low	
Population subgroups 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - intermediate term												
2	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	42	36	-	MD **3.05 lower** (5.41 lower to 0.7 lower)	⨁◯◯◯ Very low	
Population subgroup 1: gender and/or sex												
Females 1	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	36	30	-	MD **3.2 lower** (5.62 lower to 0.78 lower)	⨁◯◯◯ Very low	
Mixed 1	randomized trials	very serious[a]	not serious	not serious	very serious[g]	none	6	6	-	MD **0.5 lower** (10.57 lower to 9.57 higher)	⨁◯◯◯ Very low	

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Operant therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - long term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	5	-	MD **1.07 higher** (8.58 lower to 10.72 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (anxiety) - short term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	8	7	-	MD **3.81 higher** (8.08 lower to 15.7 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (anxiety) - intermediate term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	6	-	MD **3.17 higher** (9.5 lower to 15.84 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (anxiety) - long term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	5	-	MD **10.57 lower** (28.67 lower to 7.53 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (coping) - short term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	8	7	-	MD **1.59 higher** (33.19 lower to 36.37 higher)	⨁◯◯◯ Very low	

Web Annex D.C1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Operant therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (coping) - intermediate term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	6	-	MD **13 lower** (46.9 lower to 20.9 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning (coping) - long term												
1	randomized trials	very serious[a]	not serious[e]	not serious	very serious[g]	none	6	5	-	MD **4.5 lower** (32.34 lower to 23.34 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Social participation - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 2 levels due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Inconsistency not downgraded despite I^2 = 52%; heterogeneity may be explained by gender subgroups.

[c]Imprecision downgraded by 1 level: due to low number of participants

[d]Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.

[e]Inconsistency not assessed because only one study included in this analysis.

[f]Inconsistency downgraded by 1 level: unexplained considerable heterogeneity (I-sq = 83%)

[g]Imprecision downgraded by 2 levels: due to very low number of participants

GRADE Table 3. ***What are the benefits and harms of operant therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

No trials.

C.2 Respondent therapy

Overview of the PICO structure

Definition of the intervention	
Respondent therapy aims to modify the physiological response system to pain through the reduction of muscular tension through biofeedback, progressive relaxation and applied relaxation. This type of therapy is aligned with relaxation therapy.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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)

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; trivial; uncertain	Uncertain
Harms	Trivial; uncertain	Uncertain
Balance benefits to harms	Uncertain	Uncertain

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate; large; varies	Moderate; large; varies
Equity and human rights	Possibly reduced; no impact; uncertain; varies	Possibly reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; varies	Probably yes; probably no; varies
Feasibility	Varies	Varies

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of respondent therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
2	randomized trials	serious[a]	not serious	not serious	serious[b]	none	29	29	-	MD **6.21 lower** (14.94 lower to 2.52 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term or long term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term												
2	randomized trials	serious[a]	not serious	not serious	serious[b]	none	29	29	-	SMD **0.07 higher** (0.45 lower to 0.58 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	Placebo	Relative (95% CI)	Absolute (95% CI)		
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 1 level: due to unclear or high risk of bias in one study regarding random sequence generation, allocation concealment, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Imprecision downgraded by 1 level: low number of participants

GRADE Table 2.1. ***What are the benefits and harms of respondent therapy (biofeedback) in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
3	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	53	47	-	SMD **0.66 lower** (1.1 lower to 0.22 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	27	25	-	SMD **0.53 lower** (1.08 lower to 0.03 higher)	⨁◯◯◯ Very low	
High income 2	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	26	22	-	SMD **0.79 lower** (1.6 lower to 0.01 higher)	⨁◯◯◯ Very low	
Pain - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	43	37	-	SMD **0.62 lower** (1.07 lower to 0.17 lower)	⨁◯◯◯ Very low	

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	27	25	-	SMD **0.51 lower** (1.06 lower to 0.04 higher)	⨁◯◯◯ Very low	
High income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	16	12	-	SMD **0.85 lower** (1.64 lower to 0.06 lower)	⨁◯◯◯ Very low	
Back-specific functional status - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (anxiety) - short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	43	37	-	MD **5.15 lower** (8.74 lower to 1.57 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	27	25	-	MD **5.3 lower** (9.32 lower to 1.28 lower)	⨁◯◯◯ Very low	
High income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	16	12	-	MD **4.58 lower** (12.46 lower to 3.3 higher)	⨁◯◯◯ Very low	
Psychological functioning (depression) - short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	43	37	-	MD **3.78 lower** (8.06 lower to 0.5 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	27	25	-	MD **0.52 lower** (7.37 lower to 6.33 higher)	⨁◯◯◯ Very low	
High income 1	randomized trials	very serious[a]	not serious	not serious	serious[b]	None	16	12	-	MD **5.24 lower** (9.03 lower to 1.45 lower)	⨁◯◯◯ Very low	
Psychological functioning (coping) - short term												

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	16	12	-	MD **6.92 higher** (10.83 lower to 24.67 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Psychological functioning - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Imprecision downgraded by 1 level: low number of participants.

[c]Inconsistency not assessed because only one study included in this analysis.

GRADE Table 2.2. ***What are the benefits and harms of respondent therapy (relaxation) in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
2	randomized trials	very serious[a]	serious[b]	not serious	very serious[c]	none	31	27	-	MD **21.8 lower** (45.78 lower to 2.17 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term or long term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[d]	none	31	27	-	SMD **0.97 lower** (1.52 lower to 0.41 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies identified that reported on this outcome												

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
2	randomized trials	very serious[a]	serious[e]	not serious	very serious[c]	none	31	27	-	MD **6.8 lower** (19.73 lower to 6.12 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning - intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Inconsistency downgraded by 1 level: unexplained substantial heterogeneity I^2=57%

[c]Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.

[d]Imprecision downgraded by 1 level: low number of participants.

[e]Inconsistency downgraded by 1 level: unexplained considerable heterogeneity I^2=85%

GRADE Table 3. ***What are the benefits and harms of respondent therapy (relaxation) in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	very serious[c]	none	57	43	-	MD **11 lower** (22.22 lower to 0.22 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development (no subgroup analysis was performed; all studies performed in high income settings)												
Pain - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	54	45	-	MD **1.4 lower** (12.65 lower to 9.85 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Pain - long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	57	43	-	MD **3.3 lower** (11.6 lower to 5 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	Usual care	Relative (95% CI)	Absolute (95% CI)		
Back-specific functional status - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	54	45	-	MD **1.6 lower** (9.22 lower to 6.02 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Back-specific functional status - long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	57	43	-	MD **6.9 higher** (2.51 lower to 16.31 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Health-related quality of life - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	54	45	-	MD **2.6 lower** (11.9 lower to 6.7 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Health-related quality of life - long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.C2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Respondent therapy	Usual care	Relative (95% CI)	Absolute (95% CI)		
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	57	43	-	MD **1.5 lower** (5.87 lower to 2.87 higher)	⨁◯◯◯ Very low	
Psychological functioning (depression) - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[d]	none	54	45	-	MD **0.2 lower** (4.16 lower to 3.76 higher)	⨁◯◯◯ Very low	
Psychological functioning - long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Inconsistency not assessed because only one study included in this analysis.

[c]Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.

[d]Imprecision downgraded by 1 level: low number of participants.

C.3 Cognitive therapy

Overview of the PICO structure

Definition of the intervention	
Cognitive therapy aims to identify and modify cognition regarding pain and disability. It is proposed that beliefs about the meaning of pain and expectations regarding control over pain can be directly modified using cognitive restructuring techniques such as imagery and attention diversion.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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)

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Trivial; uncertain	Uncertain
Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Uncertain	Uncertain

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate; large; varies	Moderate; large; varies
Equity and human rights	Possibly reduced; no impact; uncertain; varies	Possibly reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; varies	Probably yes; probably no; varies
Feasibility	Varies	Varies

GRADE Table 1. ***What are the benefits and harms of cognitive therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo***?

No trials.

GRADE Table 2. ***What are the benefits and harms of cognitive therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
3	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	37	46	-	MD 2.74 **lower** (8.58 lower to 3.1 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	5	6	-	MD **0.02 higher** (0.98 lower to 1.02 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Pain - long term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	4	5	-	MD **0.08 higher** (0.93 lower to 1.09 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Back-specific functional status – short term												
4	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	133	93	-	SMD **0.1 lower** (0.37 lower to 0.17 higher)	⨁◯◯◯ Very low	

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Back-specific functional status - intermediate term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	7	6	-	MD **4.09 lower** (13.51 lower to 5.33 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Back-specific functional status - long term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	6	5	-	MD **4.41 lower** (14.11 lower to 5.29 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
General functional status – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events and serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (anxiety) - short term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	8	7	-	MD **4.56 higher** (7.66 lower to 16.78 higher)	⨁◯◯◯ Very low	

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (anxiety) - intermediate term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	7	6	-	MD **1.71 higher** (10.65 lower to 14.07 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (anxiety) - long term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	6	5	-	MD **6.23 lower** (27.59 lower to 15.13 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (depression) - short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[b]	none	24	25	-	MD **1.97 higher** (1.41 lower to 5.34 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - intermediate term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	7	6	-	MD **3.03 lower** (10.6 lower to 4.54 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (depression) - long term												

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	6	5	-	MD **4.77 lower** (12.33 lower to 2.79 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (coping) - short term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	8	7	-	MD **29.46 higher** (5.42 lower to 64.34 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (coping) - intermediate term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	7	6	-	MD **27.26 higher** (4.82 lower to 59.34 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Psychological functioning (coping) - long term												
1	randomized trials	very serious[a]	not serious[c]	not serious	serious[b]	none	6	5	-	MD **20.33 higher** (8.31 lower to 48.97 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed, only one included study on this outcome)												
Social participation – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy – short term, intermediate term or long term: no studies identified that reported on this outcome												

Web Annex D.C3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive therapy	No intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Imprecision downgraded by 1 level: low number of participants.

[c]Inconsistency not assessed because only one study included in this analysis.

GRADE Table 3. ***What are the benefits and harms of cognitive therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care***?

No trials.

C.4 Cognitive behavioural therapy (CBT)

Overview of the PICO structure

Definition of the intervention	
Cognitive behavioural therapy (CBT), is based on a multidimensional model of pain and focuses on reducing pain and distress by modifying physical sensation, catastrophic thinking and unhelpful behaviour(s). Treatment may include education about a multi-dimensional view of pain, identifying pain-eliciting and pain-aggravating situations, thoughts and behaviours, and using coping strategies and applied relaxation; in sum, integrating components of operant, respondent and cognitive therapies. Goal-setting and activity increases are encouraged as the basis of CBT to reduce feelings of helplessness and help the person gain control over their pain experience.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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)

Other Evidence-to-Decision (EtD) considerations

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

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Small; trivial; uncertain	Small; trivial; uncertain
Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Probably favours CBT; uncertain	Probably favours CBT; uncertain

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate; large; varies	Moderate; large; varies
Equity and human rights	Possibly reduced; no impact; uncertain; varies	Possibly reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; varies	Probably yes; probably no; varies
Feasibility	Varies	Varies

GRADE Table 1. ***What are the benefits and harms of cognitive behavioural therapy (CBT) in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo***?

No trials.

GRADE Table 2. ***What are the benefits and harms of cognitive behavioural therapy (CBT) in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

Certainty assessment							№ of patients		Effect		Certainty	
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
22	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	1265	1075	-	SMD **0.49 lower** (0.75 lower to 0.24 lower)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 3	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	98	99	-	SMD **0.71 lower** (1.85 lower to 0.43 higher)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 19	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	1167	976	-	SMD **0.47 lower** (0.73 lower to 0.2 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
Low/ middle income 2	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	46	45	-	MD **1.42 lower** (3.74 lower to 0.9 higher)	⨁◯◯◯ Very low	

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
High income 20	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	1219	1030	-	SMD **0.44 lower** (0.7 lower to 0.19 lower)	⨁◯◯◯ Very low	
Pain - intermediate term												
5	randomized trials	very serious[a]	serious[c]	not serious	not serious	none	570	368	-	SMD **0.08 lower** (0.32 lower to 0.16 higher)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	51	52	-	MD **0.00 lower** (0.85 lower to 0.85 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 4	randomized trials	very serious[a]	serious[c]	not serious	not serious	none	519	316	-	SMD **0.08 lower** (0.39 lower to 0.22 higher)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development (no subgroup analysis was performed)												
Pain - long term												
7	randomized trials	very serious[a]	serious[d]	not serious	serious[e]	none	799	593	-	SMD **1.06 lower** (1.66 lower to 0.47 lower)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	serious[e]	none	49	49	-	MD **1.00 lower** (1.83 lower to 0.17 lower)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 6	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	750	544	-	SMD **1.18 lower** (1.86 lower to 0.49 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development (no subgroup analysis was performed)												
Back-specific functional status – short term												
21	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	1219	1025	-	SMD **0.46 lower** (0.75 lower to 0.18 lower)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 3	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	98	99	-	SMD **0.76 lower** (1.86 lower to 0.35 higher)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 18	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	1121	926	-	SMD **0.42 lower** (0.72 lower to 0.11 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Low/ middle income 2	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	46	45	-	SMD **1.12 lower** (2.76 lower to 0.52 higher)	⨁◯◯◯ Very low	
High income 19	randomized trials	very serious[a]	serious[b]	not serious	serious[e]	none	1173	980	-	SMD **0.4 lower** (0.68 lower to 0.11 lower)	⨁◯◯◯ Very low	
Back-specific functional status - intermediate term												
5	randomized trials	very serious[a]	not serious	not serious	not serious	none	538	361	-	SMD **0.15 lower** (0.3 lower to 0)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	54	54	-	MD **0.1 lower** (1.53 lower to 1.73 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 4	randomized trials	very serious[a]	not serious	not serious	not serious	none	484	307	-	SMD **0.18 lower** (0.35 lower to 0.02 lower)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - long term												
7	randomized trials	very serious[a]	serious[d]	not serious	serious[e]	none	745	557	-	SMD **1.16 lower** (2.01 lower to 0.32 lower)	⨁◯◯◯ Very low	

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	serious[e]	none	49	49	-	MD **1.1 lower** (2.86 lower to 0.66 higher)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 6	randomized trials	very serious[a]	serious[d]	not serious	serious[e]	none	696	508	-	SMD **1.33 lower** (2.31 lower to 0.34 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development (no subgroup analysis was performed)												
General functional status – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term												
6	randomized trials	very serious[a]	serious[d]	not serious	serious[e]	none	504	519	-	SMD **0.61 higher** (0.11 higher to 1.1 higher)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Health-related quality of life - intermediate term												
2	randomized trials	very serious[a]	not serious	not serious	serious[e]	none	207	233	-	SMD **0.25 higher** (0.07 higher to 0.44 higher)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life - long term												
4	randomized trials	very serious[a]	serious[d]	not serious	serious[e]	none	311	301	-	SMD **1.06 higher** (0.03 higher to 2.1 higher)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Adverse events – narrative results only (see text)												
-	-	-	-	-	-	-	-	-	-	-	-	
Serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
8	randomized trials	very serious[a]	not serious	not serious	not serious	none	335	312	-	SMD **0.14 lower** (0.3 lower to 0.01 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	52	54	-	MD **0 lower** (1.73 lower to 1.73 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 7	randomized trials	very serious[a]	not serious	not serious	not serious	none	283	258	-	SMD **0.18 lower** (0.36 lower to 0)	⨁⨁◯◯ Low	

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroup 4: regional economic development (no subgroup analysis was performed)												
Psychological functioning (depression) - intermediate term												
3	randomized trials	very serious[a]	not serious	not serious	not serious	none	165	162	-	SMD **0.06 lower** (0.38 lower to 0.26 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	54	54	-	MD **0.7 higher** (0.59 lower to 1.99 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 2	randomized trials	very serious[a]	not serious	not serious	not serious	none	111	108	-	SMD **0.2 lower** (0.47 lower to 0.07 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - long term												
2	randomized trials	very serious[a]	not serious	not serious	not serious	none	151	149	-	SMD **0.1 lower** (0.33 lower to 0.13 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	49	49	-	MD **0.3 lower** (1.69 lower to 1.09 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	102	100	-	MD **0.46 lower** (1.63 lower to 0.71 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - short term												
4	randomized trials	very serious[a]	not serious	not serious	not serious	none	196	194	-	SMD **0.08 lower** (0.28 lower to 0.11 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	52	54	-	MD **0.6 lower** (1.75 lower to 0.55 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 3	randomized trials	very serious[a]	not serious	not serious	not serious	none	144	140	-	SMD **0.04 lower** (0.28 lower to 0.19 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Psychological functioning (anxiety) - intermediate term												
2	randomized trials	very serious[a]	not serious	not serious	not serious	none	153	152	-	SMD **0.14 lower** (0.37 lower to 0.08 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	51	52	-	MD **0.6 lower** (1.6 lower to 0.4 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	102	100	-	MD **0.56 lower** (2.1 lower to 0.98 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - long term												
2	randomized trials	very serious[a]	not serious	not serious	not serious	none	151	149	-	SMD **0.2 lower** (0.43 lower to 0.03 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	49	49	-	MD **0.6 lower** (1.76 lower to 0.56 higher)	⨁⨁◯◯ Low	

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Not specified whether radicular leg pain included 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	102	100	-	MD **0.98 lower** (2.35 lower to 0.39 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (coping) - short term												
4	randomized trials	very serious[a]	not serious	not serious	serious[e]	none	126	112	-	SMD **0.49 higher** (0.23 higher to 0.75 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Social participation - short term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - intermediate term												
2	randomized trials	very serious[a]	serious[b]	not serious	very serious[g]	none	44/64 (68.8%)	35/62 (56.5%)	**RR 1.08** (0.51 to 2.30)	**45 more per 1.000** (from 277 fewer to 734 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Social participation - long term												
2	randomized trials	very serious[a]	serious[c]	not serious	very serious[g]	none	73/137 (53.3%)	76/135 (56.3%)	**RR 1.02** (0.66 to 1.57)	**11 more per 1.000** (from 191 fewer to 321 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Self-efficacy - short term												
4	randomized trials	very serious[a]	not serious	not serious	not serious	none	148	139	-	SMD **0.04 higher** (0.19 lower to 0.28 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	52	54	-	MD **0.9 higher** (4.02 lower to 5.82 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 6	randomized trials	very serious[a]	not serious	not serious	not serious	none	96	85	-	SMD **0.03 higher** (0.26 lower to 0.32 higher)	⨁⨁◯◯ Low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Self-efficacy - intermediate term												
1	randomized trials	very serious[a]	not serious[f]	not serious	not serious	none	51	52	-	MD **0.2 higher** (4.28 lower to 4.68 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Self-efficacy - long term												
1	randomized trials	very serious[a]	not serious	not serious	not serious	none	49	49	-	MD **2.6 higher** (1.71 lower to 6.91 higher)	⨁⨁◯◯ Low	

Web Annex D.C4: 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	Combined behavioural	No intervention	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

CI: confidence interval; MD: mean difference; SMD: standardized mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, similarity of groups at baseline, co-interventions, and compliance with the intervention.

[b]Inconsistency downgraded by 1 level: unexplained considerable heterogeneity $I^2 > 80\%$

[c]Inconsistency downgraded by 1 level: unexplained substantial heterogeneity $I^2 = 50\% - 75\%$

[d]Inconsistency downgraded by 1 level: unexplained considerable heterogeneity $I^2 > 90\%$

[e]Imprecision downgraded by 1 level: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect.

[f]Inconsistency not assessed, only one study reported on this outcome.

[g]Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for harm.

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 3. ***What are the benefits and harms of combined behavioural therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
4	randomized trials	very serious[a]	not serious	not serious	not serious	none	484	485	-	MD **0.24 lower** (0.35 lower to 0.12 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term												
5	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	552	553	-	MD **0.13 lower** (0.35 lower to 0.09 higher)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[e]	not serious	not serious	none	68	68	-	MD **0.5 higher** (0.14 lower to 1.14 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 4	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	484	485	-	MD **0.18 higher** (0.38 lower to 0.03 higher)	⨁◯◯◯ Very low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Pain - long term												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
4	randomized trials	very serious[a]	not serious	not serious	not serious	none	448	448	-	MD **0.24 lower** (0.48 lower to 0.01 higher)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[e]	not serious	not serious	none	68	68	-	MD **0.1 higher** (0.66 lower to 0.86 higher)	⨁⨁◯◯ Low	
Not specified whether radicular leg pain included 3	randomized trials	very serious[a]	serious[b]	not serious	not serious	none	380	380	-	MD **0.29 lower** (0.58 lower to 0.0)	⨁◯◯◯ Very low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status – short term												
2	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	231	234	-	MD **1.46 lower** (2.34 lower to 0.58 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term												
3	randomized trials	very serious[a]	not serious	not serious	not serious	none	299	302	-	MD **1.01 lower** (1.87 lower to 0.14 lower)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[e]	not serious	serious[c]	none	68	68	-	MD **0.2 lower** (2.05 lower to 1.65 higher)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 2	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	231	234	-	MD **1.24 lower** (2.22 lower to 0.26 lower)	⨁◯◯◯ Very low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												
No subgroup analysis was performed; all studies performed in high income settings.												
Back-specific functional status - long term												
3	randomized trials	very serious[a]	not serious	not serious	not serious	none	299	302	-	MD **0.94 lower** (1.85 lower to 0.03 lower)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded radicular leg pain 1	randomized trials	very serious[a]	not serious[e]	not serious	serious[c]	none	68	68	-	MD **0.2 higher** (1.82 lower to 2.22 higher)	⨁◯◯◯ Very low	
Not specified whether radicular leg pain included 2	randomized trials	very serious[a]	not serious	not serious	serious[c]	none	231	234	-	MD **1.23 lower** (2.25 lower to 0.21 lower)	⨁◯◯◯ Very low	
Population subgroup 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
General functional status - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	none	253	251	-	MD **2.25 lower** (3.85 lower to 0.66 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Health-related quality of life - intermediate term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	none	253	251	-	MD **1.89 lower** (3.5 lower to 0.28 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Health-related quality of life - long term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	None	261	259	-	MD **0.86 lower** (2.59 lower to 0.87 higher) MD **3.43 lower** (5.28 lower to 1.58 lower)	⨁⨁◯◯ Low	Not pooled
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Adverse events – narrative results only (see text)												
-	-	-	-	-	-	-	-	-	-	-	-	
Serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
2	randomized trials	very serious[d]	not serious	not serious	not serious	None	216	218	-	MD **1.47 lower** (3.33 lower to 0.39 higher) MD **2.17 lower** (2.88 lower to 1.46 lower)	⨁⨁◯◯ Low	Not pooled
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - intermediate term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	None	216	218	-	MD **0.98 lower** (2.82 lower to 0.86 higher) MD **1.16 lower** (1.95 lower to 0.37 lower)	⨁⨁◯◯ Low	Not pooled
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - long term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	None	261	159	-	MD **0.84 lower** (1.66 lower to 0.02 lower) MD **1.61 lower** (2.68 lower to 0.54 lower)	⨁⨁◯◯ Low	Not pooled
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - short term												
1	randomized trials	very serious[d]	not serious[e]	not serious	not serious	none	112	113	-	MD **0.42 lower** (0.71 lower to 0.13 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
Psychological functioning (anxiety) - intermediate term												
1	randomized trials	very serious[d]	not serious[e]	not serious	not serious	none	112	113	-	MD **0.51 lower** (0.86 lower to 0.16 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - long term												
1	randomized trials	very serious[d]	not serious[e]	not serious	not serious	none	112	113	-	MD **0.25 lower** (0.58 lower to 0.08 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Social participation - short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	none	253	251	-	MD **2 higher** (0.01 lower to 4.01 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Self-efficacy - intermediate term												
2	randomized trials	very serious[d]	not serious	not serious	not serious	none	253	251	-	MD **1.65 higher** (0.61 lower to 3.9 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Self-efficacy - long term												

Web Annex D.C4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Combined behavioural	Usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[d]	not serious[e]	not serious	serious[c]	none	149	146	-	MD **4.23 higher** (1.84 higher to 6.62 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

CI: confidence interval; MD: mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, selective reporting, co-interventions, and compliance with the intervention.

[b]Inconsistency downgraded by 1 level: unexplained substantial heterogeneity I^2=59%

[c]Imprecision downgraded by 1 level: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect.

[d]Risk of bias downgraded by 2 levels: due to high risk of bias across all studies regarding blinding of participants, blinding of care providers, blinding of outcome assessment, and compliance with the intervention.

[e]Inconsistency not assessed, only one study reported on this outcome.

C.5 Mindfulness-based stress reduction (MBSR) therapy

Overview of the PICO structure

Definition of the intervention	
Mindfulness-based stress reduction (MBSR) therapy aims to reduce stress by developing mindfulness: a non-judgemental, moment-by-moment acceptance of awareness. The intervention is free of any cultural, religious and ideological factors, but it is associated with the Buddhist origins of mindfulness.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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)

Other Evidence-to-Decision (EtD) considerations

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

No evidence synthesis commissioned for all adults. Judgements made based on experience of GDG members	From the qualitative studies it appears that mindfulness and meditation therapies are an accepted treatment to adults aged 60 and over, although the certainty of the evidence was low or very low. **# Review findings GRADE-CERQual Assessment of confidence** 18 Mindfulness and meditation allowed some participants to increase their body awareness in relation to, for example, breathing, posture, cognition and pain. In some cases, this allowed for early recognition of pain. VERY LOW 19 Mindfulness and meditation encouraged participants to examine, assess, understand and accept their pain rather than avoid it. In some cases, this decreased the significance or power of the pain in the participants' lives, allowing some participants to take control and push pain into the background. In turn, participants were more aware of their bodies, increasing their ability to relax and handle stress in relation to their pain and in other day to day situations such as better sleep, attention, wellbeing, and general quality of life. LOW 20 Some participants were able to use mindfulness and meditation for pain management and coping to varying degrees. Some participants experienced no relief, while others had some or short-term relief and a few were able to eliminate feelings of pain. LOW

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Domain	**All adults**	**Older people**
Benefits	Uncertain	Uncertain
Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Uncertain	Uncertain

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Overall certainty	Low; very low	Low; very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Moderate; large; varies	Moderate; large; varies
Equity and human rights	Possibly reduced; no impact; uncertain; varies	Possibly reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; varies	Probably yes; probably no; varies
Feasibility	Varies	Varies

GRADE Table 1. ***What are the benefits and harms of mindfulness-based stress reduction therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo?***

No trials.

GRADE Table 2. ***What are the benefits and harms of mindfulness-based stress reduction therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

No trials.

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 3. ***What are the benefits and harms of mindfulness-based stress reduction therapy in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness-based stress reduction	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.63 lower** (1 lower to 0.26 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.45 lower** (0.89 lower to 0.01 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - long term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.63 lower** (1.06 lower to 0.2 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status – short term												

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness-based stress reduction	Usual care	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[c]	none	116	113	-	MD **1.57 lower** (2.67 lower to 0.47 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[c]	none	116	113	-	MD **1.37 lower** (2.52 lower to 0.22 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - long term												
1	randomized trials	very serious[a]	not serious[b]	not serious	serious[c]	none	116	113	-	MD **1.87 lower** (3.11 lower to 0.63 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
General functional status – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **1.48 higher** (0.04 lower to 3 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness-based stress reduction	Usual care	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.31 higher** (1.52 lower to 2.14 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Health-related quality of life - long term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.94 higher** (0.85 lower to 2.73 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Adverse events or serious adverse events: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **1.48 lower** (2.3 lower to 0.66 lower)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.68 lower** (1.43 lower to 0.07 higher)	⨁⨁◯◯ Low	

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness-based stress reduction	Usual care	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (depression) - long term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.63 lower** (1.47 lower to 0.21 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - short term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.24 lower** (0.56 lower to 0.08 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - intermediate term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.02 lower** (0.4 lower to 0.36 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning (anxiety) - long term												
1	randomized trials	very serious[a]	not serious[b]	not serious	not serious	none	116	113	-	MD **0.01 lower** (0.37 lower to 0.35 higher)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												

Web Annex D.C5: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness-based stress reduction	Usual care	Relative (95% CI)	Absolute (95% CI)		
Social participation – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy – short term, intermediate term or long term: no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference

Explanations

[a]Risk of bias downgraded by 2 levels: due to unclear or high risk of bias across all studies regarding blinding of participants, blinding of care providers, blinding of outcome assessment, co-interventions, and compliance with the intervention.

[b]Inconsistency not assessed, only one study reported on this outcome.

[c]Imprecision downgraded by 1 level: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect.

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.

D.2 Cannabis-related pharmaceutical preparations for therapeutic use

Overview of the PICO structure

Definition of the intervention	
Cannabis-related pharmaceutical preparations for therapeutic use (or 'cannabinoids') refer to a group of closely related compounds that are active in cannabis, with the two main cannabinoid compounds being tetrahydrocannabinol (THC) and cannabidiol (CBD), which are suggested to have analgesic and anti-inflammatory properties.*(1)* Cannabinoids were evaluated with short-term (< 4 weeks) and long-term (≥ 4 weeks) treatment duration, taken by various modalities including smoking or ingestion.	
PICO question	
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.D2: 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

ETD process not completed since no trials were available.

Summary of judgements

ETD process not completed since no trials were available.

References

1. McDonagh MS, Morasco BJ, Wagner J, Ahmed AY, Fu R, Kansagara D et al. Cannabis-Based Products for Chronic Pain. A Systematic Review. Annals of Internal Medicine. 2022;175:1143-53. doi: 10.7326/M21-4520.

D.3 Injectable local anaesthetics

Overview of the PICO structure

Definition of the intervention	
Injectable local anaesthetics include the subcutaneous, myofascial or intramuscular delivery of anaesthetic agents (lidocaine, articaine, bupivacaine, chloroprocaine, mepivacaine, procaine, ropivacaine and tetracaine) into local soft and/or connective tissues in the region of the lower back, between the 12th rib and gluteal fold. The injectate is delivered only to the extraspinal soft tissue and not delivered to intra-spinous structures, as is the case with intradiscal, epidural, intrathecal, facet joint and nerve root injections.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care

Web Annex D.D3: 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

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.D3: 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.D3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of judgements

Domain	All adults	Older people
Benefits	Trivial; uncertain	Trivial; uncertain
Harms	Trivial; uncertain	Trivial; uncertain
Balance benefits to harms	Probably does not favour local anaesthetic injections; uncertain	Probably does not favour local anaesthetic injections; uncertain
Overall certainty	Very low	Very low
Values and preferences	Important uncertainty or variability; possibly important uncertainty or variability	Important uncertainty or variability; possibly important uncertainty or variability
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	Probably reduced; reduced; no impact; uncertain; varies	Probably reduced; reduced; no impact; uncertain; varies
Acceptability	Probably yes; probably no; uncertain; varies	Probably yes; probably no; uncertain; varies
Feasibility	Yes; probably yes	Yes; probably yes

GRADE Table 1. ***What are the benefits and harms of local anaesthetic injections in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo/sham injections?***

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Local anaesthetic	Placebo/sham	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (assessed with: VAS; Scale from: 0 to 100)[a]												
2[b,c]	randomized trials	serious[d]	serious[e]	not serious	serious[f]	none	138	137	-	MD **10 lower** (25.44 lower to 5.43 higher)	⨁◯◯◯ Very low	Analysis 1.1
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
High income 1[g]	randomized trials	serious[h]	not serious[i]	serious[j]	very serious[k]	none	12	12	-	MD **22.4 lower** (45.51 lower to 0.71 higher)	⨁◯◯◯ Very low	
Low/middle income 1[l]	randomized trials	serious[m]	not serious[i]	serious[n]	serious[o]	none	126	125	-	MD **5 lower** (11.32 lower to 1.32 higher)	⨁◯◯◯ Very low	
Pain - short term (assessed with: decrease of at least 30% in VAS score)												
1	randomized trials	serious[m]	not serious[i]	not serious	Very serious[k]	none	71/126	62/125	RR 1.14 (0.90 to 1.44)	69 more per 1000 (50 fewer to 218 more)	⨁◯◯◯ Very low	

Web Annex D.D3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Local anaesthetic	Placebo/sham	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study reported on this outcome)												
Pain - short term (assessed with: "feeling improved" pain severity compared with baseline)												
1	randomized trials	serious[h]	not serious[i]	not serious	very serious[k]	none	7/12	1/12	RR 7.00 (1.01 to 48.53)	500 more per 1000 (1 more to 1000 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study reported on this outcome)												
Pain - intermediate or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health related quality of life – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events assessed: any unfavourable symptom, regardless of its relationship to treatment, during the treatment period												
1[l]	randomized trials	serious[m]	not serious[i]	not serious	very serious[p]	none	7/126 (5.6%)	2/125 (1.6%)	**RR 3.47** (0.74 to 16.39)	**40 more per 1,000** (from 4 fewer to 246 more)	⨁◯◯◯ Very low	Analysis 1.4
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Serious adverse events												

Web Annex D.D3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Comments
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Local anaesthetic	Placebo/sham	Relative (95% CI)	Absolute (95% CI)		
1[g]	randomized trials	serious[h]	not serious[i]	not serious	very serious[q]	none	0/12 (0.0%)	0/12 (0.0%)	not estimable		⨁◯◯◯	Analysis 1.5
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study reported on this outcome)												
Psychological functioning (depression) – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; RR: risk ratio

Explanations

a. FU time between 2–12 weeks
b. Collee 1991, Imamura 2016
d. Risk of bias downgraded by 1 level due to unclear or high risk of bias regarding random sequence generation, allocation concealment, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, compliance, and other bias.
e. Inconsistency downgraded by 1 level: substantial heterogeneity I^2=51%. Inconsistency is not clearly explained by the subgroup analyses of HIC versus LMIC setting.
f. Imprecision downgraded by 1 level: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants. This outcome was not downgraded an additional level for imprecision because it was downgraded for inconsistency, which is related to and would have contributed to the severity of the imprecision.
g. Collee 1991
h. Risk of bias downgraded by one level due to unclear or high risk of bias regarding random sequence generation, allocation concealment, incomplete outcome data, selective outcome reporting, compliance, and other bias.
i. Inconsistency not assessed as only one study included in this analysis.
j. Indirectness downgraded by 1 level: only one study included in this subgroup analysis, it is unclear whether it is representative of all high-income country settings.
k. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.
l. Imamura 2016
m. Risk of bias downgraded by one level due to unclear or high risk of bias regarding blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, and compliance.
n. Indirectness downgraded by 1 level: only one study included in this subgroup analysis, it is unclear whether it is representative of all low/middle-income country settings.
o. Imprecision downgraded by 1 level: despite narrow confidence intervals around the effect estimate showing little to no difference, downgraded due to low number of participants.
p. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for harm and the possibility for no effect and low number of participants.
q. Imprecision downgraded by 2 levels: no events in either group and a very low number of participants.

GRADE Table 2. ***What are the benefits and harms of local anaesthetic injections in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Local anaesthetic	no intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (assessed with: VAS; Scale from: 0 to 100)[a]												
1[b,c]	randomized trials	serious[d]	not serious[e]	not serious	very serious[f]	none	126	127	-	MD **5 lower** (11.65 lower to 1.65 higher)	⨁◯◯◯ Very low	Analysis 2.1
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study reported on this outcome)												
Pain - short term (assessed with: decrease of at least 30% in VAS score)												
1[b]	randomized trials	serious[d]	not serious[e]	not serious	very serious[f]	none	71/126	51/127	RR 1.40 (1.08 to 1.82)	161 more per 1000 (32 more to 329 more)	⨁◯◯◯ Very low	
Pain - intermediate or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health related quality of life – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events												

Web Annex D.D3: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Local anaesthetic	no intervention	Relative (95% CI)	Absolute (95% CI)		
1[b]	randomized trials	serious[d]	not serious[e]	not serious	very serious[g]	none	7/126 (5.6%)	4/127 (3.1%)	**RR 1.76** (0.53 to 5.88)	**24 more per 1,000** (from 15 fewer to 154 more)	⨁◯◯◯ Very low	Analysis 2.3
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study reported on this outcome)												
Serious adverse events - not reported												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-					-	

CI: confidence interval; MD: mean difference; RR: risk ratio

Explanations

a. FU time 12 weeks
b. Imamura 2016
c. The study measured the outcome on an additional scale as dichotomous outcome as decrease of at least 30% in VAS score compared with baseline at 12 weeks (Analysis 2.2): there were 71/126 events in the intervention group vs 51/127 events in the comparison group (no intervention): RR 1.40 95% CI (1.08 to 1.82)
d. Risk of bias downgraded by one level due to unclear or high risk of bias regarding, blinding of participants, blinding of care providers, blinding of outcome assessment, incomplete outcome data, selective reporting, and compliance.
e. Inconsistency not assessed as only one study included in this analysis.
f. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for no effect and low number of participants.
g. Imprecision downgraded by 2 levels: due to wide confidence interval consistent with the possibility for benefit and the possibility for harm and low number of participants.

GRADE Table 3. ***What are the benefits and harms of local anaesthetic injections in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care?***

No trials

D.4 Herbal medicines

Overview of the PICO structure

Definition of the intervention	
WHO defines herbal medicines as herbs, herbal materials, herbal preparations and finished herbal products that contain, as active ingredients, parts of plants, or other plant materials, or combinations of both. For the purpose of this guideline, herbal medicines were restricted to plants or parts of plants used for medicinal purposes, administered orally (ingestion) or applied topically. This definition does not include plant substances, smoked individual chemicals derived from plants, or synthetic chemicals based on plant constituents.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care

Web Annex D.D4: 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 herbal medicines

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** 7 Some participants adopted alternative forms of treatment (traditional or herbal medicines) as a part of their self-management approach when conventional treatments failed. Some viewed this as experimenting to find a solution. Often participants did not inform their health care provider about taking this type of treatment. LOW

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

D.4.1 Topical Cayenne pepper [Capsicum frutescens]

Domain	**All adults**	**Older people**
Benefits	Moderate; small; uncertain	Moderate; small; uncertain
Harms	Moderate; small; uncertain	Moderate; small; uncertain
Balance benefits to harms	Probably favours cayenne pepper; probably does not favour cayenne pepper; neutral; uncertain	Probably favours cayenne pepper; probably does not favour cayenne pepper; neutral; uncertain
Overall certainty	Low	Low
Values and preferences	Possibly important uncertainty or variability; probably no important uncertainty or variability	Possibly important uncertainty or variability; probably no important uncertainty or variability
Resource considerations	Moderate costs; varies	Moderate costs; varies
Equity and human rights	No impact; uncertain; varies	No impact; uncertain; varies
Acceptability	Yes; varies	Yes; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

D.4.2 Devil's claw [Harpagophytum procumbens]

Benefits	Small; trivial; uncertain	Small; trivial; uncertain

Harms	Uncertain	Uncertain
Balance benefits to harms	Probably does not favour Devil's claw; uncertain	Probably does not favour Devil's claw; uncertain
Overall certainty	Low; very low	Low; very low
Values and preferences	Possibly important uncertainty or variability; probably no important uncertainty or variability	Possibly important uncertainty or variability; probably no important uncertainty or variability
Resource considerations	Moderate; varies	Moderate; varies
Equity and human rights	No impact; uncertain; varies	No impact; uncertain; varies
Acceptability	Yes; varies	Yes; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

D.4.3 White willow [Salix spp.]

Benefits	Uncertain	Uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Uncertain	Uncertain
Overall certainty	Low; very low	Low; very low
Values and preferences	Possibly important uncertainty or variability; probably no important uncertainty or variability	Possibly important uncertainty or variability; probably no important uncertainty or variability
Resource considerations	Moderate; varies	Moderate; varies
Equity and human rights	No impact; uncertain; varies	No impact; uncertain; varies
Acceptability	Yes; varies	Yes; varies
Feasibility	Yes; probably yes; varies	Yes; probably yes; varies

D.4.4 Brazilian arnica [Solidago chilensis]

ETD process not completed based on GDG decision that too few trials contributed to the evidence base.

D.4.5 Ginger [Zingiber officinale Roscoe]

ETD process not completed based on GDG decision that too few trials contributed to the evidence base.

D.4.6 White lily [Lilium candidum]

ETD process not completed based on GDG decision that too few trials contributed to the evidence base.

D.4.7 Combination herbal compress [Zingiber cassumunar Roxb. rhizomes, Curcuma longa L. rhizomes, Cymbopogon citratus (DC.), Stapf leaves and leaf sheaths, Croton roxburghii N.P.Balakr. leaves, Tamarindus indica L. leaves, Citrus hystrix DC. peels, Blumea balsamifera (L.) DC. leaves, Vitex trifolia L. leaves and camphor]

ETD process not completed based on GDG decision that too few trials contributed to the evidence base.

D.4.8 Combination transdermal diffusional patch [Oleum thymi, Oleum limonis, Oleum nigra, Oleum rosmarini, Oleum chamomilla and Oleum lauri expressum]

ETD process not completed based on GDG decision that too few trials contributed to the evidence base.

Web Annex D.D4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of Cayenne pepper [Capsicum frutescens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to placebo?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Capsicum frutescens	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (reduction of >30% pain score) - short term												
3	randomized trials	serious[a]	not serious	not serious	Not serious	none	203/304 (66.8%)	146/307 (47.6%)	RR 1.40 (1.22 to 1.62)	**190 more per 1000** (from 105 more to 295 more)	⨁⨁⨁◯ Moderate	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain (reduction of >50% pain score) - short term												
3	randomized trials	serious[a]	not serious	not serious	Not serious	none	140/304 (46.1%)	76/307 (24.8%)	RR 1.85 (1.47 to 2.31)	**210 more per 1000** (from 116 more to 324 more)	⨁⨁⨁◯ Moderate	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term or long term – no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term, intermediate term or long term – no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events												

Web Annex D.D4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Capsicum frutescens	Placebo	Relative (95% CI)	Absolute (95% CI)		
3	randomized trials	serious[a]	not serious	not serious	serious[b]	none	36/304 (11.8%)	17/307 (5.5%)	RR 2.04 (1.19 to 3.51)	**58 more per 1000** (from 11 more to 139 more)	⨁⨁◯◯ Low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Change in medication - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; RR: risk ratio

Explanations
a. Risk of bias downgraded by 1 level due to unclear or high risk of selection bias, attrition bias, reporting bias, similar groups at baseline, and compliance.
b. Imprecision downgraded by 1 level due to few events.

GRADE Table 2. ***What are the benefits and harms of Cayenne pepper [Capsicum frutescens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to no intervention***?

No trials

GRADE Table 3. ***What are the benefits and harms of Cayenne pepper [Capsicum frutescens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to usual care***?

No trials

GRADE Table 4. ***What are the benefits and harms of Devil's claw [Harpagophytum procumbens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	H.procumbens	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (reduction of at least 30% pain intensity)												
2	randomized trials	serious[a]	not serious	not serious	serious[b]	none	25/185 (13.5%)	4/121 (3.3%)	RR 3.73 (1.29 to 10.81)	**90 more per 1000** (from 10 more to 324 more)	⨁⨁◯◯ Low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term												
2	randomized trials	serious[a]	not serious[c]	not serious	very serious[d]	none	In Chrubasik 1996 (n=118) the relative median change in the intervention group was 20% (IQR 0; 35) and in the placebo group 8% (IQR -2; 23) (p=0.059). In Chrubasik 1999 (n=197) the relative median change in the low dose group was 21% (IQR 2; 34), the high dose group 18% (IQR 0; 40) and in the placebo group 21% (IQR 6; 34) (p=0.68).				⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.D4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	H.procumbens	Placebo	Relative (95% CI)	Absolute (95% CI)		
Adverse events												
2	randomized trials	serious[a]	serious[f]	not serious	very serious[g]	none	12/185 (6.5%)	11/121 (9.1%)	RR 1.08 (0.12 to 9.94)	**7 more per 1000** (from 80 fewer to 813 more)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Change in medication - short term												
2	randomized trials	serious[a]	not serious[c]	serious[h]	very serious[e]	none	Chrubasik 1996 (n=118) reported that the intervention group consumed a mean (± SD) of 95 ± 157mg in the last three weeks of treatment while the placebo group consumed 102 ± 250mg (p=0.44). Chrubasik 1999 (n=197) reported the number of participants using Tramadol in week 4 was 13 in the placebo group; 5 in the low dose group, and 11 in the high dose group.					⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Change in medication - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; **RR:** risk ratio

Explanations

a. Risk of bias downgraded by 1 level due to high or unclear risk of bias in random sequence generation, allocation concealment, incomplete outcome data, selective reporting, cointerventions, and compliance.

b. Imprecision downgraded by 1 level due to low number of events.
c. Inconsistency not assessed; no meta-analysis performed.
d. Imprecision downgraded by 2 levels, unable to pool data reported as relative median change from baseline and small sample size.
e. Imprecision downgraded by 2 levels, unable to pool data and small sample size. Tramadol provided by trial investigators as rescue medication, unclear what instructions to participants were.
f. Inconsistency downgraded by 1 level due to substantial heterogeneity (I^2 = 73%) not explained by subgroup analyses.
g. Imprecision downgraded by 2 levels due to wide confidence intervals that encompass a potential benefit, no effect, and a potential harm.
h. Indirectness downgraded 1 level because baseline consumption of medication not reported. Tramadol provided by trial investigators as a rescue medication.

GRADE Table 5. ***What are the benefits and harms of Devil's claw [Harpagophytum procumbens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

No trials

GRADE Table 6. ***What are the benefits and harms of Devil's claw [Harpagophytum procumbens] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care***?

No trials

GRADE Table 7. ***What are the benefits and harms of White willow [Salix spp.] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with placebo?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Salix spp.	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain - short term (reduction of at least 30% pain intensity)												
1	randomized trials	serious[a]	not serious[b]	not serious	serious[c]	none	42/140 (30.0%)	4/70 (5.7%)	RR 5.25 (1.96 to 14.05)	**243 more per 1000** (from 55 more to 746 more)	⨁⨁◯◯ Low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Pain - intermediate term or long term – no studies were identified that reported on this outcome												
Back-specific functional status – short term												
1	randomized trials	serious[a]	not serious[b]	not serious	serious	none	Percentage decline in modified Aarhus score in the placebo group median 0% (IQR -13; 5); low dose group 44% (IQR 18; 60); high dose group 54% (IQR 19; 90) (p< 0.001) (n=210).				⨁⨁◯◯ Low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Back-specific functional status - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.D4: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Salix spp.	Placebo	Relative (95% CI)	Absolute (95% CI)		
Adverse events												
1	randomized trials	serious[a]	not serious[b]	serious[f]	serious	none	3/140 (2.1%)	6/70 (8.6%)	RR 0.25 (0.06 to 0.97)	**64 fewer per 1000** (from 81 fewer to 3 fewer)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Change in medication - short term												
1	randomized trials	serious[a]	not serious[b]	serious[e]	serious[c]	none	13/140 (9.3%)	33/70 (47.1%)	RR 0.20 (0.11 to 0.35)	**377 fewer per 1000** (from 420 fewer to 306 fewer)	⨁◯◯◯ Very low	
Population subgroup 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Change in medication - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; RR: risk ratio

Explanations

a. Risk of bias downgraded 1 level due to high or unclear risk of bias in allocation concealment, selective reporting, similar groups at baseline, co-interventions, and compliance.
b. Inconsistency not assessed, only one study included in this analysis.
c. Imprecision downgraded 1 level due to few events.

d. Imprecision downgraded 1 level due to small sample size.
e. Indirectness downgraded 1 level because baseline consumption of medication not reported. Tramadol provided by trial investigators as a rescue medication.
f. Indirectness downgraded 1 level because some events may be attributed to a co-intervention (Tramadol).
g. Imprecision downgraded 1 level due to wide confidence intervals that encompass a potential benefit and no effect.

GRADE Table 8. ***What are the benefits and harms of White willow [Salix spp.] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with no intervention***?

No trials

GRADE Table 9. ***What are the benefits and harms of White willow [Salix spp.] in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared with usual care***?

No trials

E.1 Weight management

Overview of the PICO structure

Definition of the intervention	
Weight management refers to nonsurgical interventions adopting unimodal or multimodal interventions that can be delivered in a primary care or community setting and are aimed at improving outcomes for adults with CPLBP. These interventions may include weight loss for adults who are overweight or obese, weight maintenance for adults of normal body weight or weight gain interventions for adults who are underweight or malnourished. The evidence synthesis for the guideline identified trials of weight loss interventions only.	
PICO question	
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 or minimal intervention, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial)

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

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

Other Evidence-to-Decision (EtD) considerations for pharmacological and non-pharmacological weight loss interventions

Summary of *values and preferences*	
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.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of *resource 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 *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	No evidence identified

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

E.1.1 Summary of judgements: pharmacological weight loss

Domain	**All adults**	**Older people**
Benefits	Uncertain	Uncertain

Harms	Uncertain	Uncertain
Balance benefits to harms	Uncertain; probably does not favour pharmacological weight loss	Uncertain; probably does not favour pharmacological weight loss
Overall certainty	Very low	Very low
Values and preferences	Probably important uncertainty or variability	Probably important uncertainty or variability
Resource considerations	Moderate costs; varies (according to country and health system)	Moderate costs; varies (according to country and health system)
Equity and human rights	Possibly increased; uncertain; possibly reduced (especially related to stigma)	Possibly increased; uncertain; possibly reduced (especially related to stigma)
Acceptability	Yes, probably yes (among health workers); uncertain for people with CPLBP	Yes, probably yes (among health workers); uncertain for people with CPLBP
Feasibility	Probably yes, probably no, uncertain, varies	Probably yes, probably no, uncertain, varies

E.1.2 Summary of judgements: non-pharmacological weight loss

Domain	**All adults**	**Older people**
Benefits	Uncertain	Uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Uncertain	Uncertain
Overall certainty	Very low	Very low
Values and preferences	Probably important uncertainty or variability	Probably important uncertainty or variability
Resource considerations	Moderate costs; varies (according to country and health system)	Moderate costs; varies (according to country and health system)

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Equity and human rights	Possibly increased; uncertain; possibly reduced (especially related to stigma)	Possibly increased; uncertain; possibly reduced (especially related to stigma)
Acceptability	Yes, probably yes (among health workers); uncertain for people with CPLBP	Yes, probably yes (among health workers); uncertain for people with CPLBP
Feasibility	Probably yes, probably no, uncertain, varies	Probably yes, probably no, uncertain, varies

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of pharmacological weight loss interventions for adults with chronic primary low back pain compared with placebo?***

Population: People with lower back pain
Setting: Varied
Intervention: Weight loss interventions
Comparator: Placebo

Certainty Assessment								Number of participants		Effect: Absolute (95%CI)	Certainty	Comment
Outcomes	**No. studies**	**Study Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**Weight loss**	**Placebo**			
Pain intensity – post-intervention												
Pharmacological weight loss intervention vs placebo assessed with: **McGill Pain Questionnaire** Follow-up: mean 10 weeks	1[a]	RCT	Very serious[b]	Serious[c]	Serious[d]	Serious[e]	-	48	48	MD -11.4 [–16.68 to – 6.12]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.1
Population subgroup 1 by intervention - not reported (no subgroup analysis was performed; only one included study for this outcome)												
Population subgroup 2 by 60 years and over - not reported (no subgroup analysis was performed; only one included study for this outcome)												
Population subgroup 3 by gender/sex - not reported (no subgroup analysis was performed; only one included study for this outcome)												
Population subgroup 4 by presence of leg pain or radicular symptoms (no subgroup analysis was performed; only one included study for this outcome)												
Population subgroup 5 by race/ethnicity (no subgroup analysis was performed; only one included study for this outcome)												
Population subgroup 6 by regional economic development (no subgroup analysis was performed; only one included study for this outcome)												
Pain intensity – long-term follow-up												
-	-	-	-	-	-	-	-	-	-	-	-	-

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Self-reported activity limitation (Disability/Function) – post-intervention												
Pharmacological weight loss intervention vs placebo assessed with: **Oswestry LBP Questionnaire** Follow-up: mean 10 weeks	1[a]	RCT	Very serious[b]	Serious[c]	Serious[d]	Serious[e]	-	48	48	MD -4.9 [–19.45 to 9.65]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.2
Population subgroups 1, 2, 3, 4 , 5 and 6 - not reported (no subgroup analysis was performed)												
Self-reported activity limitation (Disability/Function) – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Health related quality of life – post-intervention:												
Pharmacological weight loss intervention vs placebo assessed with: **Physical subscale of Short Form-36** Follow-up: 10 weeks	1[a]	RCT	Very serious[b]	Not serious	Serious[d]	Serious[e]	-	48	48	MD -8.00 [5.07 to 10.93]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.3
Pharmacological weight loss interventions vs placebo assessed with: **Psychological subscale of Short Form-36** Follow-up: 10 weeks	1[a]	RCT	Very serious[b]	Not serious	Serious[d]	Serious[e]	-	48	48	MD 5.4 [3.14 to 7.66]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.4
	-	-	-	-	-	-	-	-	-	-	-	-
Health related quality of life – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Weight – post-intervention												
Pharmacological weight loss interventions vs placebo assessed with: **Weight (kg)** Follow-up: range 10 weeks to 12 weeks	2[a,f]	RCT	Very serious[g]	Serious[h]	Not serious	Serious[i]	-	105	103	MD -1.61 [-8.53 to 5.31]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.5

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Weight/BMI – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Psychological functioning and wellbeing – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Social participation – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Self-efficacy – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Change in use of medications – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Falls – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Adverse events – post-intervention:												
Pharmacological weight loss interventions vs placebo, assessed with: **Frequency (n/N, %.)** Follow-up: 10 to 12 weeks	2[a,f]	RCT	Very serious[g]	Not serious	Not serious	Serious[e]	-	41/105 (40.35%)	28/103 (32.7%),	RR 1.41 [0.95 to 2.10]	⨁◯◯◯ Very low	Appendix 5 Analysis 2.6
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Adverse events – long-term follow-up: no studies were identified that reported for this outcome												

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

-	-	-	-	-	-	-	-	-	-	-	-	-

Explanation

a. Muehlbacher, 2006 - 10-weeks topiramate drug compared to placebo (blinded).
b. Risk of Bias: Downgrade two levels – overall high risk of bias in single study
c. Inconsistency: Downgrade one level for unexplained variability in result (SD reported likely to be SE) and unable to contact authors to confirm.
d. Indirectness: Single study
e. Imprecision: Downgraded one level for small sample size
f. Kwon, 2021- 12-weeks orlistat plus phentermine drugs compared to phentermine plus placebo.
g. Risk of Bias: Downgrade two level overall high risk of bias in all studies
h. Inconsistency: Downgrade one level due to substantial heterogeneity (I^2=74%)
i. Imprecision: Downgrade one levels – CIs show appreciable benefit and harm; not downgraded two levels due to downgrade for inconsistency would have contributed to severity of imprecision.

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 2. ***What are the benefits and harms of non-pharmacological weight loss interventions for adults with chronic primary low back pain compared with minimal or no intervention?***

Population: People with lower back pain
Setting: Varied
Intervention: Weight loss interventions
Comparator: No or minimal care

Certainty Assessment								Number of participants		Effect: (95%CI)	Certainty	Comment
Outcomes	**No. studies**	**Study Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**Weight loss**	**No or minimal intervention**			
Pain – post-intervention												
Diet (A) or Diet and extra virgin olive oil (B) vs olive oil only (C) assessed with: **Presence of severe pain n/%** Follow-up: mean 12 weeks	1[a]	RCT	Very serious[b]	Not serious	Serious[c]	Very serious[d]	-	90	43	RR 0.94 [0.68 to 1.28]	⨁◯◯◯ Very low	Effect estimate calculated by pooling A+B vs C Appendix 5 Analysis 3.1
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed; only one included study for this outcome)												
Pain– long-term follow-up												
-	-	-	-	-	-	-	-	-	-	-	-	-
Self-reported activity limitation (Disability/Function) – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Health related quality of life – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Weight and BMI – post-intervention												

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Diet (intv A) or Diet and extra virgin olive oil (intv B) vs olive oil only (control) assessed with: **BMI change** follow-up: 12 weeks	1[a]	RCT	Very serious[b]	not serious	Serious[c]	Serious[e]	-	A: 43 B:47	43	A: −2.65±5.54 kg/m2 B: −1.64±3.47 kg/m2 C: +1.66±2.94 kg/m2	⨁◯◯◯ Very low	Estimate from single study, data otherwise not usable.
Aerobic exercise and diet (A) vs no intervention control (B) Assessed with: **Weight change from baseline (kg)** Follow-up: 4 months	1[f]	RCT	Very serious[b]	not serious	Serious[c]	Very serious[e]	-	18	18	A: - 4.3 kg B: -1.4 kg [p=0.0001]	⨁◯◯◯ Very low	Estimate from single study, data otherwise not usable.
Population subgroup 1 - not reported (no subgroup analysis was performed; single study result provided above as meta-analysis not possible due to insufficient data)												
Population subgroup analysis 2 by 60 years and over												
Aerobic exercise and diet (A) vs no intervention control (B) Assessed with: **Weight change from baseline (kg)** Follow-up: 4 months Mean age: 63 years (SD2.4)	1[f]	RCT	Very serious[b]	not serious	Serious[c]	Very serious[g]	-	18	18	A: - 4.3 kg B: -1.4 kg [p=0.0001]	⨁◯◯◯ Very low	Estimate from single study, data otherwise not usable.
Population subgroup analysis 3 by gender/sex												
Aerobic exercise and diet (A) vs no intervention control (B) Assessed with: **Weight change from baseline (kg)** Follow-up: 4 months Gender: Males	1[f]	RCT	Very serious[b]	not serious	Serious[c]	Serious[g]	-	18	18	A: - 4.3 kg B: -1.4 kg [p=0.0001]	⨁◯◯◯ Very low	Estimate from single study, data otherwise not usable.
Population subgroups 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Weight/BMI – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Psychological functioning and wellbeing – post-intervention or long-term follow-up : no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Social participation – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Self-efficacy – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Change in use of medications – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Falls – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Adverse events – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-

Explanation

a. Mendonca 2021- 12 weeks individualised meal plan (5-10% energy deficit) with or without 52mls/day of olive oil compared to 52mls of daily olive oil.
b. Risk of Bias: Downgrade two levels for overall high risk of bias in single study
c. Indirectness: Single Study
d. Imprecision: Downgraded two levels as CIs show appreciable benefit and harm and small numbers of participants
e. Imprecision: Downgraded one level for small sample size
f. Irondoust 2021- 30 days; simple dietitian prescribed 30-day weight loss meal plan containing less than 1200kcal per day. Telephone call and text message follow-up every 3 days to monitor adherence, plus NSAID celecoxib 200mg/day.
g. Imprecision: Downgraded two levels for very small sample size.

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 3. ***What are the benefits and harms of non-pharmacological weight loss interventions for adults with chronic primary low back pain compared with usual care?***

Population: People with lower back pain
Setting: varied secondary care
Intervention: Weight loss interventions
Comparator: Usual care

Certainty Assessment								Number of participants		Effect: Absolute (95%CI)	Certainty	Comment
Outcomes	**No. studies**	**Study Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**Weight loss**	**Usual Care**			
Pain intensity – post-intervention												
Weight loss interventions vs usual care assessed with: **MPQ, VAS, NRS** Follow-up: range 60 days to 26 weeks.	4[a,b,c]	RCT	Serious[d]	Very serious[e]	not serious	Serious[f]	-	167	148	SMD 0.18 [-0.46, 0.81]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.1
Population subgroup analysis 1 by intervention type												
Diet only weight loss vs usual care assessed with: **MPQ, VAS** Follow-up: range 60 days to 5 weeks	3[a,b]	RCT	Very serious[g]	Very serious[e]	Not serious	Serious[f]	-	88	68	SMD 0.39 [-0.74, 1.52]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.2
Education and weight loss coaching (diet and exercise) vs usual care assessed with **NRS** Follow-up: 26 weeks	1[c]	RCT	Not serious	Not serious	Serious[h]	Very serious[i]	-	79	80	SMD -0.19 [-0.51, 0.12]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.2
Population subgroups 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 4 by presence of leg pain or radicular symptoms												

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Weight loss interventions in patients with leg pain vs usual care assessed with: MPQ, follow-up: 60 days	1[a]	RCT	Very serious[g]	Not serious	Not serious	Serious[i]	-	48	48	SMD -0.57 [-0.97to -0.16]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.3
Weight loss interventions in patients leg pain not reported vs usual care assessed with: **VAS, NPS** Follow-up: 5 weeks to 26 weeks	3[b,c]	RCT	Serious[d]	Very serious[e]	Not serious	Serious[f]	-	119	100	SMD 0.49 [-0.38 to 1.37]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.3
Population subgroup 5 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 6 by regional economic development												
Low-/middle-income countries: Diet only weight loss vs usual care assessed with: **MPQ, VAS** Follow-up: range 60 days to 5 weeks	3[a,b]	RCT	Very serious[g]	Very serious[e]	Not serious	Serious[f]	-	88	68	SMD 0.39 [-0.74, 1.52]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.4
High income country: Education and weight loss coaching (diet and exercise) vs usual care assessed with **NRS** Follow-up: 26 weeks	1[c]	RCT	Not serious	Not serious	Serious[h]	Very serious[i]	-	79	80	SMD -0.19 [-0.51, 0.12]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.4
Pain intensity – long-term follow-up												
-	-	-	-	-	-	-	-	-	-	-	-	-
Self-reported activity limitation (Disability/Function) – post-intervention												
Weight loss interventions vs usual care assessed with: **RMDQ, Barthel Index** Follow-up: range 60 days to 26 weeks	4[a,b,c]	RCT	Very serious[g]	Serious[k]	Not serious	Serious[j]	-	126	123	SMD -0.65 [-1.12 to -0.19]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.5

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Population subgroup analysis 1 by intervention type												
Diet only weight loss interventions vs usual care assessed with: **RMDQ, Barthel Index** Follow-up: range 60 days to 5 weeks	3[a,b]	RCT	very serious[g]	Not serious	Not serious	Serious[j]	-	88	68	SMD -0.88 [-1.22 to -0.54]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.6
Education and weight loss coaching (diet and exercise) vs usual care assessed with **RMDQ** Follow-up: 26 weeks	1[c]	RCT	Serious[l]	Serious[h]	Not serious	Very serious[i]	-	38	55	SMD -0.13 [-0.54, 0.28]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.6
Population subgroups 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 4 by presence of leg pain or radicular symptoms												
Diet only weight loss interventions vs usual care assessed with: **RMDQ** Follow-up: 60 days	1[a]	RCT	Serious[g]	not serious	Serious[h]	Serious[j]	-	48	48	SMD-0.86, [-1.28 to -0.44]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.7
Diet, or weight loss coaching (diet and exercise) vs usual care assessed with: **RMDQ, Barthel Index** Follow-up: 5 weeks to 26 weeks	3[b,c]	RCT	Serious[d]	Serious[k]	not serious	Serious[j]	-	78	75	SMD -0.57 [-1.18 to 0.04]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.7
Population subgroup 5 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 6 by regional economic development												
Low-/middle-income countries: Diet only weight loss interventions vs usual care assessed with: **RMDQ, Barthel Index** Follow-up: range 60 days to 5 weeks	2[a,b]	RCT	Very serious[g]	Not serious	Not serious	Serious[j]	-	88	68	SMD -0.88 [-1.22 to -0.54]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.8

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

High income country: Education and weight loss coaching (diet and exercise) vs usual care assessed with **RMDQ** Follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[i]	-	38	55	SMD -0.13 [-0.54, 0.28]	⊕◯◯◯ Very low	Appendix 5 Analysis 1.8
Self-reported activity limitation (Disability/Function) – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Health related quality of life – post-intervention												
Education and weight loss coaching (diet and exercise) vs usual care assessed with: **SF12-v2 Physical function subscale score (PCS) and Mental subscale score (MCS)** follow-up: mean 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[i]	-	43	61	MD (PCS) 1.6 [-2.53 to 5.73] MD (MCS) 2.20 [-3.11 to 7.51]	⊕◯◯◯ Very low	Appendix 5 Analysis 1.9 and 1.10
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed; only one included study for this outcome)												
Health related quality of life – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Weight and BMI – post-intervention												
Weight loss interventions vs usual care assessed with: **Weight (kg)** follow-up: range 30 days to 26 weeks	4[a,b,c]	RCT	Very serious[g]	Not serious	Not serious	Very serious[i]	-	142	131	MD 0.84 [-2.29 to 3.98]	⊕◯◯◯ Very low	Appendix 5 Analysis 1.11
Weight loss interventions vs usual care assessed with: **BMI (kg/m²)** follow-up: range 5 weeks to 26 weeks	3 [b,c]	RCT	Serious[d]	Not serious	Not serious	Very serious[i]	-	94	83	MD 0.71 [-0.54 to 1.96]	⊕◯◯◯ Very low	Appendix 5 Analysis 1.15

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Population subgroup analysis 1 by intervention type												
Diet only weight loss interventions vs usual care assessed with: **Weight (kg)** follow-up: range 30 days to 5 weeks	3[a,b]	RCT	Very serious[g]	Not serious	Not serious	Very serious[i]	-	88	68	MD 1.06 [-2.57 to 4.69]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.12
Education and weight loss coaching (diet and exercise) vs usual care assessed with: **Weight (kg)** follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Serious[j]	-	54	63	MD 0.6 [0.0 to 1.2]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.12
Weight loss interventions vs usual care assessed with: **BMI (kg/m^2)** follow-up: range 5 weeks	2[b]	RCT	Very serious[g]	not serious	Serious[h]	Very serious[i]	-	40	20	MD 1.48 [-0.51 to 3.46]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.16
Weight loss interventions vs usual care assessed with: **BMI (kg/m^2)** follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[i]	-	54	63	MD 0.20 [-1.41 to 1.81]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.16
Population subgroups 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 4 by presence of leg pain or radicular symptoms												
Diet only weight loss interventions vs usual care assessed with: **Weight (kg)** follow-up: 30 days	1[a]	RCT	Serious[g]	not serious	Serious[h]	Very serious[i]	-	48	48	SMD 0.39 [-4.47 to 5.25	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.13
Diet, or weight loss coaching (diet and exercise) vs usual care assessed with: **Weight (kg)** follow-up: 5 weeks to 26 weeks	3[b,c]	RCT	Serious[d]	Not serious	Not serious	Very serious[i]	-	94	83	SMD 1.17 [-2.94 to 5.27]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.13

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Not possible to perform for **BMI**												
Population subgroup 5 - not reported (no subgroup analysis was performed)												
Population subgroup analysis 6 by regional economic development												
Low-/middle-income countries: Diet only weight loss interventions vs usual care assessed with: **Weight (kg)** Barthel Index follow-up: range 30 days to 5 weeks	3[a,b]	RCT	Very serious[g]	Not serious	Not serious	Very serious[i]	-	88	68	MD 1.06, [-2.57, 4.69]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.14
High income country: Education and weight loss coaching (diet and exercise) vs usual care assessed with: **Weight (kg)** follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Serious[j]	-	54	63	MD 0.6 [0.0 to 1.2]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.14
Weight loss interventions vs usual care assessed with: **BMI (kg/m²)** follow-up: range 5 weeks	2[b]	RCT	Very serious[g]	Not serious	Serious[h]	Very serious[i]	-	40	20	MD 1.48 [-0.51 to 3.46]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.17
Weight loss interventions vs usual care assessed with: **BMI (kg/m²)** follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[j]	-	94	83	MD 0.20 [-1.41 to 1.81]	⨁◯◯ ◯ Very low	Appendix 5 Analysis 1.17
Psychological functioning and wellbeing – post-intervention:												

Web Annex D.E1: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Education and weight loss coaching (diet and exercise) vs usual care with: **Depression anxiety stress scale (DASS)** Depression Anxiety Stress Follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[i]	-	43	61	Depression MD 1.20 [-3.15 to 5.55] Anxiety MD 0.4 [-2.95 to 3.75] Stress MD 0.5 [-3.74 to 4.74]	⨁◯◯◯ Very low	Appendix 5 Analysis 1.18 to 1.20
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Psychological functioning and wellbeing – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Social participation – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Self-efficacy – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Change in use of medications – post-intervention												
Education and weight loss coaching (diet and exercise) vs usual care assessed with: **Frequency n/N** Follow-up: 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Very serious[i]	-	27/38	45/56	RR 0.88 (0.7 to 1.12)	⨁◯◯◯ Very low	Appendix 5 Analysis 1.21
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Change in use of medications – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-
Falls – post-intervention or long-term follow-up: no studies were identified that reported for this outcome												

-	-	-	-	-	-	-	-	-	-	-	-	-
Adverse events – post-intervention:												
Education and weight loss coaching (diet and exercise) vs usual care assessed with: **Frequency n/N** Follow-up: range 26 weeks	1[c]	RCT	Serious[l]	Not serious	Serious[h]	Serious[j]	-	32/79	45/80	RR 0.72 (0.52 to 1.00)	⨁◯◯◯ Very low	Appendix 5 Analysis 1.22
Population subgroups 1, 2, 3, 4, 5 and 6 - not reported (no subgroup analysis was performed)												
Adverse events – long-term follow-up: no studies were identified that reported for this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	-

Explanation

a. Safari 2020, 30 day Low calorie prescribed diet intervention (1200kcal/day) plus 200mg celecoxib per day vs 200mg celecoxib/day only.
b. Torlak 2022 contributes as 2 studies in the analyses as it had two weight loss intervention arms and one shared comparator group. Weight intervention consisted of a 5 week 5:2 intermittent diet consisting of two days consuming 600-700kcal/day and 5 days 1500-1700kcal per day Mediterranean diet with or without physiotherapy care (TENS and hotpack) compared to physiotherapy care only.
c. Williams 2018 One face to face pain and lifestyle education session plus 6-month telephone weight loss health coaching for diet and physical activity compared to usual care.
d. Risk of Bias: Downgrade one level for overall risk of bias in two studies (>25% of participants)
e. Inconsistency: Downgrade two levels for high, unexplained heterogeneity > 75%
f. Imprecision: Downgrade one level - CIs and point estimates show appreciable benefit and harm; not downgraded two levels due to downgrade for inconsistency would have contributed to severity of imprecision.
g. Risk of bias: Downgrade two levels for overall high risk of bias in most studies (>50% of participants)
h. Indirectness: Single study
i. Imprecision: Downgrade two levels CIs show appreciable benefit and harm and small numbers of participants
j. Imprecision: Downgrade one level for small number of participants – fewer than 400.
k. Inconsistency: Downgrade one level for inconsistency, heterogeneity > 50%
l. Risk of bias: Downgrade one level - risk of bias due to loss to follow-up for that outcome.

E.2 Multicomponent biopsychosocial care

Overview of the PICO structure

Definition of the intervention	
Multicomponent biopsychosocial care involves delivery of at least two of the three components of care from the biopsychosocial model: physical, psychological or social, delivered by a single provider or a multidisciplinary team. These components align with the biopsychosocial model of chronic pain and its applicability to older people. Multicomponent biopsychosocial care adopts a rehabilitation approach that aims to optimize function and reduce disability in individuals with health conditions in interaction with their environment. For the purpose of the guideline, trials of all types of interventions for multicomponent biopsychosocial care were included where they satisfied the criterion of a multicomponent intervention that targets *functioning* (body structures and functions, activities and participation). The intervention should target at least two domains of the biopsychosocial model: either the biological component targeting physical aspects of functioning such as body structures or functions (e.g. an exercise programme targeting an increase in muscle strength), psychological component (e.g. addressing coping with pain) or social and occupational component (e.g. addressing involvement in meaningful life roles including work).	
PICO question	
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/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 or minimal intervention or comparators, or where the effect of the intervention can be isolated c) Usual care (described as usual care in the trial) including care where the intervention can be isolated
Outcomes	Critical outcomes constructs (all adults) Critical outcomes constructs (older adults, aged ≥ 60 years) • Pain • Back-specific function/disability • General function/disability • Health-related quality of life • Psychosocial function • Social participation • Self-efficacy • 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

Other Evidence-to-Decision (EtD) considerations

Summary of *values and preferences*	
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.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Summary of *resource 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 *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	No evidence identified

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

Summary of judgements

Multicomponent biopsychosocial care (single provider)

Domain	**All adults**	**Older people**

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Benefits	Small; uncertain	Small; uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Probably favours single-provider multicomponent biopsychosocial care (single provider); uncertain	Probably favours single-provider multicomponent biopsychosocial care (single provider); uncertain
Overall certainty	Very low	Very low
Values and preferences	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies
Equity and human rights	Increased; probably increased; probably reduced; reduced; varies	Increased; probably increased; probably reduced; reduced; varies
Acceptability	Yes; probably yes; varies	Yes; probably yes; varies
Feasibility	Yes; probably yes; probably no; varies	Yes; probably yes; probably no; varies

Multicomponent biopsychosocial care (MDT provider)

Benefits	Moderate; small; uncertain	Small; uncertain
Harms	Uncertain	Uncertain
Balance benefits to harms	Probably favours multicomponent biopsychosocial care (MDT provider); uncertain	Probably favours multicomponent biopsychosocial care (MDT provider); uncertain
Overall certainty	Low; very low	Low; very low
Values and preferences	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty	Important uncertainty; possibly important uncertainty or variability; probably no important uncertainty
Resource considerations	Large costs; moderate costs; varies	Large costs; moderate costs; varies

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Equity and human rights	Increased; probably increased; probably reduced; reduced; varies	Increased, probably increased; probably reduced; reduced; varies
Acceptability	Yes; probably yes; varies	Yes; probably yes; varies
Feasibility	Yes; probably yes; probably no; varies	Yes; probably yes; probably no; varies

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

GRADE Table 1. ***What are the benefits and harms of multicomponent biopsychosocial care in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to placebo?***

No trials

GRADE Table 2. ***What are the benefits and harms of multicomponent biopsychosocial care delivered by a multidisciplinary team in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to no intervention?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	No intervention	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
3[a]	randomized trials	very serious[b]	Not serious[c]	not serious	serious[d]	none	106	107	-	SMD **0.73 SD lower** (1.22 lower to 0.24 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Pain - intermediate term or long term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status – short term												
3[a]	randomized trials	very serious[b]	not serious	not serious	serious[e]	none	106	107	-	SMD **0.49 SD lower** (0.76 lower to 0.22 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Back-specific functional status - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
General functional status - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term, intermediate term or long term: no studies were identified that reported on this outcome												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	No intervention	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events or serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term (lower score means less depression)												
3[a]	randomized trials	very serious[b]	not serious	not serious	serious[f]	none	106	107	-	SMD **0.21 SD lower** (0.59 lower to 0.18 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Psychological functioning - intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; SMD: standardized mean difference

Explanations

a. Jäckel 1990, Smeets 2006, Turner 1990

b. Risk of bias downgraded by 2 levels for unclear or serious risk of bias in all studies for random sequence generation, allocation concealment, blinding of participants, clinicians, and outcome assessors, incomplete outcome data, selective reporting, compliance, and co-interventions.

c. Despite some heterogeneity (I-sq = 64%), not downgraded for inconsistency because direction of effect was same from all studies.

d. Imprecision downgraded by 1 level for wide confidence intervals that encompass a potential benefit and little to no effect. We re-expressed the SMD as mean difference on a 0 to 100 pain scale using an SD of 22.6 (i.e. control group SD from Smeets 2006) which gave MD -16.5 (-27.6 to -5.4). The minimal important difference on the 0 to 100 pain scale is approximately 15.

e. Imprecision downgraded by 1 level for wide confidence intervals that encompass a potential benefit and little to no effect. We re-expressed the SMD as mean difference on a 0 to 24 RDQ scale using an SD of 4.78 (i.e. control group SD from Smeets 2006) which gave MD -2.3 (-3.6 to -1.1). The minimal important difference on the 0 to 24 RDQ pain scale is approximately 10%.

f. Imprecision downgraded by 1 level for small sample size.

GRADE Table 3. ***What are the benefits and harms of multicomponent biopsychosocial care delivered by a single provider in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Rehabilitation	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - short term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Pain - intermediate term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Pain - long term (two-point reduction of pain intensity from 11-point scale)												
1[a]	randomized trials	very serious[b]	not serious[c]	not serious	serious[d]	none	29/60 (48.3%)	20/54 (37.0%)	**RR 1.30** (0.84 to 2.02)	**111 more per 1000** (from 59 fewer to 378 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Back-specific functional status – short term or intermediate term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Back-specific functional status - long term (30% improvement)												
1[a]	randomized trials	very serious[b]	not serious[c]	not serious	serious[d]	none	34/60 (56.7%)	26/54 (48.1%)	**RR 1.18** (0.83 to 1.68)	**87 more per 1000** (from 82 fewer to 327 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
General functional status – short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Rehabilitation	Usual care	Relative (95% CI)	Absolute (95% CI)		
Health-related quality of life - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events												
1[a]	randomized trials	very serious[b]	not serious[c]	not serious	very serious[e]	none	0/60	0/54	RR not estimable	-	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one included study on this outcome)												
Serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Social participation - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Self-efficacy - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; RR: risk ratio

Explanations

a. van der Roer 2008

b. Risk of bias downgraded by 2 levels due to unclear or high risk of bias in blinding of participants, clinicians, and outcome assessors, selective reporting, compliance, and co-interventions.

c. Inconsistency not assessed, only one study included on this outcome

d. Imprecision downgraded by 1 level due to wide confidence intervals that encompass a potential benefit and no effect with intervention.

e. Imprecision downgraded by 2 levels due to no events reported.

GRADE Table 4. ***What are the benefits and harms of multicomponent biopsychosocial care delivered by a multidisciplinary team in the management of community-dwelling adults (including older adults aged 60 years and over) with chronic primary low back pain (with or without leg pain) compared to usual care?***

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - short term												
10[a]	randomized trials	very serious[b]	serious[c]	not serious	serious[d]	none	478	495	-	SMD **0.52 SD lower** (0.77 lower to 0.27 lower)	⨁◯◯◯ Very low	
Population subgroup 1: gender/sex												
Female only 1	randomized trials	very serious[b]	not serious[o]	not serious	serious[l]	none	44	47	-	SMD **0.61 SD lower** (1.03 lower to 0.19 lower)	⨁◯◯◯ Very low	
Mixed 9	randomized trials	very serious[b]	serious[c]	not serious	serious[dl]	none	434	448	-	SMD **0.51 SD lower** (0.79 lower to 0.23 lower)	⨁◯◯◯ Very low	
Population subgroup 2: race/ethnicity - not reported (no subgroup analysis was performed; no study included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 1	randomized trials	very serious[b]	not serious[o]	not serious	serious[l]	none	12	11	-	SMD **0.32 SD lower** (1.14 lower to 0.51 higher)	⨁◯◯◯ Very low	

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Mixed 9	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	466	484	-	SMD **0.53 SD lower** (0.8 lower to 0.27 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
Low/middle income 3	randomized trials	very serious[b]	not serious	not serious	serious[p]	none	148	155	-	SMD **0.46 SD lower** (0.69 lower to 0.23 lower)	⨁◯◯◯ Very low	
High income 7	randomized trials	very serious[b]	serious[c]	not serious	serious[d]	none	330	340	-	SMD **0.56 SD lower** (0.92 lower to 0.19 lower)	⨁◯◯◯ Very low	
Pain - intermediate term												
5[e]	randomized trials	very serious[b]	serious[c]	not serious	serious[f]	none	326	320	-	SMD **0.62 SD lower** (0.93 lower to 0.31 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2 and 3 - not reported (no subgroup analysis was performed)												
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[b]	not serious[o]	not serious	serious[l]	none	92	96	-	SMD **0.49 SD lower** (0.78 lower to 0.2 lower)	⨁◯◯◯ Very low	
High income 4	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	234	224	-	SMD **0.68 SD lower** (1.12 lower to 0.25 lower)	⨁◯◯◯ Very low	

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Pain - long term												
8[f]	randomized trials	very serious[b]	not serious	not serious	not serious	none	517	446	-	SMD **0.25 SD lower** (0.41 lower to 0.09 lower)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 1	randomized trials	very serious[b]	not serious[o]	not serious	Serious[l]	none	12	11	-	SMD **0.28 SD lower** (-1.1 lower to 0.54 higher)	⨁◯◯◯ Very low	
Mixed 7	randomized trials	very serious[b]	not serious	not serious	not serious	none	505	435	-	SMD **0.25 SD lower** (0.43 lower to 0.08 lower)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development												
Low/middle income 2	randomized trials	very serious[b]	not serious	not serious	serious[l]	none	81	88	-	SMD **0.47 SD lower** (0.77 lower to 0.16 lower)	⨁◯◯◯ Very low	
High income 6	randomized trials	very serious[b]	not serious	not serious	not serious	none	436	358	-	SMD **0.21 SD lower** (0.39 lower to 0.03 lower)	⨁⨁◯◯ Low	
Back-specific functional status – short term												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
10[g]	randomized trials	very serious[b]	serious[c]	not serious	Not serious	none	506	527	-	SMD **0.47 SD lower** (0.69 lower to 0.24 lower)	⨁◯◯◯ Very low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 2	randomized trials	very serious[b]	serious[c]	not serious	Very serious[s]	none	84	90	-	SMD **0.1 SD higher** (1.01 lower to 1.22 higher)	⨁◯◯◯ Very low	
Mixed 8	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	422	437	-	SMD **0.55 SD lower** (0.78 lower to 0.31 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
Low/middle income 2	randomized trials	very serious[b]	serious[c]	not serious	Very serious[s]	none	104	108	-	SMD **0.16 SD higher** (0.83 lower to 1.14 higher)	⨁◯◯◯ Very low	
High income 8	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	402	419	-	SMD **0.57 SD lower** (0.79 lower to 0.34 lower)	⨁◯◯◯ Very low	
Back-specific functional status - intermediate term												
6[h]	randomized trials	very serious[b]	serious[c]	not serious	Not serious	none	394	392	-	SMD **0.43 SD lower** (0.66 lower to 0.19 lower)	⨁◯◯◯ Very low	

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 1	randomized trials	very serious[b]	Not serious[o]	not serious	serious[p]	none	68	72	-	SMD **0.2 SD lower** (0.53 lower to 0.13 lower)	⨁◯◯◯ Very low	
Mixed 5	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	326	320	-	SMD **0.49 SD lower** (0.77 lower to 0.21 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[b]	Not serious[o]	not serious	serious[p]	none	92	96	-	SMD **0.32 SD lower** (0.6 lower to 0.03 lower)	⨁◯◯◯ Very low	
High income 5	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	302	296	-	SMD **0.47 SD lower** (0.77 lower to 0.17 lower)	⨁◯◯◯ Very low	
Back-specific functional status - long term												
7[i]	randomized trials	very serious[b]	not serious	not serious	not serious	none	467	397	-	SMD **0.25 SD lower** (0.4 lower to 0.11 lower)	⨁⨁◯◯ Low	
Population subgroups 1 and 2 - not reported (no subgroup analysis was performed)												
Population subgroup 3: presence of radicular leg pain												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Excluded leg pain 1	randomized trials	very serious[b]	not serious[o]	not serious	Very serious[s]	none	12	11	-	SMD **0.26 SD lower** (1.08 lower to 0.57 higher)	⨁◯◯◯ Very low	
Mixed 6	randomized trials	very serious[b]	not serious	not serious	not serious	none	455	386	-	SMD **0.26 SD lower** (0.42 lower to 0.09 lower)	⨁⨁◯◯ Low	
Population subgroup 4: regional economic development												
Low/middle income 2	randomized trials	very serious[b]	not serious	not serious	Serious[p]	none	81	88	-	SMD **0.34 SD lower** (0.65 lower to 0.04 lower)	⨁◯◯◯ Very low	
High income 5	randomized trials	very serious[b]	not serious	not serious	not serious	none	386	309	-	SMD **0.24 SD lower** (0.43 lower to 0.05 lower)	⨁⨁◯◯ Low	
General functional status - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Health-related quality of life - short term												
3[j]	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	151	143	-	SMD **0.4 SD lower** (1.11 lower to 0.31 higher)	⨁◯◯◯ Very low	
Population subgroup 1: gender/sex												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
Female 1	randomized trials	very serious[b]	Not serious[o]	not serious	serious[p]	none	37	37	-	SMD **1.08 SD lower** (1.57 lower to 0.59 lower)	⨁◯◯ ◯ Very low	
Mixed 2	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	114	106	-	SMD **0.05 SD lower** (0.49 lower to 0.38 higher)	⨁◯◯ ◯ Very low	
Population subgroup 2: race/ethnicity - not reported (no subgroup analysis was performed; no study included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 1	randomized trials	very serious[b]	Not serious[o]	not serious	serious[l]	none	73	77	-	SMD **0.14 SD higher** (0.18 lower to 0.46 higher)	⨁◯◯ ◯ Very low	
Mixed 2	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	78	66	-	SMD **0.7 SD lower** (1.45 lower to 0.05 higher)	⨁◯◯ ◯ Very low	
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[b]	Not serious[o]	not serious	serious[p]	none	37	37	-	SMD **1.08 SD lower** (1.57 lower to 0.59 lower)	⨁◯◯ ◯ Very low	
High income 2	randomized trials	very serious[b]	serious[c]	not serious	serious[l]	none	114	106	-	SMD **0.05 SD lower** (0.49 lower to 0.38 higher)	⨁◯◯ ◯ Very low	
Health-related quality of life - intermediate term												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
3[j]	randomized trials	very serious[b]	not serious[k]	not serious	not serious	none	147	137	-	SMD **0.23 SD lower** (0.58 lower to 0.12 higher)	⨁⨁◯◯ Low	
Population subgroup 1: gender/sex												
Female 1	randomized trials	very serious[b]	not serious[o]	not serious	Serious[p]	none	37	37	-	SMD **0.54 SD lower** (1.01 lower to 0.08 lower)	⨁◯◯◯ Very low	
Mixed 2	randomized trials	very serious[b]	not serious	not serious	Serious[l]	none	110	100	-	SMD **0.08 SD lower** (0.38 lower to 0.23 higher)	⨁◯◯◯ Very low	
Population subgroup 2: race/ethnicity - not reported (no subgroup analysis was performed; no study included marginalized populations)												
Population subgroup 3: presence of radicular leg pain												
Excluded leg pain 1	randomized trials	very serious[b]	not serious[o]	not serious	Serious[l]	none	69	71	-	SMD **0.04 SD higher** (0.29 lower to 0.37 higher)	⨁◯◯◯ Very low	
Mixed 2	randomized trials	very serious[b]	not serious	not serious	Serious[p]	none	78	66	-	SMD **0.42 SD lower** (0.75 lower to 0.08 lower)	⨁◯◯◯ Very low	
Population subgroup 4: regional economic development												
Low/middle income 1	randomized trials	very serious[b]	not serious[o]	not serious	Serious[p]	none	37	37	-	SMD **0.54 SD lower** (1.01 lower to 0.08 lower)	⨁◯◯◯ Very low	

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
High income 2	randomized trials	very serious[b]	not serious	not serious	Serious[l]	none	110	100	-	SMD **0.08 SD lower** (0.38 lower to 0.23 higher)	⨁◯◯◯ Very low	
Health-related quality of life - long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Adverse events or serious adverse events: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - short term												
1[l]	randomized trials	very serious[b]	not serious[o]	not serious	Serious[p]	none	13	15	-	MD **4.4 lower** (9.99 lower to 1.19 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study included on this outcome)												
Psychological functioning - intermediate term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (depression) - long term												
1[n]	randomized trials	very serious[b]	not serious[o]	not serious	Serious[p]	none	61	43	-	MD **0.7 lower** (2.27 lower to 0.87 higher)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study included on this outcome)												
Psychological functioning (anxiety) - short term												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
1[l]	randomized trials	very serious[b]	not serious[o]	not serious	serious[p]	none	13	15	-	MD **12.3 lower** (20.52 lower to 4.08 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study included on this outcome)												
Psychological functioning (anxiety) - intermediate term – no studies identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	
Psychological functioning (anxiety) - long term												
1[n]	randomized trials	very serious[b]	not serious[o]	not serious	serious[p]	none	61	43	-	MD **1.9 lower** (3.65 lower to 0.15 lower)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed; only one study included on this outcome)												
Social participation (work) - short term												
3[p]	randomized trials	very serious[b]	serious[c]	not serious	very serious[s]	none	157/212 (74.1%)	162/255 (63.5%)	**RR 1.30** (0.73 to 2.34)	**191 more per 1000** (from 172 fewer to 851 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Social participation (work) - intermediate term												
2[r]	randomized trials	very serious[b]	serious[c]	not serious	very serious[s]	none	133/167 (79.6%)	144/196 (73.5%)	**RR 1.08** (0.73 to 1.60)	**59 more per 1000** (from 198 fewer to 441 more)	⨁◯◯◯ Very low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Social participation - long term												

Web Annex D.E2: ETD summary for WHO Guideline on non-surgical management of chronic primary low back pain in adults

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	MBR	Usual care	Relative (95% CI)	Absolute (95% CI)		
7[s]	randomized trials	very serious[b]	not serious	not serious	not serious	none	526/701 (75.0%)	483/648 (74.5%)	**RR 1.00** (0.93 to 1.08)	**0 fewer per 1000** (from 52 fewer to 60 more)	⨁⨁◯◯ Low	
Population subgroups 1, 2, 3 and 4 - not reported (no subgroup analysis was performed)												
Self-efficacy - short term, intermediate term or long term: no studies were identified that reported on this outcome												
-	-	-	-	-	-	-	-	-	-	-	-	

CI: confidence interval; MD: mean difference; RR: risk ratio; SMD: standardized mean difference

Explanations

a. Abbasi 2012, Basler 1997, Bendix 1996, Lambeek 2010, Moix 2003, Morone 2011, Morone 2012, Tavafian 2007, Tavafian 2011, Von Korff 2005
b. Risk of bias downgraded by 2 levels for unclear or high risk of bias in all studies for random sequence generation, allocation concealment, blinding of participants, clinicians, and outcome assessors, incomplete outcome data, selective reporting, compliance, and co-interventions.
c. Inconsistency downgraded by 1 level for substantial statistical heterogeneity not explained by subgroup analyses (I-sq > 60%)
d. Imprecision downgraded by 1 level for wide confidence intervals that encompass a potential benefit and little to no effect. We re-expressed the SMD as mean difference on a 0 to 100 pain scale using an SD of 20 (i.e. average SD from control groups that used this scale) which gave MD -10.4 (-15.4 to -5.4). The minimal important difference on the 0 to 100 pain scale is approximately 15.
e. Lambeek 2010, Morone 2011, Morone 2012, Tavafian 2011, Von Korff 2005
f. Imprecision downgraded by 1 level for wide confidence intervals that encompass a potential benefit and little to no effect. We re-expressed the SMD as mean difference on a 0 to 100 pain scale using an SD of 20 (i.e. average SD from control groups that used this scale) which gave MD -12.4 (-18.6 to -6.2). The minimal important difference on the 0 to 100 pain scale is approximately 15.
g. Abbasi 2012, Bendix 1996, Lambeek 2010, Linton 2005, Lukinmaa 1989, Strand 2001, Tavafian 2011, Von Korff 2005
h. Abbasi 2012, Basler 1997, Bendix 1996, Lambeek 2010, Moix 2003, Morone 2011, Morone 2012, Tavafian 2011, Vollenbroek-Hutten 2004, Von Korff 2005
i. Lambeek 2010, Morone 2011, Morone 2012, Tavafian 2011, Vollenbroek-Hutten 2004, Von Korff 2005
j. Abbasi 2012, Lambeek 2010, Linton 2005, Lukinmaa 1989, Strand 2001, Tavafian 2011, Von Korff 2005
k. Morone 2011, Tavafian 2007, Vollenbroek-Hutten 2004
l. Imprecision downgraded by 1 level for wide confidence intervals that encompass a potential benefit and little to no effect.
m. Despite some statistical heterogeneity, this was largely explained by the subgroup analyses.
n. Moix 2003
o. Inconsistency not assessed, only one study included on this outcome
p. Imprecision downgraded by 1 level due to small sample size.
q. Linton 2005
r. Bendix 1996, Skouen 2002, Von Korff 2005
s. Imprecision downgraded by 2 levels for very wide confidence intervals that encompass a potential harm, no effect, and a potential benefit.
t. Skouen 2002, Von Korff 2005
u. Bendix 1996, Linton 2005, Lukinmaa 1989, Mitchell 1994, Skouen 2002, Strand 2001, Von Korff 2005.

Ageing and Health Unit, Department of Maternal, Newborn, Child and Adolescent Health and Ageing
World Health Organization
20 Avenue, Appia 1211, Geneva 27, Switzerland
Website: www.who.int

9789240085701