Grading of Recommendations Assessment, Development and Evaluation (GRADE) clinical evidence review ratings of the evidence for the key clinical questions[1] in the series of systematic clinical evidence reviews on the benefits and harms of treatments for chronic and acute pain.

Opioids for Chronic Pain. This table is based on Chou R, Hartung D, Turner J, et al. AHRQ Comparative Effectiveness Reviews. Opioid Treatments for Chronic Pain. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020. The strength of evidence ratings in the AHRQ report were converted to ACIP-adapted GRADE evidence type ratings.

Intervention	Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Study Limitations	Consistency	Findings† (95% CI)	Evidence Type
Opioid vs. placebo or no opioid therapy	Pain (short-term)	71 RCTs (continuous); 44 RCTs (dichotomous)	19,616 (continuous); 12,481 (dichotomous)	Direct	Precise	Low	Consistent	MD -0.79 (-0.93 to -0.67); RR 1.35 (1.24 to 1.48)	1
	Function (short-term)	44 RCTs	12,427	Direct	Precise	Low	Consistent	SMD -0.22 (-0.28 to -0.16)	1
	SF-36 physical (short-term)	23 RCTs	8005	Direct	Precise	Low	Consistent	MD 1.64 (1.10 to 2.17)	1
	SF-36 mental (short-term)	21 RCTs	7586	Direct	Precise	Low	Consistent	MD -0.48 (-1.39 to 0.44)	1
	Sleep quality (short-term)	25 RCTs	6720	Direct	Precise	Low	Consistent	SMD -0.25 (-0.32 to -0.19)	2
	Depression (short-term)	8 RCTs	1079	Direct	Imprecise	Low	Consistent	SMD 0.00 (-0.22 to 0.18)	2
	Anxiety (short-term)	2 RCTs	229	Direct	Imprecise	Moderate	Consistent	MD 0.60 (-3.58 to 1.82)	3
	Pain and function (long-term)	1 cohort study	529	Direct	Precise	Moderate	Unknown	No differences at 2 years	4
	Discontinuation due to AEs	61 RCTs	19,994	Direct	Precise	Low	Consistent	RR 2.25 (1.86 to 2.73)	**1**
	Serious AEs	38 RCTs	13,160	Direct	Imprecise	Low	Consistent	RR 1.23 (0.88 to 1.74)	2
	Nausea	60 RCTs	19,718	Direct	Precise	Low	Consistent	RR 2.46 (2.17 to 2.80)	1
	Vomiting	49 RCTs	17,388	Direct	Precise	Low	Consistent	RR 3.57 (2.98 to 4.34)	1
	Constipation	58 RCTs	19,351	Direct	Precise	Low	Consistent	RR 3.38 (2.96 to 3.92)	1
	Dizziness	53 RCTs	18,396	Direct	Precise	Low	Consistent	RR 2.66 (2.37 to 2.99)	1
	Headache	48 RCTs	17,405	Direct	Precise	Low	Consistent	RR 1.06 (0.95 to 1.17)	1
	Somnolence	52 RCTs	17,458	Direct	Precise	Low	Consistent	RR 2.97 (2.44 to 3.66)	1

[1] All outcomes were prioritized as important or critical relative to the GRADE thresholds.

	Pruritus	30 RCTs	11,454	Direct	Precise	Low	Consistent	RR 3.51 (2.47 to 5.16)	1
	Opioid abuse, dependence, or addiction	2 cohort studies	666,780	Direct	Precise	Moderate	Consistent	Opioids associated with increased risk	3
	Overdose	2 cohort studies	108,080	Direct	Precise	Moderate	Consistent	Opioids associated with increased risk	3
	All-cause mortality	1 cohort study	22,912	Direct	Precise	Moderate	Unknown	Opioids associated with increased risk	3
	Fracture	6 observational studies	48,250	Direct	Precise	Moderate	Consistent	Opioids associated with increased risk	3
	Cardiovascular events	3 cohort studies	505,626	Direct	Precise	Moderate	Consistent	Opioids associated with increased risk	3
	Endocrinological harms	1 cross-sectional analysis	11,327	Direct	Precise	Moderate	Unknown	Unable to determine	Insufficient
Opioids vs. nonopioids	Pain (short-term)	14 RCTs (continuous); 12 RCTs (dichotomous)	2195 (continuous); 2887 (dichotomous)	Direct	Precise	Moderate	Inconsistent	MD -0.29 (-0.61 to 0.03); RR 1.28 (0.90 to 1.85)	2
	Function (short-term)	11 RCTs	2010	Direct	Precise	Moderate	Consistent	SMD 0.00 (-0.14 to 0.12)	1
	SF-36 physical (short-term)	6 RCTs	1423	Direct	Imprecise	Moderate	Consistent	MD -1.80 (-5.45 to -0.12)	2
	SF-36 mental (short-term)	6 RCTs	1427	Direct	Precise	Moderate	Consistent	MD -0.63 (-4.27 to 0.91)	2
	Sleep quality (short-term)	7 RCTs	1694	Direct	Precise	Moderate	Consistent	SMD 0.02 (-0.10 to 0.12)	2
	Depression (short-term)	7 RCTs	748	Direct	Imprecise	Moderate	Consistent	SMD 0.05 (-0.09 to 0.22)	2
	Anxiety (short-term)	3 RCTs	414	Direct	Imprecise	Moderate	Consistent	SMD 0.00 (-0.62 to 0.36)	3
	Discontinuation due to AEs	12 RCTs	3637	Direct	Precise	Low	Inconsistent	RR 2.18 (1.48 to 3.08)	2
	Serious AEs	4 RCTs	1949	Direct	Imprecise	Low	Consistent	RR 0.63 (0.06 to 5.66)	2
	Nausea	11 RCTs	3137	Direct	Precise	Low	Consistent	RR 2.77 (2.09 to 4.18)	1
	Vomiting	6 RCTs	2644	Direct	Precise	Low	Consistent	RR 4.62 (2.94 to 7.24)	1
	Constipation	12 RCTs	3377	Direct	Precise	Low	Inconsistent	RR 2.92 (1.80 to 5.21)	2

Opioids vs. nonopioids, continued	Dizziness	12 RCTs	3377	Direct	Imprecise	Low	Inconsistent	RR 1.33 (0.78 to 2.05)[ll] • NSAID: 2.12 (1.45 to 3.00) • Gabapentinoid: 0.60 (0.15 to 1.09) • Nortriptyline: 1.31 (0.64 to 4.27)	3
	Headache	8 RCTs	2791	Direct	Precise	Low	Consistent	RR 1.35 (1.08 to 1.70)	1
	Somnolence	12 RCTs	3377	Direct	Precise	Low	Inconsistent	RR 2.11 (1.39 to 23.47)	2
	Pruritus	5 RCTs	2577	Direct	Precise	Low	Consistent	RR 4.22 (2.45 to 8.20)	1
	Opioid abuse, dependence, or addiction	No studies	--	--	--	--	--	--	--
	Overdose	No studies	--	--	--	--	--	--	--
	All-cause mortality	No studies	--	--	--	--	--	--	--
	Fracture	No studies	--	--	--	--	--	--	--
	Cardiovascular events	No Studies	--	--	--	--	--	--	--
	Endocrinological harms	No studies	--	--	--	--	--	--	--
Opioid + nonopioid vs. nonopioid	Pain (short-term)	6 RCTs (continuous); 6 RCTs (dichotomous)	628 (continuous); 765 (dichotomous)	Direct	Imprecise	Moderate	Consistent	MD -0.36 (-1.14 to 0.53); RR 1.46 (0.76 to 2.74)	3
	Function (short-term)	4 RCTs	549	Direct	Imprecise	Moderate	Consistent	SMD -0.26 (-0.63 to 0.17)	3
	SF-36 physical (short-term)	4 RCTs	297	Direct	Imprecise	Moderate	Consistent	SMD 0.58 (-4.19 to 4.37)	3
	SF-36 mental (short-term)	4 RCTs	297	Direct	Imprecise	Moderate	Consistent	SMD -2.92 (-6.30 to 0.46)	3
	Sleep quality (short-term)	3 RCTs	446	Direct	Imprecise	Moderate	Consistent	SMD 0.01 (-0.21 to 0.29)	3
	Depression (short-term)	3 RCTs	246	Direct	Imprecise	Moderate	Consistent	SMD -0.01 (-0.31 to 0.26)	3

Opioid + nonopioid vs. nonopioid, continued	Anxiety (short-term)	No studies	--	--	--	--	--	--	--
	Discontinuation due to AEs	6 RCTs	707	Direct	Imprecise	Moderate	Consistent	RR 1.99 (0.89 to 4.26)	3
	Serious AEs	1 RCT	62	Direct	Imprecise	Moderate	Unable to assess	RR 0.38 (0.02 to 8.93)	Insufficient
	Nausea	5 RCTs	330	Direct	Precise	Moderate	Consistent	RR 2.18 (1.16 to 6.49)	2
	Vomiting	2 RCTs	81	Direct	Imprecise	Moderate	Consistent	RR 1.68 (0.43 to 6.56)	3
	Constipation	6 RCTs	633	Direct	Precise	Moderate	Consistent¶	RR 2.74 (1.28 to 7.44)	2
	Dizziness	6 RCTs	633	Direct	Imprecise	Moderate	Consistent	RR 1.30 (0.12 to 2.09)	3
	Headache	3 RCTs	137	Direct	Imprecise	Moderate	Consistent	RR 1.18 (0.42 to 3.00)	3
	Somnolence	6 RCTs	663	Direct	Precise**	Moderate	Consistent¶	RR 1.39 (0.41 to 5.25); excluding poor quality trial RR 2.44 (1.32 to 4.52)	2
	Pruritus	2 RCTs	148	Direct	Imprecise	Moderate	Consistent	RR 3.49 (0.32 to 37.88)	3
	Opioid abuse, dependence, or addiction	No studies	--	--	--	--	--	--	--
	Overdose	No studies	--	--	--	--	--	--	--
	All-cause mortality	No studies	--	--	--	--	--	--	--
	Fracture	No studies	--	--	--	--	--	--	--
	Cardiovascular events	No studies	--	--	--	--	--	--	--
	Endocrinological harms	No studies	--	--	--	--	--	--	--
Opioid + nonopioid vs. opioid alone	Pain (short-term)	6 RCTs (continuous); 5 RCTs (dichotomous)	854 (continuous); 831 (dichotomous)	Direct	Imprecise	Moderate	Consistent	MD -0.18 (-0.72 to -0.36); RR 1.19 (0.97 to 1.68)	3
	Function	4 RCTs	521	Direct	Imprecise	Moderate	Consistent	SMD -0.25 (-0.49 to 0.09)	3
	SF-36 physical (short-term)	4 RCTs	553	Direct	Imprecise	Moderate	Consistent	SMD -0.19 (-2.48 to 4.08)	3

Opioid + non-opioid vs. opioid alone, continued	SF-36 mental (short-term)	6 RCTs	1381	Direct	Imprecise	Moderate	Consistent	SMD 5.73 (-0.26 to 13.84)	3
	Sleep quality (short-term)	2 RCTs	363	Direct	Imprecise	Moderate	Consistent	SMD -0.11 (-0.39 to 0.14)	3
	Depression (short-term)	4 RCTs	524	Direct	Imprecise	Moderate	Consistent	SMD -0.18 (-0.37 to -0.01)	3
	Anxiety (short-term)	1 RCT	278	Direct	Imprecise	Moderate	Consistent	SMD -0.04 (-0.28 to 0.19)	Insufficient
	Discontinuation due to AEs	5 RCTs	782	Direct	Imprecise	Moderate	Consistent	RR 0.79 (0.50 to 1.27)	3
	Serious AEs	1 RCT	313	Direct	Imprecise	Moderate	Consistent	RR 0.58 (0.14 to 2.39)	Insufficient
	Nausea	5 RCTs	585	Direct	Imprecise	Moderate	Consistent	RR 0.98 (0.57 to 1.84)	3
	Vomiting	2 RCTs	339	Direct	Imprecise	Moderate	Consistent	RR 1.68 (0.34 to 8.19)	3
	Constipation	6 RCTs	860	Direct	Imprecise	Moderate	Consistent	RR 0.91 (0.67 to 1.13)	3
	Dizziness	5 RCTs	772	Direct	Imprecise	Moderate	Consistent	RR 1.22 (0.23 to 1.99)	3
	Headache	3 RCTs	457	Direct	Imprecise	Moderate	Consistent	RR 1.12 (0.46 to 2.25)	3
	Somnolence	6 RCTs	860	Direct	Imprecise	Moderate	Inconsistent	RR 0.72 (0.35 to 1.33)	3
	Pruritus	2 RCTs	190	Direct	Imprecise	Moderate	Consistent	RR 0.25 (0.03 to 1.91)	3
	Opioid abuse, dependence, or addiction	No studies	--	--	--	--	--	--	--
	Overdose	No studies	--	--	--	--	--	--	--
	All-cause mortality	No studies	--	--	--	--	--	--	--
	Fracture	No studies	--	--	--	--	--	--	--
	Cardiovascular events	No studies	--	--	--	--	--	--	--
	Endocrinological harms	No studies	--	--	--	--	--	--	--
Opioid + cannabis vs. opioid	Pain, function, opioid discontinuation, opioid dose	1 observational study	1514	Direct	Imprecise	Moderate	Unable to assess	No association	4

Opioid + benzodiazepine vs. opioid	Overdose	3 observational studies	140,002	Direct	Precise	Moderate	Consistent	Opioid + benzodiazepine associated with increased risk	3
Opioid + gabapentinoid vs. opioid	Overdose	3 observational studies	799,013	Direct	Precise	Moderate	Consistent	Opioid + gabapentinoid associated with increased risk	3
Methods for initiating and titrating opioids	Pain	2 RCTs	81	Direct	Imprecise	Moderate	Consistent	Unable to assess	Insufficient
Methods for initiating and titrating opioids, continued	Opioid use disorder or related outcomes	No studies	--	--	--	--	--	--	--
Short-acting vs. long-acting opioids	Pain, function	2 RCTs compared short- vs. long-acting of same opioid	184	Direct	Imprecise	Moderate	Consistent	No differences	3
	Overdose	1 cohort study	840,606	Direct	Precise	Moderate	Unknown	Long-acting associated with increased risk	3
Long-acting opioid vs. a different long-acting opioid	Pain, function, and other effectiveness outcomes	16 RCTs	7356	Direct	Precise	Moderate	Inconsistent	No patterns showing differential effectiveness, with some differences in opioid dosing between arms	2
	Overdose	4 cohort studies	193,166	Direct	Precise	Moderate	Inconsistent	Methadone associated with increased risk vs. morphine in 2 studies of Medicaid patients and decreased risk in 1 study of VA patients	4
Short + long-acting opioid vs. long-acting opioid alone	All	No studies	--	--	--	--	--	--	--

Scheduled, continuous vs. as-needed dosing	All	No studies	--	--	--	--	--	--	--
Opioid dose escalation vs. dose maintenance	Pain, function	1 RCT	140	Direct	Imprecise	Moderate	Unknown	No differences; doses were similar in the 2 arms	3
	Opioid withdrawal due to misuse	1 RCT	140	Direct	Imprecise	Moderate	Unknown	No difference	3
Opioid rotation vs. maintenance of current opioid therapy	All	No studies	--	--	--	--	--	--	--
Strategies for treating acute exacerbations of chronic pain	Pain (immediate)	4 RCTs	476	Direct	Precise	Low	Consistent	Buccal fentanyl more effective than placebo or oral opioid for immediate pain relief	2
	Longer-term outcomes, addiction, abuse	No studies	--	--	--	--	--	--	--
Tapering off opioids vs. continuation of opioids	Pain, function	1 RCT	34	Direct	Imprecise	Moderate	Unknown	No differences	3
	Opioid dose	1 RCT	34	Direct	Imprecise	Moderate	Unknown	Taper associated with lower dose	3
Tapering protocols and strategies	Pain, tapering completion, opioid withdrawal symptoms	1 RCT	21	Direct	Imprecise	Moderate	Unknown	Varenicline associated with no differences vs. placebo as an adjunct to tapering	3
	Opioid-related emergency department visit	1 cohort study	494	Direct	Imprecise	Moderate	Unknown	Each additional week to discontinuation associated with 7% reduction in risk	3
Opioid Risk Tool	Diagnostic accuracy	6 studies	1025	Direct	Precise	Moderate	Inconsistent	Sensitivity: 0.20 to 0.99 Specificity: 0.16 to 0.88	3

SOAPP Version 1	Diagnostic accuracy	2 studies	203	Direct	Imprecise	High	Consistent	Sensitivity: 0.68 and 0.73 Specificity: 0.38	3
SOAPP-R	Diagnostic accuracy	4 studies	840	Direct	Precise	Moderate	Inconsistent	Sensitivity: 0.25 to 0.53 Specificity: 0.62 to 0.77	3
Brief Risk Interview	Diagnostic accuracy	3 studies	577	Direct	Precise	High	Inconsistent	Sensitivity 0.73 to 0.83 Specificity: 0.43 to 0.88	3
Naloxone co-prescription	Emergency department visits	1 nonrandomized study	1985	Direct	Precise	Moderate	Unknown	Naloxone associated with decreased risk of emergency department visits vs. no naloxone	3
	All-cause mortality, opioid poisoning deaths	1 nonrandomized study	1985	Direct	Imprecise	Moderate	Unknown	No difference	3
Prescription opioid use disorder: Taper vs. maintenance	Drug use	1 RCT	113	Indirect	Precise	Moderate	Unknown	Buprenorphine taper inferior to maintenance	3
Prescription opioid use disorder: Buprenorphine vs. methadone	Drug use, pain function	1 RCT	54	Indirect	Imprecise	Moderate	Unknown	No differences	3

Abbreviations: AE=adverse events; CI=confidence interval; MD=mean difference; RCT=randomized controlled trial; RR=risk ratio; SMD=standard mean difference; SOE=strength of evidence; SOAPP= Screening and Opioid Assessment for Patients with Pain; SOAPP-R= Screening and Opioid Assessment for Patients with Pain-Revised Version; VA=Veterans Affairs Department; vs.=versus.

*Reporting bias was undetected for all key questions/outcomes, except where noted

†Mean differences for pain are reported on a 0 to 10 scale and for SF-36 measures are reported on a 0 to 100 scale

§Graded down for potential reporting bias

‖p for interaction by nonopioid type=0.03

¶Not downgraded for inconsistency because statistical heterogeneity was eliminated by exclusion of poor-quality trial, with similar pooled estimate

**Not downgraded for precision based on the pooled estimate after excluding a poor-quality trial

Noninvasive, nonpharmacological treatments for chronic pain. This table is based on Skelly AC, Chou R, Dettori JR, et al. AHRQ Comparative Effectiveness Reviews. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review Update. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020. The strength of evidence ratings in the AHRQ report were converted to ACIP-adapted GRADE evidence type ratings.

Low back pain.

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
Exercise	*Exercise vs. usual care, attention control, or a placebo intervention*	Function *Short-term*	10 (N=940)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.31 (95% CI –0.50 to –0.13); I^2=32% (excluding an outlier trial)[a]
		Function Intermediate-term	5 (N=616)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD –0.17 (95% CI –0.39 to 0.02); I^2=0%
		Function Long-term	1 (N=124)	Moderate	Unknown	Imprecise	Undetected	3	Difference 0.0 (95% CI –11.4 to 11.4) on the 0 to 100 ODI
		Pain *Short-term*	11 (N=981)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference –1.21 (95% CI –1.77 to –0.65) on a 0 to 10 scale; I^2=64%
		Pain *Intermediate-term*	5 (N=616)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference –0.85 (95% CI –1.67 to –0.07) on a 0 to 10 scale; I^2=50%
		Pain *Long-term*	1 (N=124)	Moderate	Unknown	Imprecise	Undetected	3	Difference –1.55 on a 0 to 10 scale (95% CI –2.76 to –0.34)
		Harms	2 (N=240)	Moderate	Consistent	Imprecise	Undetected	3	No evidence of increased risk of serious harms
Psychological Therapy	*Psychological therapy vs. usual care or attention control*	Function *Short-term*	3 (N=906)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.24 (95% CI –0.38 to –0.04); I^2=0%
		Function *Intermediate-term*	3 (N=1,026)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.24 (95% CI –0.38 to –0.10); I^2=0%
		Function *Long-term*	3 (N=815)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.28 (95% CI –0.43 to –0.13); I^2=0%
		Pain *Short-term*	3 (N=906)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.75 (95% CI –1.01 to –0.41) on a 0 to 10 scale; I^2=0%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
		Pain *Intermediate-term*	3 (N=1,026)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.71 (95% CI –0.97 to –0.46); I^2=0%
		Pain *Long-term*	3 (N=816)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.55 (95% CI –0.92 to –0.23); I^2=0%
	Psychological therapy vs. exercise	Function *Intermediate- and long-term*	1 (N=49)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from 1 poor-quality trial
		Pain *Intermediate- and long-term*	1 (N=49)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from 1 poor-quality trial
		Harms	1 (N=701)	Moderate	Unknown	Imprecise	Undetected	3	One trial reported no serious adverse events and withdrawal due to adverse events in <1% of patients randomized to psychological therapy
Physical Modalities	*Short-wave diathermy vs. sham diathermy*	Pain, function, harms	1 (N=68)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial
	Ultrasound vs. sham ultrasound	Function *Short-term*	2 (N=505)	Moderate	Inconsistent	Precise	Undetected	Insufficient	Inconsistent effects on function in two trials
		Pain *Short-term*	2 (N=505)	Moderate	Consistent	Precise	Undetected	3	No effects on pain in two trials
		Harms	1 (N=455)	Moderate	Consistent	Imprecise	Undetected	3	Any adverse event: RR 1.03 (95% CI 0.49 to 2.13) Serious adverse event: RR 0.48 (95% CI 0.12 to 1.88)
	Interferential therapy vs. placebo interferential therapy	Function *Short-term*	1 (N=150)	Moderate	Unknown	Unknown	Undetected	3	Difference 0.2 to 0.3 points (CI unclear)
		Pain *Short-term*	1 (N=150)	Moderate	Unknown	Unknown	Undetected	3	Difference 0.2 to 0.4 points (CI unclear)
		Harms	1 (N=150)	Moderate	Unknown		Undetected		Withdrawals due to adverse events: RR 1.0 (95% CI 0.14 to 6.8)
	Low-level laser therapy	Function *Short-term*	1 (N=56)	Moderate	Unknown	Precise	Undetected	3	Difference –8.2 (95% CI –13.6 to –2.8) on the 0 to 100 ODI

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
	vs. sham laser	Pain *Short-term*	1 (N=56)	Moderate	Unknown	Imprecise	Undetected	3	Difference –16.0 (95% CI –28.3 to –3.7) on a 0 to 100 scale
	Low-level laser therapy vs. exercise therapy	Function *Intermediate-term*	1 (N=35)	Moderate	Unknown	Imprecise	Undetected	3	Difference –4.4 (95% CI –11.4 to 2.5) on the ODI (0 to 100 scale)
		Pain *Intermediate-term*	1 (N=35)	High	Unknown	Imprecise	Undetected	3	Difference –0.9 (95% CI –2.5 to 0.7) on a 0 to 10 scale
		Harms	3 (N=162)	Moderate	Consistent	Imprecise	Undetected	3	No adverse events were reported
Manual Therapies	*Massage vs. sham massage, usual care, or attention control*	Function *Short-term*	6 (N=694)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.38 (95% CI –0.63 to –0.20); I^2=0%
		Function *Intermediate-term*	3 (N=676)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD –0.09 (95% CI –0.26 to 0.12); I^2=0%
		Pain *Short-term*	5 (N=644)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.55 (95% CI –0.88 to –0.23) on a 0 to 10 scale; I^2=0%
		Pain *Intermediate-term*	3 (N=680)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference –0.02 (95% CI –0.56 to 0.44); I^2=0%
	Massage vs. exercise	Function Intermediate-term	1 (N=144)	Moderate	Unknown	Imprecise	Undetected	3	Difference 1.2 (95% CI –1.47 to 3.87) on the 0 to 24 Roland Disability Questionnaire
		Pain *Intermediate-term*	1 (N=144)	Moderate	Unknown	Imprecise	Undetected	3	Difference 0.60 (95% CI –0.67 to 1.87) on the 0 to 10 Von Korff pain scale
	Massage vs. sham, usual care, attention control, or exercise	Harms	7 (N=906)	Moderate	Consistent	Imprecise	Undetected	3	Four trials reported no serious adverse events and one trial reported no adverse events; in four trials the proportion of massage patients with increased pain ranged from <1% to 26%
	Traction vs. sham traction	Function *Short-term*	2 (N=211)	Moderate	Consistent	Imprecise	Undetected	3	Differences 2 points on the ODI and 0.7 points on the Roland Disability Questionnaire, p>0.05 in both trials

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
		Pain *Short-term*	2 (N=211)	Moderate	Consistent	Imprecise	Undetected	3	Differences –4 points in one trial and 4 points in one trial, p>0.05 in both trials
		Harms	No studies	--	--	--	--	--	No evidence
	Spinal manipulation vs. sham manipulation, usual care, attention control, or placebo intervention	Function *Short-term*	3 (N=704)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.34 (95% CI –0.75 to –0.02); I^2=45%
		Function *Intermediate-term*	3 (N=1,000)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.40 (95% CI –0.85 to –0.05); I^2=65%
		Pain *Short-term*	3 (N=530)	High	Inconsistent	Imprecise	Undetected	3	Pooled difference –0.36 (95% CI –0.62 to 0.25) on a 0 to 10 scale; I^2=0%
		Pain *Intermediate-term*	3 (N=978)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.64 (95% CI –0.93 to –0.35); I^2=0%
	Spinal manipulation vs. exercise	Function *Short-term*	3 (N=640)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled SMD 0.02 (95% CI –0.28 to 0.30); I^2=37%
		Function *Intermediate-term*	4 (N=1,117)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled SMD 0.01 (95% CI –0.15 to 0.21); I^2=19%
		Pain *Short-term*	3 (N=636)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled difference 0.31 (95% CI –0.42 to 1.06) on a 0 to 10 scale; I^2=34%
		Pain *Intermediate-term*	4 (N=1,093)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference 0.23 (95% CI –0.14 to 0.59); I^2=0%
		Harms	7 (N=2,201)	Moderate	Consistent	Precise	Undetected	2	No serious adverse events or withdrawals due to adverse events in 7 trials. Nonserious adverse events (primarily increased pain) reported in 3 trials
Mindfulness Practices	*Mindfulness-based stress reduction vs. usual care or*	Function *Short-term*	4 (N=581)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled SMD –0.14 (95% CI –0.51 to 0.02); I^2=0%
		Function *Intermediate-term*	1 (N=229)	Moderate	Unknown	Imprecise	Undetected	3	SMD –0.20 (95% CI –0.46 to 0.06)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
	attention control	Function *Long-term*	1 (N=229)	Moderate	Unknown	Imprecise	Undetected	3	SMD –0.09 (95% CI –0.35 to 0.16)
		Pain *Short-term*	3 (N=546)	Moderate	Consistent[b]	Precise	Undetected	2	Pooled difference –0.68 (95% CI –1.29 to –0.28) on a 0 to 10 scale; I^2=45% (excluding 2 outlier trials)[b]
		Pain *Intermediate-term*	1 (N=229)	Moderate	Unknown	Precise	Undetected	3	Difference –0.75 (95% CI –1.16 to –0.34)
		Pain *Long-term*	1 (N=229)	Moderate	Unknown	Precise	Undetected	3	Difference –0.22 (95% CI –0.63 to 0.19)
		Harms	4 (N=577)	Moderate	Consistent	Imprecise	Undetected	3	One trial reported temporarily increased pain in 29% of patients undergoing MBSR and three trials reported no adverse events
Mind-Body Practices	*Yoga vs. attention control or wait list*	Function *Short-term*	8 (N=982)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.45 (95% CI –0.69 to –0.28); I^2=31%
		Function *Intermediate-term*	3 (N=540)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD –0.29 (95% CI –0.47 to –0.11); I^2=0%
		Pain *Short-term*	7 (N=710)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference –0.87 (95% CI –1.49 to –0.24) on a 0 to 10 scale; I^2=64%
		Pain *Intermediate-term*	2 (N=268)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –1.16 (95% CI –2.16 to –0.27); I^2=0%
	Yoga vs. exercise	Function *Short-term*	4 (N=559)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD –0.04 (95% CI –0.27 to 0.16); I^2=0%
		Function *Intermediate-term*	1 (N=246)	Moderate	Unknown	Imprecise	Undetected	3	SMD –0.01 (95% CI –0.26 to 0.24)
		Pain *Short-term*	5 (N=575)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled difference –0.63 (95% CI –1.68 to 0.245) on a 0 to 10 scale; I^2=88%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
		Pain *Intermediate-term*	1 (N=246)	Moderate	Unknown	Imprecise	Undetected	3	Difference 0.30 (95% CI –0.39 to 0.99)
		Harms	3 (N=616)	Moderate	Consistent	Imprecise	Undetected	3	No difference in risk of any adverse event (primarily mild back or joint pain); three serious adverse events in yoga patients were reported by one trial each: worsening back pain related to yoga, herniated disc, and cellulitis (≤1% of patients in each trial)
	Qi Gong vs. exercise therapy	Function *Short-term*	1 (N=125)	Moderate	Unknown	Imprecise	Undetected	3	Difference 0.9 (95% CI –0.1 to 2.0) on the 0 to 24 Roland Disability Questionnaire
		Function *Intermediate-term*	1 (N=125)	Moderate	Unknown	Precise	Undetected	3	Difference 1.2 (95% CI 0.1 to 2.3) on the Roland Disability Questionnaire
		Pain *Short-term*	1 (N=125)	Moderate	Unknown	Precise	Undetected	3	Difference 7.7 (95% CI 0.7 to 14.7) on a 0 to 100 scale
		Pain *Intermediate-term*	1 (N=125)	Moderate	Unknown	Imprecise	Undetected	3	Difference 7.1 (95% CI –1.0 to 15.2) on a 0 to 100 scale
		Harms	1 (N=125)	Moderate	Unknown	Imprecise	Undetected	3	No difference in risk of adverse events
Acupuncture	*Acupuncture vs. sham acupuncture, usual care, attention control, or a placebo intervention*	Function *Short-term*	4 (N=2,066)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.23 (95% CI –0.35 to –0.04); I^2=25%
		Function *Intermediate-term*	3 (N=997)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled SMD –0.08 (95% CI –0.42 to 0.28); I^2=64%
		Function *Long-term*	1 (N=218)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference –3.4 (95% CI –7.8 to 1.0) on the 0 to 100 ODI
		Pain *Short-term*	5 (N=2,109)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.54 (95% CI –0.91 to –0.16) on a 0 to 10 scale; I^2=25%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
		Pain *Intermediate-term*	5 (N=1,264)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference –0.22 (95% CI –0.67 to 0.21) on a 0 to 10 scale; I^2=0%
		Pain *Long-term*	1 (N=218)	Moderate	Unknown	Precise	Undetected	3	Difference –0.83 (95% CI –1.53 to –0.13) on a 0 to 10 scale
		Harms	6 (N=2,525)	Moderate	Consistent	Imprecise	Undetected	3	No evidence of increased risk of serious harms
Multi-disciplinary Rehabilitation	*Multi-disciplinary rehabilitation vs. usual care*	Function *Short-term*	4 (N=907)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.30 (95% CI –0.63 to 0.00); I^2=58%
		Function *Intermediate-term*	4 (N=481)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.37 (95% CI –0.69 to –0.08); I^2=34%
		Function *Long-term*	2 (N=286)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD –0.04 (95% CI –0.36 to 0.35); I^2=0%
		Pain *Short-term*	4 (N=907)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.53 (95% CI –0.86 to –0.11) on a 0 to 10 scale; I^2=0%
		Pain *Intermediate-term*	4 (N=481)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.62 (95% CI –1.06 to –0.18); I^2=0%
		Pain *Long-term*	2 (N=286)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference –0.35 (95% CI –1.10 to 0.34); I^2=0%
	Multi-disciplinary rehabilitation vs. exercise	Function *Short-term*	6 (N=379)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD –0.21 (95% CI –0.54 to 0.01); I^2=32%
		Function *Intermediate-term*	5 (N=415)	Moderate	Consistent[c]	Precise	Undetected	2	Pooled SMD –0.20 (95% CI –0.40 to –0.00); I^2=0% (excluding an outlier trial)[c]
		Function *Long-term*	2 (N=136)	Moderate	Consistent[c]	Imprecise	Undetected	3	Pooled SMD –0.07 (95% CI –0.50 to 0.39); I^2=0% (excluding an outlier trial)[c]
		Pain *Short-term*	6 (N=377)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.69 (95% CI –1.15 to –0.22) on a 0 to 10 scale; I^2=0%
		Pain *Intermediate-term*	5 (N=409)	Moderate	Consistent[c]	Precise	Undetected	2	Pooled difference –0.55 (95% CI –1.00 to –0.11); I^2=0% (excluding an outlier trial)[c]

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction, and Magnitude of Effect
		Pain *Long-term*	2 (N=136)	Moderate	Consistent[c]	Imprecise	Undetected	3	Pooled difference 0.00 (95% CI –1.31 to 1.17); I^2=0% (excluding an outlier trial)[c]
		Harms	2 (N=94)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient data on harms from 2 trials, though no serious harms were reported

a Outlier trial exclude, Areeudomwong, 2017
b Outlier trial excluded, Banth 2015
c Outlier trial excluded, Monticone 2013

Neck pain

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Exercise	*Exercise vs. attention control, no treatment or waitlist*	Function *Short-term*	3 (N=444)	Moderate	Inconsistent[a]	Imprecise	Undetected	3	Pooled SMD –0.22, 95% CI –0.66 to 0.17, I^2=72.6% [excluding outlier trial][a] Combination exercise only (2 trials), pooled SMD –0.44, 95% CI –0.76 to –0.09
		Function *Intermediate-term*	1 (N=230)	Moderate	Unknown	Precise	Undetected	3	SMD 0.14, 95% CI –0.12 to 0.40)
		Function *Long-term*	1 (N=125)	Moderate	Unknown	Imprecise	Undetected	3	SMD –0.39, 95% CI –0.74 to –0.03
		Pain *Short-term*	3 (N=444)	Moderate	Inconsistent[a]	Imprecise	Undetected	3	Pooled difference –0.70, 95% CI –1.62 to 0.15, I^2=63.7% [excluding outlier trial][a] Combination exercise only (2 two trials), pooled difference –1.12, 95% CI –1.82 to –0.43
		Pain *Intermediate-term*	2 (N=353)	Moderate	Consistent	Precise	Undetected	3	Pooled difference –0.25, 95% CI –0.81 to 0.31, I^2=0.0%
		Pain *Long-term*	3 (N=349)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference 0.07, 95% CI –0.51 to 0.88, I^2=0%
		Harms	2 (N=201)	High	Consistent	Imprecise	Undetected	3	No evidence of increased risk of serious harms
	Exercise vs. pharmaco-logical therapy	Function *Short-term*	1 (N=40) (vs. NSAIDs + muscle relaxants)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient data from 1 poor quality trial
		Function *Short-term*	1 (N=64) (vs. acetaminophen)	Moderate	Unknown	Imprecise	Undetected	3	Difference –5.6 (95% CI –8.36 to –2.83) on the 0 to 50 NDI scale

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Short-term*	1 (N=40) (vs. NSAIDs + muscle relaxants)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient data from 1 poor quality trial
		Pain *Short-term*	1 (N=64) (vs. acetaminophen)	Moderate	Unknown	Imprecise	Undetected	3	Difference –3.11 (95% CI –4.17 to –2.05) on the 0 to 10 NPS
		Pain, Function, *Intermediate-term*	1 (N=40) (vs. NSAIDs + muscle relaxants)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient data from 1 poor quality trial
		Harms	1 (N=64) (vs. acetaminophen)	Moderate	Unknown	Imprecise	Undetected	3[c]	One trial reported no adverse events
Psychological Therapies	*Relaxation training vs. no intervention*	Function *Short-term*	1 (N=258)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference 0.1 (95% CI –2.9 to 3.2) on 0-80 scale
		Function *Intermediate-term*	1 (N=258)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference 0.2 (95% CI –2.8 to 3.1) on 0-80 scale
		Pain *Short-term*	1 (N=258)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference 0.2 (95% CI –0.4 to 0.8) on 0-10 scale
		Pain *Intermediate-term*	1 (N=258)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference 0.2 (95% CI –0.3 to 0.8) on 0-10 scale
	Relaxation training vs. exercise	Function *Short-term*	1 (N=263)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference 0.2 (95% CI –2.8 to 3.2) on 0-80 scale
		Function *Intermediate-term*	1 (N=263)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference 0.2 (95% CI –2.7 to 3.2) on 0-80 scale
		Pain *Short-term*	1 (N=263)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference –0.2 (95% CI –0.8 to 0.4) on 0-10 scale
		Pain *Intermediate-term*	1 (N=263)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference –0.2 (95% CI –0.8 to 0.3) on 0-10 scale
	Relaxation training vs. no intervention or exercise	Harms	None	–	–	–	–	–	No evidence

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Physical Modalities	*Traction vs. attention control*	Function, Pain, Harms *Short-term*	1 (N=79)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial.
	Laser vs. sham intervention	Function *Short-term*	2 (N=144)	Low	Consistent	Imprecise	Undetected	2	Pooled difference –13.60 (95% CI –26.30 to –6.30) on a 0-100 scale: I^2=0%
		Pain *Short-term*	3 (N=192)	Low	Consistent	Imprecise	Undetected	2	Pooled difference –1.89, (95% CI –3.34 to –0.06) on a 0-10 scale: I^2=61%
		Harms	1 (N=90)	Low	Unknown	Imprecise	Undetected	3	Adverse effects occurred with similar frequency in both groups. The most frequently reported adverse effects in the intervention group included mild (78%) or moderate (60%) increased neck pain, increased pain elsewhere (78%), mild headache (60%) and tiredness (24%).
	Electro-magnetic fields vs. sham intervention	Function, Pain, Harms *Short-term*	1 (N=81)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial.
Manual Therapies	*Massage vs. attention or waitlist control*	Function *Short-term*	2 (N=148)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference –3.66, (95% CI –6.58 to –0.56) on a 0-50 NDI scale: I^2=10% 1 trial (Sherman): Success (≥5 points), 39% vs. 17%; RR 2.7 (95% CI 0.99 to 7.5)
		Function *Intermediate-term*	1 (N=58)	Moderate	Unknown	Imprecise	Undetected	3	Success (≥5 points): 57% vs. 31%, RR 1.8 (95% CI 0.97 to 3.5)
		Pain *Short-term*	1 (N=92)	Moderate	Unknown	Imprecise	Undetected	3	Difference –1.8 (95% CI –2.7 to –0.9) on a 0-10 scale

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	Massage vs. exercise	Pain *Intermediate-term*	1 (N=85)	Moderate	Unknown	Imprecise	Undetected	3	Difference 0.2 (95% CI –0.82 to 1.22) on the 0-10 NRS
	Massage vs. attention control or vs. exercise	Harms	2 (N=143)	Moderate	Unknown	Imprecise	Undetected	3	No evidence of increased risk of serious harms
Mind-body Practices	*Alexander Technique plus usual care vs. usual care alone*	Function *Short-term*	1 (N=344)	Moderate	Unknown	Precise	Undetected	3	Difference –5.56 (95% CI –8.33 to –2.78) on 0-100% scale
		Function *Intermediate-term*	1 (N=344)	Moderate	Unknown	Precise	Undetected	3	Difference –3.92 (95% CI –6.87 to –0.97) on 0-100% scale
		Harms	1 (N=344)	Moderate	Unknown	Imprecise	Undetected	3	No clear difference in the risk of any non-serious adverse event (e.g., pain and incapacity, knee injury, muscle spasm, and complications after surgery): RR 2.25 (95% CI 1.00 to 5.04) No serious treatment-related adverse events reported.
	Basic body awareness therapy vs. exercise	Function *Short-term*	1 (N=113)	Moderate	Unknown	Imprecise	Undetected	3	Difference between groups in mean change from baseline –1, p>0.05
		Function *Intermediate- and long-term*	1 (N=139)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial
		Pain *Intermediate- and long-term*	1 (N=139)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial
		Harms	1 (N=113)	Moderate	Unknown	Imprecise	Undetected	3	No serious adverse effects Any non-serious adverse effects: RR 0.65 (95% CI 0.37 to 1.14)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Acupuncture	Acupuncture vs. sham, placebo or usual care	Function *Short-term*	5 (N=919)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.40 (95% CI –0.67 to –0.14); I^2=61%
		Function *Intermediate-term*	3 (N=563)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD –0.19 (95% CI –0.37 to 0.05); I^2=0%
		Function *Long- term*	1 (N=107)	Moderate	Unknown	Imprecise	Undetected	3	Difference –1.8 (95% CI –4.84 to 1.24) on a 0-50 scale
		Pain *Short-term*	4 (N=490)	Moderate	Inconsistent[b]	Precise	Undetected	3	Pooled difference –0.27 (95% CI –0.59 to 0.05) on a 0-10 scale; I^2=2% [excluding outlier trial][b]
		Pain *Intermediate-term*	3 (N=354)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled difference 0.40 (95% CI –0.45 to 1.44) on a 0-10 scale; I^2=19%
		Pain *Long- term*	1 (N=107)	Moderate	Unknown	Imprecise	Undetected	3	Pooled difference –0.35 (95% CI –1.34 to 0.64) on a 0-10 scale
	Acupuncture vs. pharmaco-logical care	Function *Short-term*	1 (N=30)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence due to study limitations, unknown consistency and imprecision from one poor-quality study
		Pain *Short-term*	2 (N=53)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence due to study limitations and imprecision from 2 poor quality studies
	Acupuncture vs. sham, placebo, usual care or pharmaco-logical care	Harms	6 (N=937)	Moderate	Consistent	Precise	Undetected	2	No serious treatment-related adverse events reported. Most common non-serious adverse effects included numbness/ discomfort, fainting and bruising.

CI = confidence interval; RCT = randomized controlled trial; RR = risk ratio; SMD = standardized mean difference.

a Outlier trial excluded, Li 2017b. Heterogeneity is explained in part by the contribution of the good quality study; the others are fair quality.
b Outlier trial excluded, Ho 2017.

Knee osteoarthritis

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Exercise	*Exercise vs. usual care, attention control, or no intervention*	Function *Short-term*	8 (N=748)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD −0.29, 95% CI −0.46 to −0.11, I^2=9.9%
		Function *Intermediate-term*	11 (N=879)	Moderate	Inconsistent[a]	Imprecise	Undetected	3	Pooled SMD −0.63, 95% CI −1.17 to −0.10, I^2=90.8% [excluding outlier trial][a]
		Function *Long-term*	4 (N=1,199)	High	Consistent	Precise	Undetected	3	Pooled SMD −0.22, 95% CI −0.34 to −0.08, I^2=0%
		Pain *Short-term*	8 (N=748)	Moderate	Consistent	Precise	Undetected	2	Pooled difference on 0-10 scale: −0.47, 95% CI −0.86 to −0.10, I^2=41.7% One fair-quality trial (Bennell 2005) found no statistical difference between exercise and sham in proportion with clinically relevant reductions (≥1.75 points) in: VAS pain on movement: 58% (34/59) vs. 42% (27/65); RR 1.4, 95% CI 1.0 to 2.0; VAS global improvement in pain: 59% (35/59) vs. 50% (33/65); RR 1.2, 95% CI 0.8 to 1.6
		Pain *Intermediate-term*	11 (N=880)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled difference on a 0-10 scale: −1.34, 95% CI −2.12 to −0.54, I^2=90%
		Pain *Long-term*	4 (N=1,200)	High	Consistent	Precise	Undetected	3	Pooled difference on a 0-10 scale: −0.30, 95% CI −0.49 to −0.00, I^2=0%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	Exercise vs. pharmacologic therapy (acetaminophen and NSAIDs)	Function, *Intermediate-term*	1 (N=93)	Moderate	Unknown	Imprecise	Undetected	3[c]	No differences between groups on any measure. Proportion achieving a clinically meaningful improvement (>10 points on KOOS ADL): 47% (22/47) versus 28% (13/46); RR 1.7, 95% CI 1.0 to 2.9 KOOS ADL (0-100): difference −3.6, 95% CI −9.2 to 2.1 KOOS Sport and Recreation (0-100): difference −2.9, 95% CI −11.4 to 5.5
		Pain *Intermediate-term*	1 (N=93)	Moderate	Unknown	Precise	Undetected	3[c]	KOOS Pain (0-100): difference 4.2, 95% CI −10.0 to 1.6

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	Exercise vs. usual care, attention control, no intervention, or pharmacologic therapy	Harms	8 (N=1097)	Moderate	Consistent	Precise	Undetected	2	One RCT in older patients reported six serious adverse events, with no significant difference between groups: five in the exercise group [four falls (1 resulting in distal radius fracture), one foot fracture from dropping a dumbbell] vs. one instance of sudden death in a control participant; 1.7% (5/290) vs. 0.7% (1/149), RR 2.57 (95% CI 0.30 to 21.79) One trial reported greater temporary, minor increases in pain in the exercise group versus a sham group; however, four trials found no difference in worsening of pain symptoms with exercise vs. comparators. No difference in adverse events was reported on the one new trial of exercise compared to standard analgesics and anti-inflammatory therapy.
Psychological Therapies	*CBT/MI/pain coping skills training vs. usual care*	Function, Pain *Short-term to long-term*	2 (N=222)	Moderate	Consistent	Imprecise	Undetected	3	No differences in one fair quality trial of CBT and one poor quality trial of pain coping skills training averaged over 6 to 12 months (intermediate to long term) and 1.5 to 10.5 months (short to intermediate term).
		Function, *Short-term*	2 (N=210)	Moderate	Consistent	Imprecise	Undetected	3	Pooled SMD on a 0-68 scale −2.09, 95% CI −8.70 to 1.61, I^2=63.3%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Short-term*	2 (N=210)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference on a 0-20 scale: −0.60, 95% CI −1.48 to −0.08, I^2 = 0.0%
		Harms	4 (N=371)	Moderate	Consistent	Imprecise	Undetected	3	No adverse events observed across four trials (3 fair quality and 1 poor quality).
	Pain coping skills training vs. exercise	Function *Short-term and intermediate term*	1 (N=149)	Moderate	Unknown	Imprecise	Undetected	3	No difference in WOMAC physical 0-68 Short-term: difference 2.0 (95% CI −2.4 to 6.4), p=0.37 Intermediate-term: MD 3.2 (95% CI −0.6 to 7.0), p=0.10
		Pain *Short-term and intermediate term*	1 (N=149)	Moderate	Unknown	Imprecise	Undetected	3	No difference in WOMAC pain 0-20) Short-term: difference −0.1 (95% CI −1.2 to 1.0) Intermediate-term: difference 0.4 (95% CI −0.8 to 1.6), p=0.49)
		Harms	1 (N=149)	Moderate	Unknown	Imprecise	Undetected	3	Knee pain was more common in the exercise group during treatment (31% versus 3%) and during short and intermediate term followup (10% versus 7%) as was overall body pain (15% versus 2%)
Physical Modalities	*Ultrasound vs. sham*	Function, *Short-term*	3 (N=249)	Moderate	Unknown	Imprecise	Undetected	3	Continuous and pulsed ultrasound vs. sham, difference −2.50, 95% CI −6.37 to 1.22, I^2=94.0%
		Function *Intermediate-term*	1 (N=60)	Moderate	Unknown	Imprecise	Undetected	3	Continuous and pulsed ultrasound vs. sham, 0-68 scale, differences: −2.9 (95% CI −9.19 to 3.39) and 1.6 (95% CI −3.01 to 6.22)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Short-term*	3 (N=249)	Moderate	Unknown	Imprecise	Undetected	3	Continuous and pulsed ultrasound vs. sham, 0-10 scale, pooled difference −1.20, 95% CI −3.71 to 1.31, I^2=91.1%
		Pain *Intermediate-term*	1 (N=60)	Moderate	Unknown	Imprecise	Undetected	3	Continuous and pulsed ultrasound vs. sham, 0-20 scale, differences: −1.6 (95% CI −3.26 to 0.06) vs. 0.2 (95% CI −1.34 to 1.74); also no difference between groups for other pain measures.
		Harms	4 (N=318)	Moderate	Unknown	Imprecise	Undetected	3	No adverse events reported during the four trials (1 good, 2 fair, and 1 poor quality)
	TENS vs. sham	Function *Intermediate-term*	1 (N=70)	Low	Unknown	Imprecise	Undetected	3	Proportion of patients who achieved MCID (≥9.1) in WOMAC function: 38% vs 39%, RR 1.2 (95% CI 0.6 to 2.2); Difference in mean change −1.9 (95% CI −9.7 to 5.9) on a 0-100 scale
		Pain *Intermediate-term*	1 (N=70)	Low	Unknown	Imprecise	Undetected	3	Proportion of patients who achieved MCID (≥20) in pain VAS: 56% vs 44%, RR 1.3 (95% CI 0.8 to 2.0) Difference in mean change 0.9 (95% CI −11.7 to 13.4) on 0-100 VAS and −5.6 (95% CI −14.9 to 3.6) on 0-100 WOMAC pain scale.

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Harms	1 (N=70)	Low	Unknown	Imprecise	Undetected	3	No evidence of increased risk of serious harms; no differences between treatments for harms (RR 1.06, 95% CI 0.38 to 2.97)
	Low-level laser therapy vs. sham laser	Function *Short-term*	1 (N=49)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one small fair quality trial
		Function *Intermediate-term*	2 (N=109)	High	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient evidence from one small fair trial and one poor quality trial
		Pain *Short-term*	2 (N=76)	High	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two small trials, one fair trial and one poor quality
		Pain *Intermediate-term*	2 (N=109)	High	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient evidence from one small fair trial and one poor quality trial
		Harms	2 (N=109)	High	Consistent	Imprecise	Undetected	Insufficient	Data for harms was insufficient. No adverse events were reported.
	Microwave diathermy vs. sham	Function *Short-term*	1 (N=63)	Moderate	Unknown	Imprecise	Undetected	Insufficient	There was insufficient evidence to determine short-term effects or harms from one small trial microwave diathermy
		Pain *Short-term*	1 (N=63)	Moderate	Unknown	Imprecise	Undetected	Insufficient	There was insufficient evidence to determine short-term effects or harms from one small trial microwave diathermy; substantial imprecision noted
		Harms	1 (N=63)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Data for harms were insufficient. However, no serious adverse events occurred in either group. Two patients in the diathermy group reported transient aggravation of symptoms.

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	Pulsed Short-wave Diathermy vs. Sham	Function *Short-term*	1 (N=115)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Function *Long-term*	1 (N=86)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Pain *Short-term*	1 (N=115)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Pain *Long-term*	1 (N=86)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Harms	2 (N=201)	High	Unknown	Imprecise	Undetected	Insufficient	Data were insufficient for harms. No adverse events were reported by either trial.
	Electro-magnetic fields vs. sham	Function *Short-term*	2 (N=180)	Moderate	Consistent	Imprecise	Undetected	3	The fair quality trial: (WOMAC) activities of daily living subscale (0-85) mean difference −3.48 (95% CI −4.44 to −2.51)
		Pain *Short-term*	2 (N=180)	Moderate	Consistent	Imprecise	Undetected	3	The fair quality trial: WOMAC-pain subscale (0-25) versus sham, −0.84 (95% CI −1.10 to −0.58)
		Harms	1 (N=90)	Moderate	Unknown	Imprecise	Undetected	3	More patients who received real versus sham electromagnetic field therapy reported throbbing or warming sensations or aggravation of pain; however the difference was not significant (RR 1.95, 95% CI 0.81 to 4.71)
	Superficial heat vs. placebo	Pain *Short-term*	1 (N=52)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one small, fair-quality trial
		Harms	1 (N=52)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Data was insufficient for harms; no adverse events were reported

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	Brace vs. usual care	Function, Pain, Harms *Intermediate- and long-term*	1 (N=118)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
Manual Therapies	*Manipulation vs. usual care*	Function, Harms *Intermediate-term*	1 (N=58 knee OA)	Moderate	Unknown	Unknown	Undetected	Insufficient	Insufficient evidence from one fair-quality trial; inadequate data to determine effect sizes or statistical significance
	Manipulation vs. exercise	Function, Harms *Intermediate-term*	1 (N=59 knee OA)	Moderate	Unknown	Unknown	Undetected	Insufficient	Insufficient evidence from one fair-quality trial; inadequate data to determine effect sizes or statistical significance
	Massage vs. usual care	Function, Pain, Harms *Short-term*	1 (N=125)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one fair-quality trial.
Mind-body Practices	*Tai Chi vs. attention control*	Function *Short-term*	2 (N=81)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, unblinded trials; (one fair, one poor quality)
		Function *Intermediate-term*	1 (N=40)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, unblinded trials (one fair, one poor quality)
		Pain *Short-term*	2 (N=81)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, unblinded trials (one fair, one poor quality)
		Pain *Intermediate term*	1 (N=40)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, unblinded trials (one fair, one poor quality)
		Harms	2 (N=81)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, unblinded trials(one fair, one poor quality)
Acupuncture	*Acupuncture vs. usual care, no*	Function *Short-term*	4 (N=871)	Moderate	Inconsistent[b]	Precise	Undetected	3	Pooled SMD −0.05, 95% CI −0.32 to 0.38) [Excluding outlier][b]

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	treatment, waitlist, or sham	Function *Intermediate-term*	4 (N=767)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD[c] −0.15, 95% CI −0.31 to 0.02, I^2=0%
		Pain *Short-term*	6 (N=1065)	Moderate	Inconsistent	Precise	Undetected	3	Pooled SMD −0.27, 95% CI −0.67 to 0.12, I^2=79.3%
		Pain *Intermediate term*	4 (N=767)	Moderate	Consistent	Precise	Undetected	2	Pooled SMD −0.16, 95% CI −0.32 to −0.01, I^2=0%); Individually no trial reached statistical significance.
		Harms	9 (N=1796)	Moderate	Consistent	Imprecise	Undetected	2	There is no apparent difference in risk of serious adverse events between any form of acupuncture and the control group. Worsening of symptoms (7%-14%), mild bruising, swelling or pain at the acupuncture site (1%-18%) were most common; One case of infection at an electroacupuncture site was reported.
	Acupuncture vs. exercise	Function, Pain, Harms *Short-term*	1 (N =120)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor-quality trial.

CI = confidence interval; OA: osteoarthritis; MCID = minimal clinically important difference; MI = motivation interveiewing; NSAIDs = non-steroidal anti-inflammatory drugs; RCT = randomized controlled trial; RR = risk ratio; SMD = standardized mean difference; TENS = transcutaneous electrical stimulation; VAS = visual analog scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

a Outlier excluded, Dias 2003.

b Outlier excluded, Berman 1999.

c Results for all trials individually were not statistically significant.

Hip osteoarthritis

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Exercise	*Exercise vs. usual care*	Function *Short-term*	3 (N=377)	Moderate	Consistent	Precise	Undetected	3	Pooled SMD −0.33, 95% CI −0.58 to −0.11, I^2=0%
		Function *Intermediate-term*	2 (N=307)	Moderate	Consistent	Precise	Undetected	3	Pooled SMD −0.28, 95% CI −0.55 to 0.02, I^2=0%
		Function *Long-term*	1 (N=118)	Moderate	Unknown	Imprecise	Undetected	Insufficient	SMD −0.37, 95% CI −0.74 to −0.01
		Pain *Short-term*	3 (N=371)	Moderate	Inconsistent	Imprecise	Undetected	3	Pooled SMD −0.30, 95% CI −0.70 to −0.02, I^2=0%
		Pain *Intermediate-term*	2 (N=307)	Low	Consistent	Imprecise	Undetected	3	Pooled SMD −0.14, 95% CI −0.40 to 0.12, I^2=0%
		Pain *Long-term*	1 (N=118)	Moderate	Unknown	Imprecise	Undetected	Insufficient	SMD −0.25, 95% CI −0.62 to 0.11
		Harms	2 (N=170)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient data from two trials although no serious harms were reported in two trials.
Manual Therapies	*Manipulation vs. usual care*	Function *Intermediate-term*	1 (N=47)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one fair quality trial. No effect size could be calculated.
		Harms	1 (N=47)	Moderate	Unknown	Imprecise	Undetected	Insufficient	No treatment-related serious adverse events were detected
	Manipulation vs. exercise	Function *Short-term*	1 (N=109)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted difference 11.1 (95% CI 4.0 to 18.6) on 0-100 scale
		Function *Intermediate-term*	2 (N=155)	Moderate	Consistent	Imprecise	Undetected	3	Adjusted difference 9.7, 95% CI, 1.5 to 17.9 on 0-100 scale; no effect size could be calculated in the other trial but direction of effect was similar
		Pain *Short-term*	1 (N=109)	Moderate	Unknown	Imprecise	Undetected	3	Adjusted differences −0.72 (95% CI −1.38 to −0.05) for pain at rest and −1.21 (95% CI −2.29 to −0.25) for pain walking on 0-10 scale

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Intermediate-term*	1 (N=109)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Adjusted differences −0.70 (95% CI −2.03 to 0.59) for pain at rest and −1.27 (95% CI −2.40 to −0.19) for pain walking on 0-10 scale; impact on pain is unclear from different measures
		Harms	2 (N=155)	Moderate	Consistent	Imprecise	Undetected	3	No treatment-related serious adverse events were detected in one trial; similar rates of study withdrawal due to symptom aggravation were seen in the second trial (5% vs. 4%; RR 1.42, 95% CI 0.25 to 8.16)

CI = confidence interval; RCT = randomized controlled trial; RR = risk ratio; SMD = standardized mean difference.

Hand osteoarthritis

Intervention	Comparator	Outcome	N RCTs (patients)	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Exercise	*Exercise vs. usual care*	Function, Pain, Harms *Short-term*	1 (N=130)	High	Unknown	Imprecise	Undetected	Insufficient	Poor quality trial of exercise vs waitlist; high attrition rate in exercise arm (29%). No serious adverse events; increased pain (hand or neck/shoulders) in eight patients (6%), not reported by group.
Physical Modalities	*Low level laser therapy vs. sham intervention*	Function *Short-term*	1 (N=88)	Low	Unknown	Imprecise	Undetected	3	No differences observed in one good quality trial (difference 0.2, 95% CI −0.2 to 0.6).
		Pain *Short-term*	1 (N=88)	Low	Unknown	Imprecise	Undetected	3	No differences observed in one good quality trial (difference 0.1, 95% CI −0.3 to 0.5).
		Harms	1 (N=88)	Low	Unknown	Imprecise	Undetected	3	No serious adverse events identified in one good quality trial.
	Superficial heat (paraffin) vs. no treatment	Function, Pain, Harms *Short-term*	1 (N=56)	Moderate	Unknown	Imprecise	Possible	Insufficient	Insufficient evidence from one small trial
Multidisci-plinary Rehabilitation	*Multidisci-plinary rehabilitation vs. waitlist*	Function *Short-term*	1 (N=151)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference 0.49 (95% CI −0.09 to 0.37); OASRI-OMERACT Responder: OR 0.82 (95% CI 0.42 to 1.61)
		Pain *Short-term*	1 (N=151)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference 0.40 (95% CI −0.5 to 1.3)
		Harms	1 (N=151)	Moderate	Unknown	Imprecise	Undetected	Insufficient	No serious adverse events identified.

CI = confidence interval; OASRI-OMERACT = Osteoarthritis Research Society International-Outcome Measures in Rheumatology; OR = odds ratio; RCT = randomized controlled trial.

Fibromyalgia

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Exercise	*Exercise vs. usual care, attention control, or a placebo intervention*	Function *Short-term*	7 (N=410)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference, −7.68 on a 0 to 100 scale, 95% CI, −13.04 to −1.84, I^2=59.9%
		Function *Intermediate-term*	8 (N=461)	Moderate	Consistent	Precise	Undetected	2	Pooled difference on 0-100 scale, −6.04 95% CI −9.25 to −3.01, I^2=0%
		Function *Long-term*	3 (N=178)	Moderate	Consistent	Imprecise	Undetected	3	Pooled difference, on 0-100 scale, −4.33, 95% CI −10.46 to 1.97, I^2=0%)
		Pain *Short-term*	6 (N=337)	Moderate	Consistent[a]	Imprecise	Undetected	2	Pooled difference −0.88, 95% CI −1.33 to −0.27, I^2=1.5%; (Excluding outlier)[a]
		Pain *Intermediate-term*	8 (N=382)	Moderate	Consistent	Precise	Undetected	2	Pooled difference −0.51, 95% CI −0.92 to −0.06, I^2=0%
		Pain *Long-term*	4 (N=241)	Moderate	Consistent	Precise	Undetected	2	Pooled difference –0.18, 95% CI −0.77 to 0.42, I^2=0%
		Harms	3 (N=132)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Insufficient data on harms. Most trials of exercise did not report on adverse events at all. One trial reported one non-study-related adverse event. Two trials reported no adverse events.
	Exercise vs. pharma-cologial therapy	Pain *Intermediate-term*	1 (n=32)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small, poor-quality trial

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Psychological Therapies	*Psycho-logical therapy vs. usual care, waitlist, or attention control*	Function *Short-term*	Any therapy: 5 (N=258) CBT: 2 (N=96) CBT/ACT: 1 (n=169) EMG Biofeedback: 1 (n=59) Imagery: 1 (n=70)	Moderate	Inconsistent	Imprecise	Undetected	3 (CBT) Insufficient (biofeed-back, imagery)	FIQ total score 0-100 scale **Any therapy** Pooled mean difference −2.82 (95% CI −9.79 to 2.81, I^2=70.6% **CBT only:** More CBT recipients with clinically important improvement, 2 trials, RR 2.2 (0.5 to 9.3) and RR 2.8 (1.3 to 6.1) Pooled mean difference (3 trials [1 new] −6.14, 95% CI −16.86 to 3.74 **Other therapies**: No clear difference for guided imagery (1 poor quality trial) or EMG biofeedback (1 poor quality trial, 1 small fair quality trial)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Function *Intermediate-term*	CBT: 3 (N=280) EMG Biofeedback: 1 (n=85)	Moderate	Inconsistent	Imprecise	Undetected	3 (CBT) Insufficient (biofeed-back)	CBT: Pooled difference on FIQ Total (0-100): −12.81, 95%CI −24.07 to −2.33, I^2 = 94.2%) Difference on FIQ Physical Function Scale (0-10) (1 trial, Thieme): −1.8, 95% CI −2.9 to −0.70 More CBT recipients with a clinically important improvement RR 2.9 (95% CI 1.4 to 6.3) in one trial (Castel) New trials: No difference between CBT and waitlist on Pain Disability Index (McCrae) or West Haven - Yale Multidimensional Pain Inventory (MPI) pain interference subscale (Karlsson) Trial of biofeedback vs. usual care: unclear difference, mean changes −1.6 (95% CI −3.4 to 0.2) versus −0.6 (95% CI −2.9 to 1.7)
		Function *Long-term*	CBT: 2 (N=227) EMG Biofeedback: 1 (n=59)	High	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient data from three poor quality trials
		Pain *Short-term*	CBT: 4 (N=197) EMG Biofeedback: 1 (n=53)	High	Consistent	Precise	Undetected	3 (CBT) Insufficient (biofeed-back)	CBT: Pooled mean difference −0.62, 95% CI −1.08 to −0.14, 0-10 scale No clear difference for EMG biofeedback (1 poor quality trial)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Intermediate-term*	CBT/ACT: 6 (N=551) EMG Biofeedback: 1 (n=65)	Moderate	Consistent	Imprecise	Undetected	2 (CBT) Insufficient (biofeed-back)	CBT: Pooled mean difference −0.55, 95% CI −1.13 to −0.06,, 0-10 scale Mean difference −1.11, 95% CI −2.06 to −0.16 for EMG biofeedback (1 poor quality trial)
		Pain *Long-term*	CBT: 1 (n=40) EMG Biofeedback: 1 (n=53)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient data from two poor quality trials
		Harms	5 (N=482)	High	Unknown	Imprecise	Undetected	Insufficient	Data were insufficient; withdrawals due to adverse events were reported by three trials: 0% vs. 3.6% (2 cases) and 7% vs. 8% (1 in each group for pain during testing) for CBT vs. usual care, respectively, in two trials (1 fair, poor quality), and in 5% (2 cases of depression) vs. 50% (worsening of symptoms in 20 patients) for CBT vs. attention control in one poor quality trial. Two (1 new) fair quality trials reported no adverse events for CBT.
	Psychological therapy vs. pharma-cological therapy	Function *Short-term*	CBT plus amitriptyline vs. amitriptyline 1 (n=51) EEG Biofeedback vs. escitalopram, 1 (n=36)	Moderate	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient data from one fair and one poor quality trial

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Function *Intermediate-term*	CBT or ACT: 2 (N=212) vs. pregabalin (plus duloxetine for depressed patients)	Moderate	Inconsistent	Imprecise	Undetected	3	FIQ 0-100 scale Pooled difference −9.81, 95%CI −21.2 to 1.58, I^2=96% Improvement in function reported for both trials of CBT versus pregabalin (plus duloxetine as needed) (small improvement in one trial, difference −4.0, 95% CI −7.4 to −0.56; moderate improvement in the second trial, difference −15.6, 95% CI −19.0 to −12.2). Different magntitude of effects resulted in substantial heterogeneity.
		Pain *Short-term*	CBT: 1 (n=51) EEG Biofeedback 1 (n=36)	Moderate	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient data from one fair and one poor quality trial
		Pain *Intermediate-term*	CBT or ACT: 2 (N=212) vs. pregabalin (plus duloxetine for depressed patients)	Moderate	Inconsistent	Precise	Undetected	3	VAS 0-10 scale, pooled difference, −0.31, 95% CI −1.15 to 0.51, I^2= 63.5%)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Harms	CBT or ACT: 2 (N=216)	Moderate	Unknown	Imprecise	Undetected	3	Withdrawals due to adverse events, CBT vs. pregabalin: 0% vs. 5.5%; events included two digestive problems, and one dizziness in one trial. In the second (new) trial, for ACT vs. pregabalin, withdrawals due to lack of efficacy (5.9% vs. 1.9, respectively) or patients decision (3.9% vs. 0%, respectively); adverse events reported in the pregabalin group only included nausea (25%), dry mouth (23%), drowsiness, headache and fatigue (21% each) and constipation (19%).
	Psychological therapy vs. exercise	Function *Short-term*	1 (n=51)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one small, poor quality trial
		Function *Intermediate-term*	CBT: 1 (n=40) EMG Biofeedback: 1 (n=114)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two poor quality trials
		Function Long-term	CBT: 2 (N=40) Relaxation: 1 (n=130) EMG Biofeedback: 1 (n=51)	High	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient evidence from three poor quality trials; inconsistency in findings noted.
		Pain *Short-term*	EMG Biofeedback: 1 (n=51)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one small, poor quality trial
		Pain *Intermediate-term*	CBT: 1 (n=40) EMG Biofeedback: 1 (n=114)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence two poor quality trials

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Pain *Long-term*	CBT: 2 (N=80) Relaxation 1 (n=130) EMG Biofeedback 1 (n=51)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from four poor quality trials
		Harms	2 (N=170)	High	Consistent	Imprecise	Undetected	Insufficient	Data were insufficient for harms. In one trial no patient had an adverse event in relaxation group compared to five (7.5%) in the strengthening exercise group (increased pain, three of which withdrew). In the other trial, withdrawals due to adverse events were similar between groups and none of the events were related to treatment.
Physical Modalities	*Magnetic fields vs. usual care or sham*	Function and Pain *Short-term*	1 (n=33)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Function *Intermediate-term*	1 (n=119)	Moderate	Unknown	Imprecise	Undetected	3	Difference −5.0 (95% CI −14.1 to 4.1) vs. sham and −5.5 (95% CI −14.4 to 3.4) vs. usual care on the 0-80 scale FIQ
		Pain *Intermediate-term*	1 (n=119)	Moderate	Unknown	Imprecise	Undetected	3	Difference −0.6 (95% CI −1.9 to 0.7) vs. sham and −1.0 (95% CI −2.2 to 0.2) vs. usual care on a 0-10 NRS
		Harms	1 (n=119)	Moderate	Unknown	Imprecise	Undetected	3	No differences in adverse events between the functional and sham magnetic groups (data not reported); none of the events were deemed to be related to the treatments
Manual Therapies	*Massage/ myofascial*	Function *Intermediate-term*	1 (n=94)	Moderate	Unknown	Imprecise	Undetected	3	Mean 58.6 (SD 16.3) vs. 64.1 (SD 18.1) on the FIQ (0-100 scale), p=0.048

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	release vs. sham	Function *Long-term*	1 (n=94)	Moderate	Unknown	Imprecise	Undetected	3	Mean 62.8 (SD 20.1) vs. 65.0 (19.8) on the FIQ (0-100 scale), p=0.329
		Pain *Short-term*	1 (n=64)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Pain *Intermediate-term*	2 (N=158)	Moderate	Inconsistent	Imprecise	Undetected	Insufficient	Insufficient evidence from one fair and one poor quality trial due to inconsistency in the estimates
		Pain *Long-term*	1 (n=94)	Moderate	Unknown	Imprecise	Undetected	3	MPQ sensory domain, mean 18.2 (SD 8.3) vs. 21.2 (7.9) on a 0-33 scale, p=0.038; MPQ evaluative domain, mean 23.2 (SD 7.6) vs. 26.7 (SD 6.9) on a 0-42 scale, p=0.036
		Harms	1 (n=94)	Moderate	Inconsistent	Imprecise	Undetected	Insufficient	Data for harms were insufficient; however, no adverse effect occurred in one fair quality trial
Mindfulness Practices	*Mindfulness-based stress*	Function *Short-term*	2 (N=1258)	Moderate	Consistent	Precise	Undetected	2	No clear effect: difference 0 to 0.06 on a 0-10 scale

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
	reduction or "Meditation Awareness Training: vs. waitlist or attention control	Pain *Short-term*	2 (N=1258)	Moderate	Consistent	Precise	Undetected	2	No clear effect: difference 0.1 on a 0-100 VAS pain scale in one poor quality trial; difference −1.38 to −1.59 on the affective and −0.28 to −0.71 on the sensory dimension (scales not reported) of the Pain Perception Scale in one fair-quality trial; Clinically meaningful improvement in function (≥14% on the FIQ total, 0-100 scale) was not different for MBSR versus either comparator in that trial; vs. AC%, RR 1.21 (95% CI 0.79 to 1.82; vs. WL, RR 1.37 (95% CI 0.83 to 1.94)
		Function *Intermediate term*	1 (n=148)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference −7.9, 95% CI −8.2 to −4.3 on 0-100 FIQ-R
		Pain *Intermediate term*	1 (n=148)	Moderate	Unknown	Precise	Undetected	3	Adjusted difference −3.0, 95% CI −4.1 to −1.9 on 0-45 SF-MPQ
		Harms	No studies	--	--	--	--	--	No evidence
Mind-Body Therapies	*Tai Chi, Qigong vs. waitlist or attention control*	Function *Short-term*	2 (N=154)	Moderate	Consistent[e]	Imprecise	Undetected	3	FIQ total score (0-100): Qigong, mean difference −7.5 (95% CI −13.3 to −1.68); Tai chi, mean difference −23.5 (95% CI −30 to −17) Heterogeneity may be explained by duration and intensity of intervention and control group
		Pain *Short-term*	2 (N=154)	Moderate	Consistent[e]	Imprecise	Undetected	3	Pooled difference −1.44, 95% CI −2.96 to −0.23; I^2=46%, scale 0-10

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Harms	2 (N=154)	Moderate	Inconsistent	Imprecise	Undetected	Insufficient	Data for harms were insufficient. One trial reported two adverse events judged to be possibly related to Qigong practice: an increase in shoulder pain and plantar fasciitis; neither participant withdrew from the study. In the trial of Tai chi, no adverse events were reported.
	Tai Chi vs. aerobic exercise	Function *Short to intermediate term*	1 (n=181)	Moderate	Unknown	Precise	Undetected	3	FIQ (Revised) 0-100 scale Short to intermediate term : Any tai chi (12 or 24 weeks of sessions) (N= 181) Difference in change scores −5.5, 95% CI −0.6 to −10.4
		Function *Intermediate term*	1 (n=89)	Moderate	Unknown	Imprecise	Undetected	3	FIQ (Revised) 0-100 scale 2 sixty-minute tai chi sessions/week for 24 weeks vs aerobic exercise 2 sixty-minute sessions/week for 24 weeks (N= 89): difference in change scores −16.2, 95% CI −8.7 to −23.6
		Function *Intermediate to long term*	1 (n=158)	Moderate	Unknown	Precise	Undetected	3	Intermediate to long term : Any tai chi (12 or 24 weeks of sessions) (N=158) Difference in change scores: −2.7 (95% CI −2.3 to 7.7); p=0.29
		Function *Long term*	1 (n=78)	Moderate	Unknown	Imprecise	Undetected	3	FIQ (Revised) 0-100 scale 2 sixty-minute sessions/week for 24 weeks. vs aerobic exercise 2 sixty-minute sessions/week for 24 weeks (N=78): Difference in change scores −11.1, 95% CI −2.7 to −19.6)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Harms	1 (n=226)	Moderate	Unknown	Imprecise	Undetected	3	No severe adverse events were reported for either treatment. Mild/moderate adverse events were reported for 5.3% of the tai chi participants and 5.3% of the aerobic exercise participants.
Acupuncture	*Acupuncture vs. sham*	Function *Short-term*	3 (N=283)	Moderate	Consistent	Precise	Undetected	2	Pooled difference −9.21, 95% CI −13.65 to −5.78, I^2=0%, 0-100 scale
		Function *Intermediate-term*	2 (N=211)	Moderate	Consistent	Precise	Undetected	2	Pooled difference −9.82, 95% CI −14.35 to −3.01, I^2=27.4%, 0-100 scale
		Pain *Short-term*	Sham or attention control 5 (N= 399) Sham control 4 (N=369)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference, all control conditions (5 trials): −1.14, 95% CI −2.56 to 0.33, I^2=91.6%, 0-10 scale. Pooled difference, sham only (4 trials): −0.86, 95% CI −2.73 to 0.92, I^2=88.9%,
		Pain *Intermediate-term*	3 (N=297)	Moderate	Inconsistent	Precise	Undetected	3	Pooled difference −0.65, 95% CI −1.15 to 0.17, I^2=45.5%, 0-10 scale
		Harms	4 (N=369)	Moderate	Consistent	Precise	Undetected	2	Discomfort and bruising were the most common reported adverse events and were more common in the true acupuncture groups. Discomfort was substantially more common for acupuncture or sham needling (61%to 70%) compared with simulated acupuncture (29%). Vasovagal symptoms and aggravation of fibromyalgia symptoms were less common (4% of sessions)

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
Multidisciplinary Rehabilitation	*Multi-disciplinary rehabilitation vs. usual care or waitlist*	Function *Short-term*	3 (N=381)	Moderate	Consistent[j]	Imprecise	Undetected	3	Pooled mean difference −6.08, 95% CI −14.17 to 0.16, I^2=48.9%, on 0-100 FIQ Proportion with clinically meaningful improvement in FIQ total score compared with usual care at short (OR 3.1, 95% CI 1.6 to 6.2)
		Function *Intermediate-term*	3 (N=394)	High	Consistent	Precise	Undetected	3	Pooled difference −7.77, 95% CI −12.22 to −3.83, I^2=0% Proportion with clinically meaningful improvement in FIQ total score compared with usual care at short (OR 3.1, 95% CI 1.5 to 6.4)
		Function *Long-term*	2 (N=311)	Moderate	Consistent	Precise	Undetected	3	Pooled difference −8.54, 95% CI −15.00 to −1.30, I^2=0% Proportion with clinically meaningful improvement in FIQ total score compared with usual care at short (OR 8.8, 95% CI 2.5 to 30.9)
		Pain *Short-term*	2 (N=341)	Moderate	Consistent[b]	Precise	Undetected	3	Pooled difference on 0-10 scale −0.24, 95%CI −0.63 to 0.15, I^2=0% (Excluding outlier)[c]
		Pain *Intermediate-term*	3 (N=394)	High	Consistent	Precise	Undetected	3	Pooled difference −0.68, 95% CI −1.10 to −0.27, I^2=0%
		Pain *Long-term*	2 (N=311)	Moderate	Consistent	Precise	Undetected	3	Pooled difference −0.25, 95% CI −0.79 to 0.36, I^2=0%

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Evidence Type	Findings, Direction and Magnitude of Effect
		Harms	1 (n=164)	High	Unknown	Imprecise	Undetected	Insufficient	Data were insufficient for harms; however, one poor quality trial reported that 19% (16/84) in the multidisciplinary group withdrew (versus 0% for waiting list), two gave increased pain as the reason. Reasons for other withdrawals were not given and there was not systematic reporting of adverse events
	Multi-disciplinary rehabilitation vs. exercise	Function *Long-term*	1 (n=155)	Moderate	Unknown	Precise	Undetected	3	Difference −1.10, 95% CI −8.40 to 6.20, on a 0-100 scale
		Pain *Long-term*	1 (n=155)	Moderate	Unknown	Precise	Undetected	3	Difference 0.10, 95% CI −0.67 to 0.87, on a 0-10 scale
		Harms	1 (n=155)	Moderate	Unknown	Imprecise	Undetected	Insufficient	Data were insufficient. Harms not reported

CBT = cognitive behavioral therapy; CI = confidence interval; EMG = electromyography; FIQ = Fibromyalgia Impact Questionnaire; MD = mean difference; MPQ = McGill Pain Questionnaire; NDI = Neck Disability Index; PSFS = Patient Specific Functional Scale; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; VAS = visual analog scale.

[a] Outlier excluded, Baptista 2012.

[b] I^2 >40% but not downgraded for inconsistency because direction of effect consistent across >75% of trials or heterogeneity explainable in subgroup/stratified/sensitivity analyses.

[c] Outlier excluded, Saral 2016.

Chronic tension headache

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Strength of Evidence	Findings, Direction and Magnitude of Effect
Psychological Therapies	*CBT vs. waitlist, attention control, or placebo*	Function *Short- and intermediate term*	1 (n=60)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
		Pain *Short-term*	2 (N=105)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from two small poor quality trials
		Pain *Intermediate-term*	1 (n=60)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
		Harms	1 (n=60)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial. The risk of withdrawal due to adverse events did not differ between CBT plus placebo and placebo alone (2% vs. 6%).
	Relaxation vs. waitlist of attention control	Pain, Harms *Short-term*	1 (n=55)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
	CBT vs. amitriptyline	Function *Short- and intermediate term*	1 (n=60)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
		Pain *Short-term*	2 (N=96)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from two small poor quality trials
		Pain *Intermediate-term*	1 (n=60)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
		Harms	2 (N=96)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from two small poor quality trial. Lower risk of “at least mild” adverse events in the CBT group (0% vs. 59%) in one poor quality trial; similar risk of withdrawal due to adverse events (2% in each group).
Physical Modalities	Occipital transcutaneous electrical stimulation vs. sham	Function, Pain, *Short-term*	1 (n=83)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one small poor quality trial
		Harms	1 (n=83)	High	Unknown	Precise	Undetected	Insufficient	Data for harms were insufficient; however, no adverse events occurred in either the real or the sham OTES group

Intervention	Comparator	Outcome	Number of RCTs (patients) Author Year	Study Limitations	Consistency	Precision	Reporting Bias	Strength of Evidence	Findings, Direction and Magnitude of Effect
Manual Therapies	Spinal manipulation vs. usual care	Function *Short-term*	1 (n=75)	Moderate	Unknown	Precise	Undetected	3	Difference −5.0, 95% CI −9.02 to −1.16, on the Headache Impact Test, scale 36-78; Difference −10.1, 95% CI −19.5 to −0.64, on the Headache Disability Inventory, scale 0-100
		Pain *Short-term*	1 (n=75)	Moderate	Unknown	Precise	Undetected	3	Difference −1.4 on a 0-10 NRS scale, 95% CI −2.69 to −0.16
		Harms	1 (n=75)	Moderate	Unknown	Precise	Undetected	3	No adverse events occurred in either group.
	Spinal manipulation vs. amitriptyline	Pain *Short-term*	1 (n=126)	High	Unknown	Precise	Undetected	Insufficient	Insufficient evidence from one poor quality trial
		Harms	1 (n=126)	High	Unknown	Precise	Undetected	3	Fewer adverse events with manipulation versus amitriptyline (RR 0.05, 95% CI 0.02 to 0.16), though the risk of withdrawal due to adverse events was not significantly different (RR 0.16, 95% CI 0.02 to 1.33). Common complaints were neck stiffness in the manipulation group and dry mouth, dizziness, and weight gain in the medication group
Acupuncture	Traditional Chinese needle acupuncture vs. sham	Pain *Short-term*	2 (N=69)	High	Consistent	Imprecise	Undetected	Insufficient	Insufficient evidence from two small, poor quality trials
		Pain *Intermediate- and long-term*	1 (n=30)	High	Unknown	Imprecise	Undetected	Insufficient	Insufficient evidence from one small, poor quality trial
		Harms	No studies	--	--	--	--	--	No evidence
	Laser acupuncture vs. sham laser	Pain *Short-term*	1 (n=50)	Moderate	Unknown	Precise	Undetected	3	Median difference −2, IQR 6.3, on a 0-10 VAS scale for pain intensity median difference −8, IQR 21.5, for number of headache days per month
		Harms	1 (n=50)	Moderate	Unknown	Precise	Undetected	3	No adverse events occurred in either group.

CBT = cognitive behavioral therapy; CI = confidence interval; IQR = interquartile range; NRS = numerical rating scale; RCT = randomized controlled trial; RR = risk ratio; VAS = visual analog scale

Nonopioid pharmacological treatments for chronic pain. This table is based on McDonagh MS, Selph SS, Buckley DI, et al. AHRQ Comparative Effectiveness Reviews. Nonopioid Pharmacologic Treatments for Chronic Pain. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020. The strength of evidence ratings in the AHRQ report were converted to ACIP-adapted GRADE evidence type ratings.

Efficacy

Neuropathic pain - placebo controlled trials

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	15 (N=4,832)	Fair	Direct	Consistent	Precise	Undetected	Small effect (NRS) Pregabalin/gabapentin vs. placebo MD -0.61 (-0.87 to -0.36), I^2=72%	2
		2 (N=493)	Fair	Direct	Consistent	Imprecise	Unknown	Small effect (VAS) Oxcarbazepine vs. placebo MD -0.89 (-1.50 to -0.37), I^2=0%	2
		6 (N=2,082)	Fair	Direct	Consistent	Precise	Unknown	Small effect (NRS) Duloxetine vs. placebo MD -0.79 (-1.10 to -0.49), I^2=43%	2
		2 (N=486)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (NRS) Cannabis vs. placebo no difference between groups (p=0.68 and 0.14)	3
		3 (N=1,519)	Fair	Direct	Consistent	Precise	Unknown	No effect (NRS) Capsaicin vs. Placebo MD -0.33 (-0.60 to -0.004), I^2=0%	2
		1 (n=45)	Fair	Direct	Unknown	Imprecise	Unknown	VAS Memantine vs. placebo mean change 1.82 (SD 2.77) vs. -2.36 (SD 3.35), p=0.87	Insufficient

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Response[a] (Dichotomous)	Short	15 (N=4,576)	Fair	Direct	Consistent	Imprecise	Undetected	Small effect (≥30%) Pregabalin/gabapentin vs. placebo RR 1.27 (1.12 to 1.50), I^2=72%	2
		1 (n=144)	Fair	Direct	Unknown	Imprecise	Unknown	Small effect Oxcarbazepine vs. placebo 45.6% vs. 28.9%, p=0.028	3
		6 (N=2,075)	Fair	Direct	Consistent	Imprecise	Unknown	Small effect (≥30%) Duloxetine vs. placebo RR 1.39 (1.22 to 1.62), I^2=39%	2
		1 (m=246)	Fair	Direct	Unknown	Imprecise	Unknown	Moderate effect Cannabis vs. placebo 28% vs. 16%; RR 1.70 (1.04 to 2.78), p=0.03	3
		3 (N=1,519)	Fair	Direct	Consistent	Precise	Unknown	No effect Capsaicin vs. placebo RR 1.17 (0.98 to 1.37), I^2=0%	2
Function	Short	1 (n=371)	Fair	Direct	Unknown	Imprecise	Unknown	No effect BPI Interference Gabapentin enacarbil vs. placebo MD -0.23 (-0.70 to 0.23)	3
		6 (N=2,082)	Fair	Direct	Consistent	Imprecise	Unknown	Small effect (BPI Interference) Duloxetine vs. placebo SMD -0.31 (-0.42 to -0.20), I^2=0%	3
		1 (n=303)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (BPI) Cannabis vs. placebo p=0.18	3

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Quality of Life	Short	3 (N=1,015) 3 (N=1,400) 3 (N=1,400)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (EQ-5D) Pregabalin/gabapentin vs. placebo SMD 0.24 (-0.07 to 0.54), I^2=58% (SF-36 MCS) Pregabalin/gabapentin vs. placebo MD 0.22 (-1.93 to 2.37) (SF-36 PCS) Pregabalin/gabapentin vs. placebo MD 0.80 (-0.29 to 2.07)	3
		2 (N=493)	Fair	Direct	Inconsistent	Precise	Unknown	(SF-36 MCS) Oxcarbazepine vs. placebo 47.2 vs. 50.2; p=0.03 (1 trial); No difference for other SF-36 scales	3
		3 (N=9,444)	Fair	Direct	Consistent	Precise	Unknown	Small effect (EQ-5D) Duloxetine vs. placebo MD 0.22 (0.05 to 0.38), I^2=0%	2
		2 (N=486)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (EQ-5D) Cannabis vs. placebo p=0.62 (SF-36) Cannabis vs. placebo p=not significant	3

[a] Pain Response main findings, percentages represent threshold for Pain Response

BPI = Brief Pain Inventory; CI = confidence interval; EQ-5D = EuroQol five dimensions; MCS = mental component score; MD = mean difference; NRS = numeric rating scale; PCS = physical component score; RR = risk ratio; SD = standard deviation; SF-36 = Short Form-36; SMD = standard mean difference; VAS = visual analogue scale

Neuropathic pain - cross-class comparisons

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	1 (n=152)	Fair	Direct	Unknown	Imprecise	Unknown	(VAS) Gabapentin vs. duloxetine No difference between groups (p=not reported)	Insufficient

VAS = visual analogue scale

Neuropathic pain - head-to-head comparisons

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	1 (N=301)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (NRS) Pregabalin vs. gabapentin enacarbil (p-values NR)	3
		2 (N=132)	Fair	Direct	Unknown	Imprecise	Unknown	(VAS) Pregabalin vs. gabapentin (p-value NR)	Insufficient
Function	Short	1 (N=301)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (BPI Interference) Pregabalin vs. gabapentin enacarbil (p-values NR)	3
Quality of Life	Short	1 (N=301)	Fair	Direct	Unknown	Imprecise	Unknown	No differences between Pregabalin vs. gabapentin enacarbil (p-values NR)	3

BPI = Brief Pain Inventory; MCS = mental component score; NR = not reported; NRS = numeric rating scale; PCS = physical component score; SF-36 = Short Form-36; VAS = visual analogue scale

Fibromyalgia - antidepressants

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	11 (N=5,936)	Fair	Direct	Consistent	Precise	Undetected	Small effect (0-10 scale) Antidepressants vs. placebo MD -0.59 (-0.80 to -0.43), I^2=26%	2
		1 (N=87)	Fair	Direct	Unknown	Imprecise	Unknown	(VAS 0-10) Amitriptyline vs. placebo MD -0.7 (Endpoint VAS 4.5 vs. 5.2, p=NR)	Insufficient
	Intermediate	3 (N=1,357)	Fair	Direct	Consistent	Precise	Unknown	Small effect (0-10 scale) Antidepressants vs. placebo MD -0.67 (-0.99 to -0.34), I^2=0%	2
Pain Response[a] (Dichotomous)	Short	10 (N=5,853)	Fair	Direct	Consistent	Precise	Undetected	Small effect (≥30%) Antidepressants vs. placebo RR 1.36 (1.26 to 1.46), I^2=0%	2
		1 (N=87)	Fair	Direct	Unknown	Imprecise	Unknown	(Physician's global assessment) Amitriptyline vs. placebo 74% vs. 49%, p=0.017	Insufficient
	Intermediate	3 (N=1,715)	Fair	Direct	Consistent	Precise	Unknown	Small effect (≥30%) Antidepressants vs. placebo RR 1.29 (1.08 to 1.52), I^2=0%	2
Function	Short	11 (N=6,240)	Fair	Direct	Consistent	Precise	Undetected	Small effect Antidepressants vs. placebo SMD -0.24 (-0.32 to -0.17), I^2=22%	2
	Intermediate	3 (N=1,724)	Fair	Direct	Consistent	Precise	Unknown	No effect Antidepressants vs. placebo SMD -0.13 (-0.24 to -0.02), I^2=0%	2
Quality of Life	Short	8 (N=5,487)	Fair	Direct	Consistent	Precise	Undetected	Small effect (SF-36 MCS or PCS, 0-100) Antidepressants vs. placebo MCS: SMD 0.19 (0.13 to 0.27), I^2=12% PCS: SMD 0.16 (0.10 to 0.22), I^2=0%	2

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
	Intermediate	3 (N=1,716)	Fair	Direct	Consistent	Precise	Unknown	Small effect (SF-36 MCS or PCS, 0-100) Antidepressants vs. placebo MCS: SMD 0.18 (0.08 to 0.30), I^2=0% PCS: SMD 0.07 (-0.10 to 0.24), I^2=0%	2

[a] Pain Response main findings, percentages represent threshold for Pain Response

BPI = brief pain inventory; CI = confidence interval; MCS = mental component score; MD = mean difference; NR = not reported; PCS = physical component score; SD = standard deviation; SE = standard error; SEM = standard error of the mean; SF-36 = Short Form-36; SMD = standard mean difference

Fibromyalgia - anticonvulsants

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	8 (N=4,747)	Fair	Direct	Consistent	Precise	Undetected	Small effect (0-10 scale) MD -0.57 (-0.75 to -0.40), I^2=30%	2
Pain Response[a] (Dichotomous)	Short	8 (N=4,773)	Fair	Direct	Consistent	Precise	Undetected	Small effect (≥30%) RR 1.30 (1.20 to 1.43), I^2=0%	2
Function	Short	8 (N=4,740)	Fair	Direct	Consistent	Precise	Undetected	Small effect (FIQ 0-80 or 0-100) SMD -0.22 (-0.29 to -0.15), I^2=0%	2
Quality of Life	Short	4 (N=2,520)	Fair	Direct	Consistent	Precise	Unknown	No effect (SF-36 MCS or PCS, 0-100) Pregabalin vs. placebo MCS: SMD 0.13 (0.04 to 0.22), I^2=0% PCS: SMD 0.17 (0.04 to 0.31), I^2=39%	2

[a] Pain Response main findings, percentages represent threshold for Pain Response

CI = confidence interval; FIQ = Fibromyalgia Impact Questionnaire; MCS = mental component score; MD = mean difference; PCS = physical component score; RR = risk ratio; SMD = standard mean difference; SF-36 = short form 36

Fibromyalgia - memantine and cross-class comparisons

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Intermediate	1 (N=63)	Good	Direct	Unknown	Imprecise	Unknown	Moderate effect (VAS, 0-10) Memantine vs. placebo: 4.87 vs. 7.01, p=0.001	3
		1 (N=208)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (MPQ present pain intensity, range NR) Cyclobenzaprine vs. amitriptyline vs. placebo 2.11 (SD 0.93, p<0.001 vs. baseline) vs. 2.17 (SD 1.02, p<0.001 vs. baseline) vs. 2.47 (SD 0.97, p<0.05 vs. baseline)	3
Function	Intermediate	1 (N=63)	Good	Direct	Unknown	Imprecise	Unknown	Moderate effect (FIQ, range NR) Memantine vs. placebo 50.02 vs. 69.57, p<0.001	3
		1 (N=208)	Fair	Direct	Unknown	Imprecise	Unknown	Magnitude of effect uncertain (HAQ, range NR) Cyclobenzaprine vs. amitriptyline vs. placebo 0.53 (SD 0.40) vs. 0.60 (SD 0.49) vs. 0.70 (SD 0.65)	Insufficient
Quality of Life	Intermediate	1 (N=63)	Good	Direct	Unknown	Imprecise	Unknown	Moderate effect (EQ-5D, 0-100) Memantine vs. placebo 60.48 vs. 43.75, Cohen's d -1.09, p=0.001	3

EQ-5D = EuroQol five dimensions; FIQ = Fibromyalgia Impact Questionnaire; HAQ = Health Assessment Questionnaire; MPQ = McGill Pain Questionnaire; NR = not reported; SD = standard deviation; VAS = visual analogue scale

Osteoarthritis – oral NSAIDS vs. placebo

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	27 (N=13,478)	Fair	Direct	Consistent	Precise	Possible	Small effect (NRS 0-10) MD -0.73 (-0.84 to -0.62), I^2=27%	2
Pain Response[a] (Dichotomous)	Short	15 (N=8,253)	Fair	Direct	Consistent	Precise	Undetected	Small effect RR 1.23 (1.18 to 1.31), I^2=0%	1
Function	Short	28 (N=13,473)	Fair	Direct	Consistent	Precise	Undetected	Small effect (WOMAC, LI) SMD -0.32 (-0.37 to -0.28), I^2=24%	1
Quality of Life	Short	3 (N=1,027)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (SF-36 MCS or PCS) MCS: MD 0.61 (-0.50 to 1.79) PCS: MD 2.95 (1.79 to 4.18)	2

[a] Pain Response main findings, percentages represent threshold for Pain Response

CI = confidence interval; LI = Lequesne Index; MD = mean difference; NRS = numeric rating scale; RR = risk ratio; SMD = standard mean difference; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Osteoarthritis - topical diclofenac vs. placebo

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	4 (N=1,541)	Fair	Direct	Consistent	Precise	Unknown	Small effect (WOMAC) MD -0.58 (-0.81 to -0.35), I^2=0%	2
Pain Response[a] (Dichotomous)	Short	3 (N=1,232)	Good	Direct	Consistent	Imprecise	Unknown	Small effect RR 1.20 (1.09 to 1.38), I^2=0%	2
Function	Short	4 (N=1,538)	Fair	Direct	Inconsistent	Precise	Unknown	No effect (WOMAC) MD -0.51 (-1.06 to 0.04), I^2=94%	3

[a] Pain Response main findings, percentages represent threshold for Pain Response

CI = confidence interval; MD = mean difference; RR = risk ratio; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Osteoarthritis - oral NSAIDs: head-to-head comparisons

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	4 (N=1,313)	Fair	Direct	Consistent	Imprecise	Unknown	Moderate effect (VAS, WOMAC pain Subscale) Diclofenac vs. celecoxib MD -12.2 (2.2 to 22.1) Small effect Diclofenac vs. 3.75 mg/d meloxicam No effect 7 mg vs. 15 mg No effect with other comparisons	3
	Intermediate	1 (n=586)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (VAS, WOMAC pain subscale) Celecoxib vs. naproxen	3
	Long	1 (n=916)	Fair	Direct	Unknown	Precise	Unknown	No significant differences between groups at endpoint (VAS) Celecoxib vs. diclofenac	3
Pain Response[a] (Dichotomous)	Short	2 (N=849)	Fair	Direct	Consistent	Imprecise	Unknown	No effect Diclofenac dispersible vs. enteric coated: RR 0.82 (0.73 to 1.09) Ibuprofen vs. nabumetone: RR 1.2 (0.88 to 1.66)	3
	Intermediate	1 (n=586)	Fair	Direct	Consistent	Imprecise	Unknown	No effect Celecoxib vs. naproxen	3
Function	Short	2 (N=301)	Fair	Direct	Consistent	Imprecise	Unknown	Moderate effect (WOMAC 0-68) Diclofenac vs. celecoxib RR 2.06 (1.37 to 3.08) No effect: Diclofenac vs. meloxicam 7 or 15 mg/d, but small effect over meloxicam 3.75 mg/d	3
	Intermediate	2 (N=921)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (WOMAC) Celecoxib vs. naproxen Meloxicam vs. diclofenac	3

[a] Pain Response main findings, percentages represent threshold for Pain Response
CI = confidence interval; MD = mean difference; RR = risk ratio; VAS = visual analogue scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Osteoarthritis - antidepressants: duloxetine vs. placebo

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	6 (N=1,508)	Good	Direct	Consistent	Precise	Unknown	Small effect (0 to10 scale) MD -0.75 (-1.05 to -0.53), I^2=15%	1
Pain Response[a] (Dichotomous)	Short	4 (N=1,247)	Good	Direct	Consistent	Precise	Unknown	Moderate effect RR 1.37 (1.24 to 1.52), I^2=0%	1
Function	Short	5 (N=1,480)	Good	Direct	Consistent	Precise	Unknown	Small effect (WOMAC, BPI) SMD -0.27 (-0.41 to -0.12), I^2=27%	1
Quality of Life	Short	2 (N=570)	Good	Direct	Consistent	Precise	Unknown	Small effect (EQ-5D, 0 to 1 scale) MD 0.05 (0.02 to 0.08), I^2=0%	1

[a] Pain Response main findings, percentages represent threshold for Pain Response
BPI = brief pain inventory; EQ-5D = EuroQol five dimensions; MD = mean difference; RR = risk ratio; SMD = standardized mean difference; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Osteoarthritis - acetaminophen vs. placebo

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	3 (N=1,082)	Fair	Direct	Consistent	Imprecise	Unknown	No effect MD -0.34 (-0.66 to 0.03), I^2=0%	3
	Intermediate	1 (n=212)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (WOMAC subscale) MD -0.30 (-0.77 to 0.17)	3
Pain Response[a] (Dichotomous)	Intermediate	1 (n=212)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (OARSI-A criteria) RR 1.58 (1.00 to 2.49), p=0.051	3
Function	Short	3 (N=1,081)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (WOMAC) SMD -0.14 (-0.29 to 0.04), I^2=0%	3
	Intermediate	1 (n=212)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (< small) (WOMAC, 1 to 100) MD -3.7 (-6.9 to -0.5)	3

[a] Pain Response main findings, percentages represent threshold for Pain Response
MD = mean difference; OARSI = Osteoarthritis Research Society International; SMD = standardized mean difference; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Osteoarthritis - anticonvulsants vs. antidepressants: duloxetine vs. pregabalin

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	1 (n=65)	Good	Direct	Unknown	Imprecise	Unknown	(NRS scale, 0 to 10): Pregabalin 300 mg/d: -2.7 (-3.5 to -1.9) vs. duloxetine 60 mg/d: -2.3 (-3.8 to -0.9) vs. placebo: -0.9 (-2.0 to 0.2); Pregabalin vs. placebo = 0.023 and 0.19	Insufficient
Function	Short	1 (n=65)	Good	Direct	Unknown	Imprecise	Unknown	(AUSCAN Function scale, 0 to 900): Pregabalin 300 mg/d: -46.4 (-341.7 to -151.0) vs. duloxetine 60 mg/d: -101.8 (-248.4 to -44.7) vs. placebo: -67.3 (-156.4 to -21.8); Pregabalin vs. placebo = 0.009 and >0.05	Insufficient

AUSCAN = Australian Canadian osteoarthritis hand index; CI = confidence interval; NRS = numeric rating scale

Osteoarthritis - acetaminophen vs. NSAIDs

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	1 (n=85)	Fair	Direct	Inconsistent	Imprecise	Unknown	(WOMAC) Diclofenac 150 mg/d vs. acetaminophen 4000 mg/d; diclofenac shows greater pain improvement (-53.9 vs. -23.8 WOMAC; p=0.003)	Insufficient
Function	Short	1 (n=25)	Fair	Direct	Unknown	Imprecise	Unknown	(WOMAC) Diclofenac 150 mg/d vs. acetaminophen 4000 mg/d; diclofenac shows greater function improvement (-163.0 vs. -41.8 WOMAC; p<0.001)	Insufficient

WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Inflammatory arthritis – oral NSAIDS vs. placebo

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	9 (N=4,543)	Fair	Direct	Consistent	Precise	Undetected	Small effect (0 to 10 scale) MD -0.97 (-1.33 to -0.74), I^2=39%	2
	Intermediate	1 (n=563)	Fair	Direct	Unknown	Precise	Unknown	Small effect (0 to 10 scale) Naproxen 1000 mg/d MD -0.53 (-0.93 to -0.13)	3
	Long	1 (n=365)	Fair	Direct	Unknown	Precise	Unknown	Large effect (0 to10 scale) Meloxicam 15-22.5 mg/d MD -2.10 (-2.72 to -1.48)	3
Pain Response[a] (Dichotomous)	Short	7 (N=3,434)	Fair	Direct	Consistent	Precise	Unknown	Moderate effect (ACR 20; ASAS 20) RR 1.58 (1.34 to 2.06), I^2=52%	2
	Intermediate	1 (n=563)	Fair	Direct	Unknown	Precise	Unknown	Small effect (ACR 20) Naproxen 1000 mg/d: RR 1.28 (1.03 to 1.60)	3
	Long	1 (n=365)	Fair	Direct	Unknown	Precise	Unknown	Large effect (≥ 50%) Meloxicam 15- 22.5 mg/d: RR 3.05 (1.98 to 4.71)	3
Function	Short	7 (N=4,284)	Fair	Direct	Consistent	Precise	Undetected	Small effect (HAQ; BASFI) SMD -0.34 (-0.51 to -0.20), I^2=67%	2
	Intermediate	1 (n=563)	Fair	Direct	Unknown	Precise	Unknown	Small effect (HAQ-DI, 0-3) Naproxen 1000 mg/d: MD -0.18 (-0.35 to -0.02)	3
	Long	1 (n=365)	Fair	Direct	Unknown	Precise	Unknown	No effect (ASFI[117], 0-40) Meloxicam 15-22.5 mg/d: MD -0.63 (-0.85 to -0.40)	3
Quality of Life	Short	2 (N=1,204)	Fair	Direct	Inconsistent	Imprecise	Unknown	(ASQoL, 0 to 18) Naproxen 1000 mg/d: MD -2.9; p=0.04 (SF-36 PCS and MCS) Celecoxib 200-800 mg/d or Naproxen 1000 mg/d	Insufficient

[a] Pain Response main findings, percentages represent threshold for Pain Response

ACR = American College of Rheumatology; ASAS = Assessment of SpondyloArthritis international Society; ASFI = Ankylosing Spondylitis Functional Index; ASQoL = Ankylosing Spondylitis Quality of Life; BASFI = Bath Ankylosing Spondylitis Functional Index; HAQ = Health Assessment Questionnaire; MCS = mental component score; MD = mean difference; PCS = physical component score; RR = risk ratio; SMD = standardized mean difference

Inflammatory arthritis – oral NSAIDS: head-to-head comparisons

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	3 (N=1,453)	Fair	Direct	Consistent	Precise	Unknown	No effect (0-10 scale) Celecoxib 200-400 mg/d vs. diclofenac 150 mg/d: NS	2
		2 (N=1,132	Fair	Direct	Consistent	Imprecise	Unknown	No effect (0-10 scale) Celecoxib 200-400 mg/d vs. naproxen 1000 mg/d: NS	3
		1 (n=103)	Fair	Direct	Unknown	Imprecise	Unknown	(Non-visual scale, 1-5) Diclofenac vs. etodolac	Insufficient
		1 (n=717)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (VAS, 0-100) Diclofenac vs. meloxicam	3
		1 (n=39)	Fair	Direct	Unknown	Imprecise	Unknown	(Non-visual scale, 1-5) Etodolac vs. naproxen	Insufficient
		2 (N=621)	Fair	Direct	Consistent	Precise	Unknown	No effect (0-10 scale) Nabumetone 2000 mg/d vs. naproxen 1000 mg/d: NS	3
	Intermediate	1 (n=379)	Fair	Direct	Unknown	Precise	Unknown	No effect (VAS, 0-100) Meloxicam vs. naproxen	3
		1 (n=47)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (% better, %same, %worse) Nabumetone vs. naproxen	3
Pain Response[a] (Dichotomous)	Short	3 (N=1,443)	Fair	Direct	Consistent	Precise	Unknown	No effect (ACR 20; ASAS 20) Celecoxib 200-400 mg/d vs. diclofenac 150 mg/d: NS	2
		2 (N=1,133)	Fair	Direct	Inconsistent	Imprecise	Unknown	No effect (ACR 20; ASAS 20) Celecoxib 200-400 mg/d vs. naproxen 1000 mg/d: NS	3
		1 (n=344)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (Pt global assess, % improved) Nabumetone 2000 mg/d vs. naproxen 1000 mg/d: NS	3
Function	Short	3 (N=1,448)	Fair	Direct	Consistent	Precise	Unknown	No effect (mHAQ; BASFI) Celecoxib 200-400 mg/d vs. diclofenac 150 mg/d: NS	2

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
		2 (N=1,373)	Fair	Direct	Consistent	Imprecise	Unknown	No effect (HAQ; BASFI) Celecoxib 200-400 mg/d vs. naproxen 1000 mg/d: NS	3
		1 (n=103)	Fair	Direct	Unknown	Imprecise	Unknown	(Non-visual scale, 1-4) Diclofenac vs. etodolac	Insufficient
		1 (n=717)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (mHAQ, max 3) Diclofenac vs. meloxicam	3
		1 (n=346)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (ACR class) Nabumetone 2000 mg/d vs. naproxen 1000 mg/d: NS	3
Quality of Life	Short	1 (n=917)	Fair	Direct	Unknown	Precise	Unknown	No effect (SF-36 PCS and MCS) Celecoxib vs. naproxen	3

[a] Pain Response main findings, percentages represent threshold for Pain Response

ACR = American College of Rheumatology; ASAS = Assessment of SpondyloArthritis international Society; BASFI = Bath Ankylosing Spondylitis Functional Index; (m)HAQ = (modified) Health Assessment Questionnaire; MCS = mental component score; NS = not significant; PCS = physical component score; SF-36 = Short Form-36; SMD = standardized mean difference; VAS = visual analogue score

Inflammatory arthritis – antidepressants: placebo controlled trials

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	1 (n=36)	Fair	Direct	Unknown	Imprecise	Unknown	(Non-visual scale, 0-4) Amitriptyline 50-75 mg/d	Insufficient

Low back pain – antidepressants and anticonvulsants: placebo controlled trials

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Pain Improvement (Continuous)	Short	3 (N=1,491)	Fair	Direct	Consistent	Precise	Unknown	Small effect (BPI Pain Scale) Duloxetine vs. placebo MD -0.50 (-0.71 to -0.29), I^2=0%	2
		1 (n=78)	Fair	Direct	Unknown	Imprecise	Unknown	(DDS, 0-20) Despiramine vs. placebo MD -0.80 (-2.64 to 1.04)	Insufficient
		1 (n=86)	Fair	Direct	Unknown	Imprecise	Unknown	(DDS, 0-20) Fluoxetine vs. placebo MD 0.70 (-1.40 to 2.80)	Insufficient
		1 (n=108)	Fair	Direct	Unknown	Imprecise	Unknown	(DDS, 0-20) Gabapentin vs. placebo p=0.42	Insufficient
	Intermediate	1 (n=146)	Good	Direct	Unknown	Imprecise	Unknown	No effect (VAS) Amitriptyline vs. placebo MD -7.81 (-15.7 to 0.10)	3
Pain Response[a] (Dichotomous)	Short	3 (N=1,235)	Fair	Direct	Consistent	Imprecise	Unknown	Small effect Duloxetine vs. placebo RR 1.25 (1.11 to 1.40), I^2=0%	3
Function	Short	3 (N=1,214)	Fair	Direct	Consistent	Precise	Unknown	No effect (BPI Interference Scale) Duloxetine vs. placebo MD -0.36 (-0.73 to -0.04), I^2=34%	2
		1 (n=78)	Fair	Direct	Unknown	Imprecise	Unknown	(RMDQ score) Despiramine vs. placebo 2.3 vs. 4.1, p=0.05 (Physician-rated CGI) Despiramine vs. placebo 5.9 vs. 4.8, p=0.003	Insufficient
		1 (n=108)	Fair	Direct	Unknown	Imprecise	Unknown	(CGI-C "minimal improvement") Gabapentin vs. placebo 37% vs. 33%, p=0.95	Insufficient
	Intermediate	1 (n=146)	Good	Direct	Unknown	Imprecise	Unknown	No effect (RMDQ) Amitriptyline vs. placebo MD -0.98 (-2.42 to 0.46)	3
Quality of Life	Short	3 (N=1,198)	Fair	Direct	Consistent	Precise	Unknown	No effect Duloxetine vs. placebo SMD 0.18 (-0.03 to 0.39), I^2=38%	2

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
		1 (n=108)	Fair	Direct	Unknown	Imprecise	Unknown	(BDI-II) Gabapentin vs. placebo p=0.52	Insufficient

[a] Pain Response main findings, percentages represent threshold for Pain Response

BDI-II = Beck Depression Inventory-II; BPI = brief pain inventory; CGI = clinical global impression scale; CI = confidence interval; DDS = Descriptor Differential Scale; MD = mean difference; RMDQ = Roland-Morris Disability Questionnaire; SMD = standard mean difference; VAS = visual analogue scale

Low back pain - antidepressants: head-to-head trials

Outcome	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size	Evidence Type
Pain Improvement (Continuous)	Short	1 (n=200)	Fair	Direct	Unknown	Imprecise	Unknown	Small effect (VAS) Amitriptyline vs. pregabalin Mean change from baseline: 2.9 vs. 3.9, p=0.03	3
Function	Short	1 (n=200)	Fair	Direct	Unknown	Imprecise	Unknown	No effect (ODI) Amitriptyline vs. pregabalin Mean change from baseline: p=0.09	3

ODI = Oswestry Disability Index; VAS = visual analogue scale;

Harms

Adverse events - antidepressants

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	vidence Type
SAE	SNRI Antidepressants vs. Placebo	Short	19 (N=8,832)	Fair	Direct	Consistent	Imprecise	Undetected	No effect RR 0.88 (0.62 to 1.24), I^2=0%	3
		Intermediate	2 (N=1,218)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 0.86 (0.35 to 2.24), I^2=0%	3
WAE	SNRI Antidepressants vs. Placebo	Short	24 (N=9,971)	Fair	Direct	Consistent	Precise	Undetected	Moderate effect RR 1.99 (1.71 to 2.35), I^2=18%	2
		Intermediate	3 (N=1,738)	Fair	Direct	Consistent	Precise	Unknown	Moderate effect RR 1.83 (1.23 to 2.61), I^2=4%	2
	TCA Antidepressants vs. placebo	Short	5 (N=478)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 1.49 (0.89 to 3.01)	3
		Intermediate	1 (n=126)	Far	Direct	Unknown	Imprecise	Unknown	RR 1.75 (0.38 to 8.06)	Insufficient
Nausea	SNRI Antidepressants vs. Placebo	Short	19 (N=8,929)	Fair	Direct	Consistent	Precise	Undetected	Large effect RR 3.10 (2.50 to 4.06), I^2=60%	2
		Intermediate	3 (N=1,738)	Fair	Direct	Consistent	Imprecise	Unknown	Moderate effect RR 1.98 (1.57 to 2.82), I^2=0%	3
Sedation	SNRI Duloxetine vs. Placebo	Short	16 (N=5,831)	Fair	Direct	Consistent	Precise	Undetected	Large effect RR 2.46 (2.00 to 3.01), I^2=0%	2
		Intermediate	2 (N=850)	Fair	Direct	Consistent	Imprecise	Unknown	Large effect RR 3.51 (1.46 to 11.05), I^2=0%	3
Dry Mouth	TCA Antidepressants vs Placebo	Short	1 (n=131)	Fair	Direct	Unknown	Imprecise	Unknown	RR 1.80, (1.14 to 2.85)	Insufficient
Cognitive effects	SNRI Antidepressants vs. Placebo	Short	2 (N=805)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 3.24 (0.26 to 40.17), I2=0	3

CI = confidence interval; RR = risk ratio; SAE = serious adverse event; SNRI = serotonin-norepinephrine reuptake inhibitor; TCA = tricyclic antidepressant; WAE = withdrawal due to adverse event

Adverse events – anticonvulsants

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study = Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	Oxcarbazepine vs. Placebo	Short	2 (N=493)	Fair	Direct	Consistent	Imprecise	Unknown	No effect 8.9% vs. 4.8% RR 1.82 (0.74 to 5.05), I^2=0%	3
	Pergabalin/Gabapentin vs. Placebo	Short	19 (N=7,982)	Fair	Direct	Consistent	Imprecise	Undetected	No effect 2.3% vs. 2.5% RR 0.90 (0.63 to 1.30), I^2=0%	3
WAE	Oxcarbazepine vs. Placebo	Short	2 (N=493)	Fair	Direct	Consistent	Imprecise	Unknown	Large effect 25.7% vs. 7.2% RR 3.64 (1.86 to 7.12), I^2=0%	3
	Pergabalin/Gabapentin vs. Placebo	Short	26 (N=9,754)	Fair	Direct	Consistent	Precise	Undetected	Moderate effect 14.4% vs. 7.0% RR 1.73, (1.48 to 2.01), I^2=5%	2
Blurred Vision	Pregabalin/Gabapentin vs. Placebo	Short	12 (N=5,127)	Fair	Direct	Consistent	Imprecise	Undetected	Large effect 5.8% vs. 1.4% RR 3.79 (2.20 to 7.19), I^2=29%	3
Cognitive Effects	Pregabalin/Gabapentin vs. Placebo	Short	8 (N=3,801)	Fair	Direct	Consistent	Imprecise	Undetected	Large effect 4.8% vs. 1.3% RR 3.15 (1.86 to 5.51), I^2=0%	3
Dizziness	Pregabalin/Gabapentin vs. Placebo	Short	25 (N=9,696)	Fair	Direct	Consistent	Precise	Undetected	Large effect 25.6% vs. 7.4% RR 2.97 (2.53 to 3.50, I^2=31%	2
Peripheral Edema	Pregabalin/Gabapentin vs. Placebo	Short	22 (N=9,005)	Fair	Direct	Consistent	Precise	Undetected	Large effect 8.8% vs. 3.7% RR 2.32 (1.80 to 3.09), I^2=26%	2
Sedation	Pregabalin/Gabapentin vs. Placebo	Short	24 (N=9,652)	Fair	Direct	Consistent	Precise	Undetected	Large effect 17% vs. 5.4% RR 3.03 (2.62 to 3.67), I^2=0%	2
	Oxcarbazepine vs. Placebo	Short	2 (N=490)	Fair	Direct	Consistent	Imprecise	Unknown	No effect 8.6% vs. 3.0% RR 3.13 (0.74 to 16.08), I^2=0%	3
Weight Gain	Pregabalin/Gabapentin vs. Placebo	Short	21 (N=8,620)	Fair	Direct	Consistent	Precise	Undetected	Large effect 10.1% vs. 2.8% RR 3.57 (2.77 to 4.91), I^2=7%	2

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study = Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Hyponatremia	Oxcarbazepine vs. Placebo	Short	2 (N=490)	Fair	Direct	Consistent	Imprecise	Unknown	No effect 2.8% vs. 0.0% RR 5.93 (0.55 to 63.8), I^2=0%	3

CI = confidence interval; RR = risk ratio; SAE = serious adverse event; WAE = withdrawal due to adverse event

Adverse events – NSAIDs

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	NSAIDs vs. Placebo	Short	23 (N=13,082)	Fair	Direct	Consistent	Imprecise	Undetected	No effect RR 0.96 (0.72 to 1.29), I^2=0%	3
		Intermediate	1 (n=563)	Fair	Direct	Unknown	Imprecise	Unknown	RR 0.51 (0.05 to 5.58)	Insufficient
	Topical diclofenac vs placebo	Short	2 (N=912)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 1.03 (0.29 to 27.01), I^2=0%	3
WAE	NSAIDs vs. Placebo	Short	38 (N=20,060)	Fair	Direct	Consistent	Precise	Undetected	Small effect RR 1.30 (1.14 to 1.49), I^2=13%	2
		Intermediate	2 (N=941)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 1.59 (0.89 to 3.08), I^2=0%	3
		Long	1 (n=365)	Fair	Direct	Unknown	Imprecise	Unknown	RR 1.59 (0.81 to 3.12)	Insufficient
	Topical diclofenac vs placebo	Short	4 (N=1,549)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 1.03 (0.29 to 27.01), I^2=0%	3
CV Events	NSAIDs vs. Placebo	Short	1 SR (639 RCTs, unclear N patients)	Fair	Direct	Consistent	Precise	Unknown	Small effect Diclofenac RR 1.41 (1.12 to 1.78) No effect Ibuprofen RR 1.44 (0.89 to 2.33) Naproxen RR 0.93 (0.69 to 1.27) Celecoxib RR 1.36 (1.00 to 1.84)	2

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
	Celecoxib vs. nonselective NSAIDs	Intermediate	3 RCTs (N=33,064)	Fair	Direct	Consistent	Precise	Unknown	No effect Cardiovascular death, nonfatal myocardial infarction, or nonfatal stroke: celecoxib 1.7%; ibuprofen 1.9%; naproxen 1.8% (p<0.001 for noninferiority between drugs)	2
		Long	1 RCT (n=7,297)	Good	Direct	Unknown	Precise	Unknown	No effect Hospitalization for non-fatal MI or other biomarker positive acute coronary syndrome, non-fatal stroke or CV death hazard ratio 1.12 (0.81 to 1.55)	2
Serious GI Events	NSAIDs vs. Placebo	Short	1 SR (639 RCTs, unclear N patients); 13 RCTs (N=7,262)	Fair	Direct	Consistent/ Inconsistent	Precise	Undetected	Moderate effect EPC meta-analysis NSAIDs vs. placebo RR 3.04 (1.73 to 5.11), I^2=73% IPD meta-analysis coxibs RR 1.81 (1.17 to 2.81); Diclofenac RR 1.89 (1.16 to 3.09); Ibuprofen RR 3.97 (2.22 to 7.10); Naproxen RR 4.22 (2.71 to 6.56); Celecoxib vs. placebo: 1.02 (0.47 to 1.56; 3 RCTs, N=1,877), I^2= 0%	2 (non-selectives) 3 (celecoxib)
	Coxibs (celecoxib) vs. nonselective NSAIDs	Short	1 SR (639 RCTs, unclear N patients); 13 RCTs (N=7,262)	Fair	Direct and Indirect	Inconsistent	Imprecise	Undetected	No clear effect SR (4 RCTs, N=1,755) OR 0.61 (0.15 to 2.43), I^2=38% Placebo trials: Celecoxib RR 1.04 (0.67 to 1.54), I^2=0% Nonselective NSAIDs RR 4.29 (2.75 to 6.93), I^2=46%; p<0.001 for interaction	Insufficient
		Intermediate	1 RCT (n=8,067)	Fair	Direct	Unknown	Precise	Unknown	Moderate effect OR 1.82 (1.31 to 2.55)	3

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
Serious Hepatic Events	NSAIDs vs. Placebo	Intermediate	1 SR (64 RCTs)	Fair	Direct	Unclear	Imprecise	Unknown	Large effect Aminotransferase >3x upper limit of normal: Diclofenac 3.55% (3.12% to 4.03%) vs. 0.29% (0.17% to 0.51%) Large effect Liver-related discontinuations: Diclofenac 2.17% (1.78% to 2.64%) vs. 0.08% (0.02% to 0.29%) No effect Liver-related SAE: Naproxen 0.06% (0.02% to 0.15%) vs. 0.00% (0.00% to 0.08%)	3

CI = confidence interval; CV = cardiovascular; GI = gastrointestinal; IPD = individual patient data; NSAIDs = nonsteroidal anti-inflammatory drug; RCT = randomized controlled trial; RR = risk ratio; SAE = serious adverse event; SR = systematic review; WAE = withdrawal due to adverse event

Adverse events - acetaminophen vs. placebo

	Duration	Number of Studies (n) participants	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	Short	2 (N=1,023)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 2.57 (0.60 to 10.8); I^2=0%	3
	Intermediate	1 (n=212)	Fair	Direct	Unknown	Imprecise	Unknown	No effect RR 0.96 (0.29 to 3.23)	3
WAE	Short	2 (N=1,023)	Fair	Direct	Consistent	Imprecise	Unknown	No effect RR 1.14 (0.67 to 1.95); I^2=0%	3
	Intermediate	1 (n=212)	Fair	Direct	Unknown	Imprecise	Unknown	No effect RR 1.28 (0.56 to 2.92)	3

SAE = serious adverse event; WAE = withdrawal due to adverse event; RR = risk ratio

Adverse events - capsaicin vs. placebo

	Duration	Number of Studies (n) participants	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	Short	3 (N=1,051)	Good	Direct	Consistent	Imprecise	Unknown	No effect RR 1.32 (0.71 to 3.47), I^2=0%	2
WAE	Short	2 (N=896)	Good	Direct	Consistent	Imprecise	Unknown	No effect RR 1.04 (0.08 to 17.1), I^2=0%	2
Application Site Erythema	Short	3 (N=1,051)	Good	Direct	Consistent	Imprecise	Unknown	2 effect RR 1.46 (1.29 to 1.66) , I^2=0%	2
Application Site Pain	Short	3 (N=1,051)	Good	Direct	Consistent	Imprecise	Unknown	Large effect RR 2.26 (1.61 to 2.82) , I^2=0%	2
Application Site Pruritus	Short	3 (N=1,051)	Good	Direct	Consistent	Imprecise	Unknown	No effect RR 1.70 (0.92 to 3.35) , I^2=0%	2

SAE = serious adverse event; WAE = withdrawal due to adverse event; RR = risk ratio

Adverse events - cannabis vs. placebo

Adverse Event	Cannabis type	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	Dronabinol 7.5-15mg/d	Short	1 (n=240)	Good	Direct	Unknown	Imprecise	Unknown	No effect 13.7% vs. 8.9% RR 1.58 (0.75 to 3.30)	3
	THC 2.7m/microL + CBD 2.5gm/microL	Short	1 (n=246)	Fair	Direct	Unknown	Imprecise	Unknown	No effect 8% vs. 5% RR 1.54 (0.58 to 4.10)	3
WAE	Dronabinol 7.5-15mg/d	Short	1 (n=240)	Good	Direct	Unknown	Imprecise	Unknown	No effect 14.5% vs. 14.0% RR 1.05 (0.56 to 1.96)	3
	THC 2.7m/microL + CBD 2.5gm/microL	Short	1 (n=246)	Fair	Direct	Unknown	Imprecise	Unknown	Large effect 19% vs. 6% RR 3.16 (1.41 to 7.06)	3
Dizziness	Dronabinol 7.5-15mg/d	Short	1 (n=240)	Good	Direct	Unknown	Imprecise	Unknown	Large effect 20% vs. 4.3% RR 4.68 (1.85 to 11.8)	3
	THC 2.7m/microL + CBD 2.5gm/microL	Short	1 (n=246)	Fair	Direct	Unknown	Imprecise	Unknown	Large effect 39% vs. 9% RR 4.55 (2.48 to 8.32)	3
Nausea	Dronabinol 7.5-15mg/d	Short	1 (n=240)	Good	Direct	Unknown	Imprecise	Unknown	No effect 4.2% vs. 6.8% RR 1.39 (0.40 to 4.80)	3
	THC 2.7m/microL + CBD 2.5gm/microL	Short	1 (n=246)	Fair	Direct	Unknown	Imprecise	Unknown	Large effect 17% vs. 8% RR 2.25 (1.8 to 4.70)	3
Sedation	THC 2.7m/microL + CBD 2.5gm/microL	Short	1 (n=246)	Fair	Direct	Unknown	Imprecise	Unknown	3% vs. 0% RR 8.30 (0.45 to 152.58)	Insufficient

SAE = serious adverse event; WAE = withdrawal due to adverse event; RR = risk ratio; min = minute; THC = tetrahydrocannabinol; CBD = Cannabidiol

Adverse events - skeletal muscle relaxants

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
WAE	Cyclobenzaprine vs. Placebo	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	No effect 16% vs. 5%; p=0.20 RR 2.82 (0.65 to 12.1)	3
	Cyclobenzaprine vs. Amitriptyline	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	No effect 16% vs. 8% RR 2.25 (0.82 to 6.20)	3
Dizziness	Cyclobenzaprine vs. Placebo	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	6.1% vs. 2.4%; p=0.38 RR 2.56 (0.31 to 21.22)	Insufficient
	Cyclobenzaprine vs. Amitriptyline	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	61.% vs. 0% RR 11.27 (0.63 to 200.53)	Insufficient
Sedation	Cyclobenzaprine vs. Placebo	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	3.7% vs. 2.4%; p=0.23 RR 2.00 (023 to 17.34)	Insufficient
	Cyclobenzaprine vs. Amitriptyline	Intermediate	1 (n=208)	Fair	Direct	Unknown	Imprecise	Unknown	3.7% vs. 4.8% RR 1.30 (0.30 to 5.64)	Insufficient

WAE = withdrawal due to adverse event

Adverse events – memantine vs. placebo

Adverse Effect	Comparison	Duration	Number of Studies (n participants)	Study Quality	Directness	Consistency (1 study= Unknown)	Precision	Publication Bias	Main Findings Effect Size (95% CI)	Evidence Type
SAE	Neuropathic Pain	Short	1 (n=45)	Fair	Direct	Unknown	Imprecise	Unknown	Reported as “no differences”	Insufficient
	Fibromyalgia	Medium	1 (n=63)	Good	Direct	Unknown	Imprecise	Unknown	Reported as “no serious adverse events”	Insufficient
WAE	Neuropathic Pain	Short	1 (n=45)	Fair	Direct	Unknown	Imprecise	Unknown	Reported as “no differences”	Insufficient
	Fibromyalgia	Medium	1 (n=63)	Good	Direct	Unknown	Imprecise	Unknown	6% vs. 3%; p=0.55	Insufficient
Dizziness	Fibromyalgia	Medium	1 (n=63)	Good	Direct	Unknown	Imprecise	Unknown	25.8% vs. 12.5%; RR 2.06 (0.69 to 6.16), p=0.22	Insufficient
Sedation	Fibromyalgia	Medium	1 (n=63)	Good	Direct	Unknown	Imprecise	Unknown	0% vs. 6%; RR 0.21 (0.01 to 4.13), p=0.30	Insufficient

CI = confidence interval; RR = risk ratio; SAE = serious adverse event; WAE = withdrawal due to adverse event

Treatments for acute pain. This table is based on Chou R, Wagner J, Ahmed AY, et al. AHRQ Comparative Effectiveness Reviews. Treatments for Acute Pain: A Systematic Review. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020. The strength of evidence ratings in the AHRQ report were converted to ACIP-adapted GRADE evidence type ratings.

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Key Question 1: Acute low back pain										
Opioid vs. NSAID	Pain	<1 d; 1 d to <1 w	1	113	Direct	Precise	Fair	Unable to assess	No differences	3
	Serious AEs, study withdrawal due to AEs, any AE	<1 d; 1 d to 1 w	1	113	Direct	Precise	Fair	Unable to assess	Higher risk with opioid	3
Opioid vs. muscle relaxant	Pain, function	1 to <2 w; ≥4 w	1	216	Direct	Imprecise	Good	Unable to assess	No differences	3
	Dizziness, nausea or vomiting	1 to <2 w; ≥4 w	1	216	Direct	Precise	Good	Unable to assess	Higher risk with opioid	2
Muscle relaxant vs. benzodiazepine	Pain	1 d to <1 w	2	110	Direct	Imprecise	Fair	Consistent	Small to moderate decrease in pain with muscle relaxant	3
	AEs	1 d to <1 w	2	110	Direct	Imprecise	Fair	Inconsistent	Unable to determine	Insufficient
NSAID or muscle relaxant vs. manipulation	Pain, function	1 to <2, 2 to <4, and ≥4 w	3	320	Direct	Precise	Fair	Inconsistent	Likely no differences	3
	AEs	1 to <2, 2 to <4, and ≥4 w	3	320	Direct	Imprecise	Fair	Inconsistent	Unable to determine	Insufficient
Acupuncture vs. NSAID	Pain, function	2 to <4 w and ≥4 w	1	58	Direct	Imprecise	Fair	Unable to assess	Moderate improvement in pain and function with acupuncture	3
Exercise vs. usual care	Pain, function	1 to 52 w	2	194	Direct	Imprecise	Fair	Consistent	No differences	3

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
	AEs	--	--	--	--	--	--	--	No evidence	Insufficient
Exercise vs. bed rest	Pain, function	1 to 52 w	3	384	Direct	Precise	Fair	Consistent	No differences	2
	Sick days	2 to 4 w and ≥4 w	1	100	Direct	Precise	Fair	Unable to assess	Fewer sick days with opioid	3
	AEs	1 to 52 w	3	384	Direct	Imprecise	Fair	Unable to assess	Sparse data	Insufficient
Traditional Chinese acupuncture vs. sham or usual care	Pain, function	2 to <4 w	1	261	Direct	Imprecise	Fair	Inconsistency based on type of sham	Acupuncture decreased persistent pain vs. nonpenetrating sham or usual care, but not needle sham	3
	Pain, function	≥4 w	1	261	Direct	Imprecise	Fair	Unable to assess	No differences	3
	Serious AEs, study withdrawal due to AEs	2 to <4 w, ≥4 w	1	261	Direct	Imprecise	Fair	Unable to assess	No events reported	3
Brace vs. no brace, osteoporotic compression fracture	Pain, function, opioid use	2 to <4 w, ≥4 w	1	85	Direct	Precise	Fair	Unable to assess	No differences	3
Heat therapy vs. usual care or placebo	Pain, function	1 d to <1 w, 1 to <2 w, 2 to <4 w	6	425	Direct	Imprecise to precise	Fair	Consistent	Moderate improvement in pain and function with heat therapy	3 to 2
	Adverse events	1 d to <1 w, 1 to <2 w, 2 to <4 w	6	425	Direct	Imprecise	Fair	Consistent	No serious AEs and few non-serious AEs	3
Manipulation vs. inactive controls, no radiculopathy	Pain, function	1 d to <1 w, 1 to <2 w, 2 to <4 w, or ≥4 w	6	555	Direct	Imprecise	Fair	Consistent	No differences	3 to 2

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
	Adverse events	1 d to <1 w, 1 to <2 w, 2 to <4 w, or ≥4 w	6	555	Direct	Imprecise	Fair	Consistent	Limited reporting, few or no serious AEs	3
Manipulation vs. sham, radiculopathy	Pain	2 to <4 w, ≥4 w	1	102	Direct	Precise	Good	Unable to assess	Decreased likelihood of pain with manipulation	3
	AEs	2 to <4 w, ≥4 w	1	102	Direct	Imprecise	Good	Unable to assess	No AEs reported in either group	3
Key Question 2: Acute neck pain										
Collar vs. usual activity, neck pain with radiculopathy	Pain, function	2 to <4 w, ≥4 w	1	135	Direct	Imprecise	Fair	Unable to assess	Moderate to large decrease in pain with collar, no difference in function	3
Brace vs. exercise, neck pain with radiculopathy	Pain, function	2 to <4 w, ≥4 w	1	139	Direct	Imprecise	Fair	Unable to assess	No differences	3
Exercise vs. usual activity, neck pain with radiculopathy	Pain, function	2 to <4 w, ≥4 w	1	136	Direct	Imprecise	Fair	Unable to assess	Moderate to large decrease in pain with exercise, no difference in function	3
Ultrasound vs. sham, neck pain with radiculopathy	Pain	1 to <2 w, 2 to <4 w	1	54	Direct	Imprecise	Fair	Unable to assess	No difference at 1 to <2 w, small decrease with ultrasound at 2 to <4 w	3
Collar vs. usual activity, whiplash neck strain	Pain, health status	≥4 weeks	1	303	Direct	Precise	Fair	Unable to assess	No difference at ≥4 weeks	3

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Collar vs. exercise, whiplash neck strain	Pain, health status	≥4 weeks	1	297	Direct	Precise	Fair	Unable to assess	No difference at ≥4 weeks	3
Exercise vs. usual activity, whiplash neck strain	Pain, health status	≥4 weeks	1	296	Direct	Precise	Fair	Unable to assess	No difference at ≥4 weeks	3
Key Question 3: Other musculoskeletal pain										
NSAID vs. acetaminophen	Pain	<1 d, 1 d to <1 w, 1 to <2 w, ≥4 w	8	1,100	Direct	Imprecise	Good	Consistent	No differences	2
Ultrasound vs. sham	Pain	1 d to <1 w, ≥4 w	3	190	Direct	Imprecise	Fair	Consistent	No differences	3
Acupressure vs. sham acupressure or usual care	Pain, health status	1 d to <1 w, ≥4 w	1	62	Direct	Precise	Fair	Unable to assess	Moderate decrease in pain and small improvement in health status with acupressure	3
Key Question 4: Acute neuropathic pain										
Opioid vs. gabapentin, herpes zoster	Pain	1 to <2 w, ≥4 w	1	45	Direct	Imprecise	Fair	Unable to assess	Increased likelihood of improvement in pain	3
	Constipation	1 to <2 w, ≥4 w	1	45	Direct	Imprecise	Fair	Unable to assess	Increased risk of constipation with opioid	3
Key Question 5: Postoperative pain										

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Opioid vs. NSAID, single dose, various surgeries	Pain, rescue medication use	<1 d	2	421	Direct	Imprecise	Fair	Consistent	No differences	3
Opioid vs. NSAID, multidose course, various surgeries	Pain	1 d to <1 w	4	830	Direct	Imprecise	Fair	Inconsistent	Unable to determine	Insufficient
	Rescue medication use	1 d to <1 w	4	860	Direct	Imprecise	Fair	Consistent	RR 1.22 to 2.04	2
Opioid vs. acetaminophen, single dose, cesarean section	Pain, re-medication	<1 d	1	96	Direct	Imprecise	Fair	Unable to assess	No difference	3
Opioid vs. acetaminophen, multidose course, various surgeries	Study wthdrawal due to AEs	<1 d, 1 d to <1 w	3	252	Direct	Imprecise	Fair	Consistent	Increased risk with opioid	3
NSAID vs. acetaminophen, single dose, various surgeries	Pain, rescue medication use	<1 d	2	113	Direct	Imprecise	Fair	Inconsistent	Unable to determine	Insufficient
Acupuncture vs. sham, various surgeries	Pain	1 d to <1 w	2	106	Direct	Imprecise	Fair	Inconsistent	Unable to determine	Insufficient
Acupressure vs. sham, knee surgeries	Pain	<1 d, 1 d to <1 w	2	130	Direct	Imprecise	Fair	Consistent	Unable to determine	Insufficient
	Pain medication use	<1 d, 1 d to <1 w	2	130	Direct	Imprecise	Fair	Consistent	Decreased with acupuncture	3

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Cold therapy vs. sham or usual care, knee surgeries	Pain intensity	<1 w	3	168	Direct	Imprecise	Fair	Consistent	No differences	3
	Pain intensity; function, QoL	2 to <4 w, ≥4 w	1	60	Direct	Imprecise	Good	Unable to assess	No differences	3
	Pain medication use	<1 w	1	60	Direct	Imprecise	Good	Unable to assess	Decreased with cold therapy	3
Massage vs. no massage, various surgeries	Pain intensity, decreased pain medication use, anxiety	<1 d	2 to 5	733	Direct	Precise	Poor	Consistent	Moderate to large decrease with massage at <1 day, decreased pain medication use, and decreased anxiety	3
Music therapy vs. no music therapy, various surgeries	Pain	<1 d, 1 d to <1 w	2	148	Direct	Imprecise	Fair	Consistent	Small to moderate decrease in pain intensity	3
Exercise vs. no exercise, thyroid surgery	Function	1 to <2 w, ≥4 w	1	80	Direct	Imprecise	Fair	Unable to assess	Large decrease with exercise at 1 week, no difference at 1 month	3
TENS vs. sham TENS, liposuction	Pain intensity, analgesic use	<1 d, 1 d to <1 w	1	42	Direct	Imprecise	Fair	Unable to assess	Moderate to large decrease in pain intensity and decreased analgesic use with TENS	3

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Key Question 6. Dental pain										
Opioid + acetaminophen vs. acetaminophen, single dose	Pain	<1 d	11	828	Direct	Precise	Fair	Inconsistent	Inconsistent effects on pain intensity, but larger sum of pain intensity differences with opioid	2 (for sum of pain intensity differences)
	Rescue or repeat medication use	<1 d	7	484	Direct	Precise	Fair	Consistent	RR 0.81, 95% CI 0.56 to 0.97	2
Opioid vs. acetaminophen, single dose	Pain, rescue medication use	<1 d	2	149	Direct	Imprecise	Fair	Consistent	No differences	3
Opioid (with or without acetaminophen) vs. acetaminophen, single dose	Any AE, nausea, drowsiness, dizziness	<1 d	4 to 8	445 to 769	Direct	Imprecise	Fair	Consistent	Increased risk with opioid	3
Opioid plus acetaminophen or NSAID vs. NSAID, single dose	Pain, rescue or repeat medication use	<1 d	8 to 12	926 to 2,021	Direct	Precise	Fair	Inconsistent (pain intensity); consistent (rescue or repeat medication use)	Small to moderate increase in pain intensity with opioids, increased likelihood of rescue or repeat medication use (RR 1.35, 95% CI 1.23 to 1.48)	3 for pain; 2 for rescue or repeat medication use
Opioid vs. acetaminophen, multidose course	Pain intensity	1 d to <1 w	1	20	Direct	Imprecise	Fair	Consistent	No difference	3

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
Opioid (with or without acetaminophen) vs. NSAID	Any AE, nausea, dizziness, drowsiness	<1 d, 1 d to <1 w	9 to 12	1,959 to 2,784	Direct	Precise	Fair	Consistent	RR 1.72 (95% CI 1.29 to 2.28) for any AE, 2.72 (95% CI 1.84 to 4.01) for nausea, 2.97 (95% CI 1.59 to 5.54) for dizziness, and 1.76 (95% CI 1.00 to 3.10) for drowsiness	2
NSAID vs. acetaminophen, single dose	Pain intensity, rescue or repeat medication use	<1 d	11 to 15	2,014 to 2,506	Direct	Precise	Fair	Consistent	Moderate to large decrease in pain with NSAID, decreased likelihood of rescue or repeat medication use (RR 0.64, 95% CI 0.58 to 0.71)	2
	Any AE	<1 d	12	2,512	Direct	Precise	Fair	Consistent	RR 0.85 (95% CI 0.72 to 1.00)	2
Key Question 7: Kidney stone pain										
Morphine vs. NSAID, single dose	Pain, rescue medication use	<1 d	1	1,097	Direct	Precise	Good	Unable to assess	Increased likelihood of pain, and rescue medication use with morphine	2

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
	Any AE	<1 d	1	1,097	Direct	Precise	Good	Unable to assess	3% vs. 1%, RR 2.70 (95% CI 1.15 to 6.38)	2
Meperidine vs. NSAID, single dose	Pain, rescue medication use	<1 d	4 to 6	475 to 671	Direct	Precise	Fair	Inconsistent	Moderate to large increase in pain intensity with meperidine, increased likelihood of rescue medication use	2
	Any AE, somnolence, nausea	<1 d	4 to 5	471 to 573	Direct	Imprecise	Fair	Inconsistent	RR 1.71 (95% CI 0.99 to 2.96) for any AE, RR 1.98 (95% CI 0.82 to 4.79) for somnolence, and RR 1.84 (95% CI 1.02 to 3.31) for nausea	3
Morphine vs. acetaminophen, single dose	Pain, rescue medication use	<1 d	1	1,096	Direct	Precise	Good	Unable to assess	Increased likelihood of pain with morphine, similar rescue medication use	2
	Any AE	<1 d	1	1,096	Direct	Precise	Good	Unable to assess	3% vs. 1%, RR 2.71 (95% CI 1.15 to 6.39)	2

Intervention	Outcomes	Timing of Outcomes	Number of Studies	Number of Subjects	Directness	Precision	Quality	Consistency	Findings	Evidence Type
NSAID vs. acetaminophen, single dose	Pain, rescue medication use	<1 d	2 to 3	1,145 to 1,225	Direct	Imprecise (for pain)	Fair	Inconsistent (pain)	Inconsistent effects on pain; decreased likelihood of rescue medication use with NSAID	3 for rescue medication use, insufficient for pain
Acupuncture vs. NSAID or acetaminophen	Pain	<1 d	1	160	Direct	Imprecise	Fair	Unable to assess	Moderate increase in pain with acupuncture	3
Key Question 8: Sickle cell pain										
Insufficient evidence	--	--	--	--	--	--	--	--	--	--

Abbreviations: AE = adverse event; CI = confidence interval; D = day; NSAID = nonsteroidal anti-inflammatory drug; QoL = quality of life; RR = relative risk; SOE = strength of evidence; TENS = transcutaneous electrical nerve stimulation; W = week

Treatments for acute episodic migraine. This table is based on Singh RBH, VanderPluym JH, Morrow AS, et al. AHRQ Comparative Effectiveness Reviews. Acute Treatments for Episodic Migraine. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020. The strength of evidence ratings in the AHRQ report were converted to ACIP-adapted GRADE evidence type ratings.

Opioid therapy

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Any opioid vs. any nonopioid	Pain free	2 hours	RR 0.88 (95% CI 0.65 to 1.20)	1 comparative observational study (161)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 1.34 (95% CI 0.82 to 2.18)	1 comparative observational study (161)	High risk of bias and severe imprecision	Insufficient
Butorphanol vs. placebo	Pain free	2 hours	RR 2.90 (95% CI 1.20 to 7.01)	1 RCT (157)	High risk of bias and imprecision	3
	Pain free	1 day	RR 1.83 (95% CI 1.10 to 3.05)	1 RCT (157)	High risk of bias and imprecision	3
	Pain free	1 week	RR 2.08 (95% CI 1.27 to 3.43)	1 RCT (157)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 3.37 (95% CI 1.83 to 6.22)	1 RCT (157)	High risk of bias and imprecision	3
	Pain relief	1 day	RR 2.07 (95% CI 1.43 to 2.98)	1 RCT (157)	High risk of bias and imprecision	3
	Pain relief	2 week	RR 2.09 (95% CI 1.45 to 3.02)	1 RCT (157)	High risk of bias and imprecision	3
Hydromorphone vs. diphenhydramine	Pain free	2 hours	RR 0.54 (95% CI 0.33 to 0.90)	1 RCT (127)	High risk of bias and imprecision	3

plus prochlorperazine						
	Restored function	2 hours	RR 0.45 (95% CI 0.27 to 0.74)	1 RCT (127)	High risk of bias and imprecision	3
	Restored function	1 week	RR 0.80 (95% CI 0.61 to 1.06)	1 RCT (127)	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 week	RR 0.53 (95% CI 0.35 to 0.81)	1 RCT (127)	High risk of bias and imprecision	3
Hydromorphone vs. metoclopramide	Pain scale	2 hours	SMD -0.56 (95% CI -0.90 to -0.21)	1 comparative observational study (200)	High risk of bias and imprecision	3
	Pain scale	1 day	SMD -0.32 (95% CI -0.66 to 0.03)	1 comparative observational study (200)	High risk of bias and imprecision	Insufficient
Meperidine plus dimenhydrinate vs. chlorpromazine	Pain relief	2 hours	RR 0.65 (95% CI 0.36 to 1.18)	1 RCT (46)	High risk of bias and imprecision	Insufficient
	Pain scale	2 hours	SMD -1.09 (95% CI -1.71 to -0.47)	1 RCT (46)	High risk of bias and imprecision	3
Meperidine plus hydroxyzine vs. dihydroergotamine plus metoclopramide	Pain relief	2 hours	RR 0.23 (95% CI 0.08 to 0.64)	1 RCT (28)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.06 (95% CI -0.24 to 0.36)	1 RCT (170)	High risk of bias and severe imprecision	Insufficient

	Restored function	1 day	RR 0.44 (95% CI 0.24 to 0.82)	1 RCT (170)	High risk of bias and imprecision	3
	Serious AE	N/A	Rate ratio 1.00 (95% CI 0.02 to 50.40)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient
Meperidine vs. droperidol	Pain scale	2 hours	P=0.33	1 RCT (29)	High risk of bias and severe imprecision	Insufficient
Methotrimeprazine vs. dimenhydrinate plus meperidine	Pain scale	2 hours	SMD 0.26 (95% CI -0.20 to 0.72)	1 RCT (74)	High risk of bias and severe imprecision	Insufficient
Morphine vs. intravenous dexamethasone	Pain scale	2 hours	SMD -0.35 (95% CI -0.64 to -0.06)	1 RCT (190)	High risk of bias and imprecision	3
	Pain scale	1 day	SMD -0.38 (95% CI -0.66 to -0.09)	1 RCT (190)	High risk of bias and imprecision	3
Tramadol vs. placebo	Pain free	2 hours	RR 2.50 (95% CI 0.56 to 11.16)	1 RCT (34)	Moderate risk of bias and imprecision	Insufficient
	Pain relief	2 hours	RR 2.00 (95% CI 0.98 to 4.08)	1 RCT (34)	Moderate risk of bias and imprecision	Insufficient
	Pain scale	2 hours	SMD -0.25 (95% CI -0.43 to 0.92)	1 RCT (34)	Moderate risk of bias and imprecision	Insufficient
Tramadol plus acetaminophen vs. placebo	Pain free	2 hours	RR 2.42 (95% CI 1.34 to 4.35)	1 RCT (375)	High risk of bias and imprecision	3

	Pain free	1 day	RR 1.43 (95% CI 1.09 to 1.88)	1 RCT (375)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 1.68 (95% CI 1.27 to 2.22)	1 RCT (375)	High risk of bias and imprecision	3
	Pain relief	1 day	RR 1.75 (95% CI 1.35 to 2.25)	1 RCT (375)	High risk of bias and imprecision	3
	Serious AE	N/A	Rate ratio 0.99 (95% CI 0.02 to 50.13)	1 RCT (375)	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 day	RR 2.26 (95% CI 1.15 to 4.46)	1 RCT (375)	High risk of bias and imprecision	3
	Sustained pain relief	1 day	RR 1.56 (95% CI 1.08 to 2.272)	1 RCT (375)	High risk of bias and imprecision	3

AE=adverse event; CI=confidence interval; N/A=not applicable; RCT=randomized controlled trial; RR=relative risk; SMD=standardized mean difference
SMD>0 indicates the intervention mentioned first in the comparison is better

Systematic reviews of triptans compared with placebo

Outcome	Conclusion	Evidence Type
Pain	Improvement in pain resolution at 2 hours and 1 day	1[a]
Adverse events[b]	Increased risk of mild and transient adverse events	1

Evidence base: 186 randomized controlled trials summarized in 9 systematic reviews (101,276 patients). The most studied triptan is sumatriptan, followed by zolmitriptan, eletriptan, naratriptan, almotriptan, rizatriptan, and frovatriptan.

[a]Some older trials do not report methods of allocation concealment. However, this concern was not sufficient to rate down the evidence type particularly in the presence of a large relative effect (relative risk >2)

[b]The number of events is small, particularly for adverse events analyses

Systematic reviews of nonsteroidal anti-inflammatory drugs compared with placebo

Outcome	Conclusion	Evidence Type
Pain	Improvement in pain resolution at 2 hours and 1 day	2[a]

Adverse events	Increased risk of mild and transient adverse events	2[a, b]

Evidence base: 5 systematic reviews (13,124 patients). The most studies NSAID is ibuprofen (9 randomized controlled trials, 4,373 patients), followed by diclofenac and ketorolac.

[a]Some older trials do not report the methods of allocation concealment. However, this concern was not sufficient to rate down strength of evidence particularly in the presence of a large relative effect (relative risk >2)

[b]The number of events is small, particularly for adverse events analyses

Ergot alkaloids

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Dihydroergotamine vs. chlorpromazine	Pain free	2 hours	RR 0.69 (95% CI 0.28 to 1.70)	1 RCT (50)	High risk of bias and severe imprecision	Insufficient
Dihydroergotamine vs. lidocaine	Pain free	2 hours	RR 3.03 (95% CI 0.67 to 14.29)	1 RCT (50)		Insufficient
Dihydroergotamine vs. placebo	Pain free	2 hours	RR 2.89 (95% CI 2.07 to 4.03); I2=0%	2 RCTs (989)	High risk of bias	2
	Pain free	1 day	RR 1.74 (95% CI 1.43 to 2.12)	1 RCT (903)	Possible imprecision, single trial	2
	Pain free	1 week	RR 1.54 (95% CI 1.25 to 1.89)	1 RCT (903)	Possible imprecision, single trial	2
	Pain relief	2 hours	RR 1.83 (95% CI 1.58 to 2.13); I2=0%	3 RCTs (1,299)	N/A	1
	Pain relief	1 day	RR 1.79 (95% CI 1.54 to 2.08), I2=0%	2 RCTs (1,213)	N/A	1
	Pain relief	1 week	RR 1.48 (95% CI 1.22 to 1.80	1 RCT (903)	Possible imprecision, single trial	2
	Pain scale	2 hours	SMD -0.14 (95% CI -0.82 to 0.53)	1 RCT (34)	High risk of bias and severe imprecision	Insufficient

	Restored function	2 hours	RR 2.38 (95% CI 1.44 to 3.94)	1 RCT (348)	Imprecision	2
	Restored function	1 day	RR 2.80 (95% CI 1.82 to 4.40	1 RCT (348)	Imprecision	2
	Serious AE	N/A	Rate ratio 0.69 (95% CI -0.03 to 16.62); I2=0%	4 RCTs	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 day	RR 3.51 (95% CI 2.33 to 5.28); I2=0%	2 RCTs (989)	N/A	1
	Sustained pain free	1 week	RR 2.96 (95% CI 1.90 to 4.62); I2=0%	2 RCTs (989)	N/A	1
	Sustained pain relief	1 day	RR 2.23 (95% CI 1.76 to 2.81)	2 RCTs (989)	N/A	1
	Sustained pain relief	1 week	RR 2.11 (95% CI 1.62 to 2.76)	2 RCTs (989)	N/A	1
Ergotamine plus caffeine vs. placebo	Improved function	2 hours	RR 1.38 (95% CI 0.91 to 2.10)	1 RCT (309)	Severe imprecision	3
	Pain free	2 hours	RR 2.08 (95% CI 0.81 to 5.40)	1 RCT (309)	Severe imprecision	3
	Pain relief	2 hours	RR 1.61 (95 5 CI 1.05 to 2.49)	1 RCT (309)	Imprecision	2
	Pain scale	2 hours	SMD 0.01 (95% CI -1.01 to 1.02)	1 RCT (15)	High risk of bias and severe imprecision	Insufficient
Ergotamine plus caffeind vs. prochlorperazine	Pain scale	2 hours	SMD -0.58 (95% CI -1.46 to 0.28)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better

Antiemetics

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Chlorpromazine vs. placebo	Improved function	2 hours	RR 2.01 (95% CI 0.76 to 5.36)	1 RCT (36)	High risk of bias and imprecision	3
	Pain free	2 hours	RR 7.25 (95% CI 3.20 to 16.42); I2=0%	2 RCTs (123)	High risk of bias and imprecision	3
	Pain free	1 day	RR 1.37 (95% CI 1.09 to 1.74); I2=17%	2 RCTs (123)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 5.46 (95% CI 2.97 to 10.05); I2=0%	2 RCTs (123)	High risk of bias and imprecision	3
	Pain relief	1 day	RR 1.22 (95% CI 1.02 to 1.47); I2=0%	2 RCTs (123)	High risk of bias and imprecision	3
Diphenhydramine plus metoclopramide vs. diphenhydramine plus haloperidol	Pain scale	2 hours	SMD -0.41 (95% CI -0.90 to 0.08)	1 RCT (64)	High risk of bias and severe imprecision	Insufficient
Droperidol vs. placebo	Pain free	2 hours	RR 1.60 (95% CI 1.06 to 2.41)	1 RCT (305)	High risk of bias and imprecision	3
Granisetron vs. placebo	Pain free	2 hours	RR 1.29 (95% CI 0.06 to 28.65)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient
	Pain scale	2 hours	SMD 1.10 (95% CI 0.23 to 1.97)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 0.40 (95% CI 0.01 to 20.16)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient
Haloperidol vs. placebo	Pain relief	2 hours	RR 5.33 (95% CI 1.84 to 15.49)	1 RCT (40)	High risk of bias and imprecision	3
Magnesium sulfate vs. dexamethasone	Pain scale	2 hours	SMD 0.82 (95% CI 0.33 to 1.31)	1 RCT (70)	High risk of bias and imprecision	3

plus metoclopramide						
Metoclopramide vs. chlorpromazine	Pain relief	2 hours	RR 0.98 (95% CI 0.48 to 1.99)	1 RCT (91)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 0.84 (95% CI 0.65 to 1.09)	1 RCT (91)	High risk of bias and severe imprecision	Insufficient
	Pain scale	2 hours	SMD -0.20 (95% CI -0.61 to 0.21)	1 RCT (91)	High risk of bias and severe imprecision	Insufficient
Metoclopramide vs. diphenhydramine plus metoclopramide	Pain scale	2 hours	SMD -0.26 (95% CI -0.54 to 0.01)	1 RCT (208)	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 week	RR 0.82 (95% CI 0.42 to 1.48)	1 RCT (208)	High risk of bias and severe imprecision	Insufficient
	Sustained pain relief	1 week	RR 0.95 (95% CI 0.67 to 1.35)	1 RCT (208)	Severe imprecision	3
Metoclopramide vs. granisetron	Pain scale	2 hours	SMD -1.10 (95% CI -1.44 to -0.75)	1 RCT (148)	High risk of bias and imprecision	3
	Pain scale	1 day	SMD -0.41 (95% CI -0.74 to -0.09)	1 RCT (148)	High risk of bias and imprecision	3
Metoclopramide vs. magnesium sulfate plus metoclopramide	Pain relief	2 hours	RR 1.34 (95% CI 1.01 to 1.78)	1 RCT (44)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.4 (95% CI -0.06 to 1.15)	1 RCT (44)	High risk of bias and severe imprecision	Insufficient
	Restored function	2 hours	RR 1.94 (95% CI 1.07 to 3.52)	1 RCT (44)	High risk of bias and imprecision	3
Metoclopramide vs. placebo	Pain free	2 hours	RR 2.00 (95% CI 0.40 to 10.08)	1 RCT (86)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 1.91 (95% CI 1.47 to 2.48); I2=67%	3 RCTs (268)	High risk of bias and imprecision	3

	Pain scale	2 hours	SMD -0.12 (95% CI -0.40 to 0.17); I2=90%	2 RCTs (198)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 1.08 (95% CI 0.02 to 54.60)	1 RCT (50)	High risk of bias and severe imprecision	Insufficient
Prochloroperazine vs. ergotamine plus caffeine	Pain scale	2 hours	SMD 0.58 (95% CI -0.28 to 1.46)	1 RCT (28)	High risk of bias and severe imprecision	Insufficient
Prochlorperazine vs. metoclopramide	Pain free	2 hours	RR 1.56 (95% CI 1.00 to 2.45); I2=0%	2 RCTs (163)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 0.89 (95% CI 0.72 to 1.10); I2=0.8%	2 RCTs (147)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.18 (95% CI -0.27 to 0.63)	1 RCT (77)	High risk of bias and severe imprecision	Insufficient
	Pain scale	1 day	SMD 0.29 (95% CI -0.16 to 0.74)	1 RCT (77)	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 day	RR 1.46 (95% CI 0.45 to 4.77)	1 RCT (77)	High risk of bias and severe imprecision	Insufficient
	Sustained pain relief	1 day	RR 1.26 (95% CI 0.81 to 1.97)	1 RCT (77)	High risk of bias and severe imprecision	Insufficient
Prochloperazine vs. octreotide	Pain relief	2 hours	RR 1.66 (95% CI 1.12 to 2.47)	1 RCT (44)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.84 (95% CI 0.22 to 1.46)	1 RCT (44)	High risk of bias and severe imprecision	Insufficient
Prochlorperazine vs. placebo	Pain free	2 hours	RR 4.66 (95% CI 1.10 to 19.70)	1 RCT (86)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 1.80 (95% CI 1.10 to 2.94); I2=0%	2 RCTs (90)	High rhisk of bias and imprecision	3
	Pain scale	2 hours	SMD 1.29 (95% CI 0.58 to 2.01); I2=91%	2 RCTs (49)	High risk of bias and inconsistency	3

Valproate vs. prochlorperazine	Pain scale	2 hours	SMD -1.38 (95% CI -2.07 to -0.69)	1 RCT (40)	High risk of bias and imprecision	3

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better

Calcitonin gene-related peptide receptor antagonists (gepants)

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Rimegepant vs. placebo	Pain free	2 hours	RR 1.80 (95% CI 1.52 to 2.13); I2=0%	3 RCTs (3,336)	High risk of bias	2
	Pain free	1 day	RR 1.52 (95% CI 1.33 to 1.74)	1 RCT (1,186)	High risk of bias	2
	Pain relief	2 hours	RR 1.36 (95% CI 1.26 to 1.46); I2=0%	3 RCTs (3,336)	High risk of bias	2
	Restored function	2 hours	RR 1.43 (95% CI 1.26 to 1.62); I2=0%	2 RCTs (2,652)	High risk of bias	2
	Serious AE	N/A	Rate ratio 0.54 (95% CI 0.13 to 2.28); I2=0%	3 RCTs (3,336)	High risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 day	RR 2.24 (95% CI 1.65 to 3.05); I2=71%	2 RCTs (1,870)	High risk of bias	2
	Sustained pain free	1 week	RR 2.23 (955 cI 1.60 to 3.09); I2=71%	2 RCTs (1,870)	High risk of bias	2
	Sustained pain relief	1 day	RR 1.65 (95% CI 1.47 to 1.85); I2=0%	2 RCTs (2,150)	High risk of bias	2
	Sustained pain relief	1 week	RR 1.64 (95% CI 1.40 to 1.93)	1 RCT (1,466)	High risk of bias	2
	Sustained restored function	1 day	RR 1.73 (95% CI 1.41 to 2.12)	1 RCT (1,466)	High risk of bias	2
	Sustained restored function	1 week	RR 1.66 (95% CI 1.33 to 2.07)	1 RCT (1,466)	High risk of bias	2

Ubrogepant vs. placebo	Improved function	2 hours	RR 1.26 (95% CI 1.12 to 1.42); I2=0%	2 RCTs (3,358)	N/A	1
	Improved function	1 day	RR 1.16 (95% CI 1.09 to 1.24); I2=0%	2 RCTs (3,358)	N/A	1
	Pain free	2 hours	RR 1.58 (95% CI 1.31 to 1.90); I2=0%	3 RCTs (4,192)	N/A	1
	Pain relief	2 hours	RR 1.21 (95% CI 1.12 to 1.31); I2=0%	3 RCTs (4,192)	N/A	1
	Pain relief	1 day	RR 1.63 (95% CI 1.33 to 2.01)	1 RCT (1,686)	N/A	1
	Sustained pain free	1 day	RR 1.63 (95% CI 1.29 to 2.07); I2=0%	3 RCTs (4,192)	N/A	1
	Sustained pain free	1 week	RR 1.89 (95% CI 0.88 to 4.02)	1 RCT (834)	Severe imprecision	3
	Sustained pain relief	1 day	RR 1.55 (95% CI 1.30 to 1.85)	2 RCTs (2,506)	Inconsistency	2
	Sustained pain relief	1 week	RR 1.29 (95% CI 0.91 to 1.84)	1 RCT (833)	Severe imprecision	3
	Restored function	2 hours	RR 1.27 (95% CI 1.13 to 1.42); I2=0%	2 RCTs (3,358)	N/A	1
	Restored function	1 day	RR 1.17 (95% CI 1.09 to 1.25); I2=0%	2 RCTs (3,358)	N/A	1
	Satisfied with pain relief	2 hours	RR 1.43 (95% CI 1.24 to 1.64); I2=0%	2 RCTs (3,358)	N/A	1
	Satisfied with pain relief	1 day	RR 1.55 (95% CI 1.39 to 1.72); I2=31%	2 RCTs (3,358)	N/A	1
	Serious AE	N/A	Rate ratio 2.54 (95% CI 0.28 to 23.11)	2 RCTs (3,358)	Severe imprecision	3

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better

5-HT1F receptor agonists (ditans)

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Lasmiditan vs. placebo	Function scale	2 hours	SMD 3.34 (95% CI 3.04 to 3.64)	1 RCT (512)	N/A	1
	Pain free	2 hours	RR 1.67 (95% CI 1.25 to 2.24); I2=21%	4 RCTs (5,742)	N/A	1
	Pain relief	2 hours	RR 1.38 (95% CI 1.14 to 1.68); I2=34%	4 RCTs (5,742)	N/A	1
	Pain scale	2 hours	SMD 2.68 (95% CI 2.41 to 2.95)	1 RCT (512)	Possible imprecision, single trial	2
	Restored function	2 hours	RR 1.42 (95% CI 1.26 to 1.61); I2=0%	2 RCTs (5,100)	N/A	1
	Serious AE	N/A	Rate ratio 4.05 (95% CI 1.75 to 9.41); I2=33%	2 RCTs (2,743)	N/A	1
	Sustained pain free	1 day	RR 1.38 (95% CI 1.10 to 1.72); I2=33%	2 RCTs (2,999)	N/A	1
	Sustained pain free	1 week	RR 1.38 (95% CI 1.07 to 1.78)	1 RCT (2,869)	Possible imprecision, single trial	2
	Sustained pain relief	1 day	RR 1.76 (95% CI 1.08 to 2.87)	1 RCT (130)	Imprecision	2

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better

Other nonopioid pharmacological interventions

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Acetaminophen vs. placebo	Function scale	2 hours	SMD 0.38 (95% CI 0.18 to 0.59)	1 RCT (378)	Imprecision	2
	Pain free	2 hours	RR 1.89 (95% CI 1.24 to 2.86); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Pain free	1 day	RR 1.78 (95% CI 1.38 to 2.30); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Pain relief	2 hours	RR 1.61 (95% CI 1.33 to 1.95); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Pain relief	1 day	RR 1.71 (95% CI 1.42 to 2.04); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Pain scale	2 hours	SMD 0.39 (95% CI 0.25 to 0.54); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Pain scale	1 day	SMD 0.31 (95% CI 0.10 to 0.52)	1 RCT (351)	Moderate risk of bias	2
	Restored function	1 day	RR 1.75 (95% CI 1.41 to 2.17); I2=0%	2 RCTs (729)	Moderate risk of bias	2
	Serious AE	N/A	Rate ratio 0.99 (95% CI 0.06 to 15.86); I2=0%	2 RCTs (729)	Moderate risk of bias and severe imprecision	Insufficient
Dexamethasone vs. placebo	Pain free	2 hours	RR 1.09 (95% CI 0.83 to 1.44)	1 RCT (205)	Severe imprecision	3
	Restored function	2 hours	RR 0.87 (95% CI 0.73 to 1.04)	1 RCT (205)	Severe imprecision	3
	Restored function	1 day	RR 1.12 (95% CI 0.89 to 1.40)	1 RCT (205)	Severe imprecision	3
	Restored function	1 week	RR 1.49 (95% CI 1.04 to 2.13)	1 RCT (115)	Moderate risk of bias and imprecision	3
	Sustained pain free	1 day	RR 1.23 (95% CI 0.72 to 2.09)	1 RCT (205)	Severe imprecision	3

Dipyrone vs. placebo	Pain free	2 hours	RR 7.14 (95% CI 3.02 to 16.86)	1 RCT (134)	High risk of bias and imprecision	3
	Pain free	1 day	RR 1.28 (95% Ci 1.01 to 1.63)	1 RCT (134)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 4.32 (95% CI 2.31 to 8.08)	1 RCT (134)	High risk of bias and imprecision	3
	Pain relief	1 day	RR 1.09 (95% CI 0.90 to 1.33)	1 RCT (134)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 0.47 (95% CI 0.01 to 23.66)	1 RCT (72)	High risk of bias and severe imprecision	Insufficient
Greater occipital nerve block vs. sham injection	Pain free	2 hours	RR 10.29 (95% CI 0.61 to 174.70)	1 RCT (28)	Moderate risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 2.08 (95% CI 0.93 to 4.63)	1 RCT (28)	Moderate risk of bias and severe imprecision	Insufficient
	Pain scale	2 hours	SMD 0.74 (95% CI -0.03 to 1.51)	1 RCT (28)	Moderate risk of bias and severe imprecision	Insufficient
Ketamine vs. placebo	Function scale	2 hours	SMD 0.23 (95% CI -0.44 to 0.91)	1 RCT (34)	Moderate risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 0.75 (95% CI 0.14 to 3.94)	1 RCT (34)	Moderate risk of bias and severe imprecision	Insufficient
	Pain scale	2 hours	SMD -0.43 (95% CI -1.11 to 0.25)	1 RCT (34)	Moderate risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 1.13 (95% CI 0.02 to 56.70)	1 RCT (34)	Moderate risk of bias and severe imprecision	Insufficient
Lidocaine vs. chlorpromazine	Pain free	2 hours	RR 0.23 (95% CI 0.05 to 0.98)	1 RCT (50)	High risk of bias and imprecision	3
Lidocaine vs. placebo	Function scale	2 hours	SMD 0.39 (95% CI -0.07 to 0.86)	1 RCT (81)	Moderate risk of bias and severe imprecision	Insufficient

	Pain free	1 week	RR 1.45 (95% CI 0.93 to 2.27)	1 RCT (162)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 2.14 (95% CI 1.16 to 3.96); I2=65%	2 RCTs (130)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.02 (95% CI -0.21 to 0.26); I2=85%	3 RCTs 9292)	Moderate risk of bias, inconsistency, and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 1.00 (95% CI 0.02 to 50.40)	1 RCT (162)	High risk of bias and severe imprecision	Insufficient
Magnesium sulfate vs. caffeine citrate	Pain scale	2 hours	SMD 1.62 (95% CI 1.08 to 2.17)	1 comparative observational study (70)	High risk of bias and imprecision	3
Magnesium sulfate vs. placebo	Pain free	2 hours	RR 5.73 (95% CI 2.43 to 13.50); I2=55%	1 RCT and 1 crossover RCT (150)	High risk of bias and imprecision	3
	Pain free	1 day	RR 1.25 (95% CI 0.97 to 1.61)	1 RCT (120)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 3.86 (95% CI 2.11 to 7.07); I2=60%	1 RCT and 1 crossover RCT (150)	High risk of bias and imprecision	3
	Pain relief	1 day	RR 1.14 (95% CI 0.93 to 1.39)	1 RCT (120)	High risk of bias and severe imprecision	Insufficient
Octreotide vs. placebo	Pain relief	1 day	RR 3.06 (95% CI 1.11 to 8.44)	1 RCT (29)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 1.09 (95% CI 0.30 to 1.88)	1 RCT (29)	High risk of bias and imprecision	3
	Pain scale	1 day	SMD 1.51 (95% CI 0.67 to 2.35)	1 RCT (29)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 1.15 (95% CI 0.02 to 57.96)	1 RCT (43)	High risk of bias and severe imprecision	Insufficient

Propofol vs. standard therapy (chlorpromazine, metoclopramide, ondansetron, lignocaine, magnesium sulphate, or morphine)	Pain scale	2 hours	SMD 0.00 (95% CI -0.72 to 0.72)	1 RCT (30)	Moderate risk of bias and imprecision	Insufficient
	Pain scale	1 day	SMD 0.53 (95% CI -0.18 to 1.28)	1 RCT (30)	Moderate risk of bias and severe imprecision	Insufficient
Propofol vs. dexamethasone	Pain scale	2 hours	SMD 1.01 (95% CI 0.58 to 1.45)	1 RCT (90)	Moderate risk of bias and imprecision	3
Secobarbital vs. placebo	Pain relief	1 day	RR 1.88 (95% CI 1.09 to 3.21)	1 RCT (30)	High risk of bias and imprecision	3
	Pain scale	1 day	SMD 0.79 (95% CI 0.04 to 1.53)	1 RCT (30)	High risk of bias and severe imprecision	Insufficient
Valproate vs. dexamethasone	Pain free	1 day	RR 1.25 (95% CI 0.39 to 3.99)	1 RCT (40)	High risk of bias and severe imprecision	Insufficient
	Pain relief	2 hours	RR 0.83 (95% CI 0.68 to 1.02)	1 RCT (80)	Severe imprecision	3
	Pain relief	1 day	RR 0.92 (95% CI 0.82 to 1.04)	1 RCT (80)	Severe imprecision	3
	Pain scale	2 hours	SMD -0.16 (95% CI -0.46 to 0.15); I2=0%	2 RCTs (166)	High risk of bias and severe imprecision	Insufficient
	Pain scale	1 day	SMD -0.15 (95% CI -0.51 to 0.22); I2=74%	2 RCTs (120)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 1.00 (95% CI 0.02 to 50.40)	1 RCT (86)	High risk of bias and severe imprecision	Insufficient
Valproate vs. prochlorperazone	Pain scale	2 hours	SMD -1.38 (95% CI -2.07 to -0.69)	1 RCT (40)	Imprecision	2

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better

Nonpharmacological therapies

Comparison	Outcome	Time	Findings	Study Design and Sample Size	Rationale for Evidence Type	Evidence Type
Acupuncture vs. sham acupuncture	Pain free	1 day	RR 2.53 (95% CI 1.27 to 5.02)	1 RCT (175)	Moderate risk of bias and imprecision	3
	Pain relief	1 day	RR 0.74 (95% CI 0.56 to 0.97)	1 RCT (175)	Moderate risk of bias and imprecision	3
	Pain scale	2 hours	SMD 0.19 (95% CI -0.10 to 0.49); I2=78%	2 RCTs (235)	Moderate risk of bias and severe imprecision	Insufficient
	Pain scale	1 day	SMD 0.49 (95% CI 0.25 to 0.73); I2=0%	2 RCTs (325)	Moderate risk of bias and imprecision	3
	Sustained pain free	1 day	RR 2.14 (95% CI 0.93 to 4.95)	1 RCT (150)	Moderate risk of bias and severe imprecision	Insufficient
	Sustained pain free	1 week	RR 1.12 (95% CI 0.96 to 1.32)	1 RCT (150)	Moderate risk of bias and severe imprecision	Insufficient
	Serious AEs	N/A	RR 1.03 (95% CI 0.02 to 52.13)	1 RCT (175)	Moderate risk of bias and severe imprecision	Insufficient
Chamomile vs. placebo	Pain scale	2 hours	SMD 1.51 (95% CI 1.07 to 1.96)	1 RCT (98)	Moderate risk of bias and imprecision	3
	Pain scale	1 day	SMD 1.16 (95% CI 0.74 to 1.58)	1 RCT (98)	Moderate risk of bias and imprecision	3
Eye movement desensitization reprocessing vs. standard care	Pain free	2 hours	RR 17.00 (95% CI 2.44 to 118.55)	1 RCT (52)	High risk of bias and imprecision	3
	Pain scale	2 hours	SMD 2.28 (95% CI 1.58 to 2.99)	1 RCT (52)	High risk of bias and imprecision	3

	Pain scale	1 day	SMD 0.60 (95% CI 0.04 to 1.16)	1 RCT (52)	High risk of bias and severe imprecision	Insufficient
	Pain scale	1 week	SMD 0.52 (95% CI -0.03 to 1.08)	1 RCT (52)	High risk of bias and severe imprecision	Insufficient
External trigeminal nerve stimulation vs. sham	Pain free	2 hours	RR 2.34 (95% CI 0.77 to 7.12)	1 RCT (106)	Severe imprecision	3
	Pain free	1 day	RR 2.23 (95% CI 0.99 to 5.01)	1 RCT (106)	Severe imprecision	3
	Pain relief	2 hours	RR 1.32 (95% CI 0.88 to 1.99)	1 RCT (106)	Severe imprecision	3
	Pain relief	1 day	RR 1.24 (95% CI 0.87 to 1.77)	1 RCT (106)	Severe imprecision	3
	Pain scale	2 hours	SMD 1.25 (95% CI 0.90 to 1.60); I2=99%	2 RCTs (189)	Moderate risk of bias and imprecision	3
	Pain scale	1 day	SMD 0.53 (95% CI 0.14 to 0.92)	1 RCT (106)	Imprecision	2
	Serious AE	N/A	Rate ratio 1.04 (95% CI 0.02 to 52.34)	1 RCT (106)	Severe imprecision	3
	Sustained pain free	1 day	RR 7.26 (95% CI 0.38 to 137.28)	1 RCT (106)	Severe imprecision	3
	Sustained pain relief	1 day	RR 1.95 (95% CI 0.90 to 4.20)	1 RCT (106)	Severe imprecision	3
Magnetic stimulation vs. sham stimulation	Function scale	1 week	SMD 0.00 (95% CI -0.28 to 0.27)	1 RCT (201)	High risk of bias and severe imprecision	Insufficient
	Pain free	2 hours	RR 1.73 (95% CI 1.04 to 2.86)	1 RCT (201)	High risk of bias and imprecision	3
	Pain relief	2 hours	RR 1.04 (95% CI 0.82 to 1.33)	1 RCT (201)	High risk of bias and severe imprecision	Insufficient
	Serious AE	N/A	Rate ratio 0.97 (95% CI 0.02 to 48.91)	1 RCT (201)	High risk of bias and severe imprecision	Insufficient

	Sustained pain free	1 week	RR 1.94 (95% CI 0.99 to 3.79)	1 RCT (201)	High risk of bias and severe imprecision	Insufficient
Noninvasive vagus nerve stimulation vs. sham stimulation	Pain free	2 hours	RR 1.43 (95% CI 0.92 to 2.22)	1 RCT (248)	Imprecision	2
	Pain relief	2 hours	RR 1.49 (95% CI 1.04 to 2.13)	1 RCT (248)	Imprecision	2
	Serious AE	N/A	Rate ratio 1.04 (95% CI 0.02 to 52.05)	1 RCT (248)	Severe imprecision	3
Remote electrical neuromodulation vs. sham stimulation	Pain free	2 hours	RR 1.95 (95% CI 1.19 to 3.19)	1 RCT (252)	Imprecision	2
	Pain relief	2 hours	RR 1.65 (95% CI 1.22 to 2.24)	1 RCT (252)	Imprecision	2
	Serious AE	N/A	Rate ratio 1.00 (95% CI 0.02 to 50.40)	1 RCT (252)	Severe imprecision	3
	Sustained pain free	1 week	RR 2.57 (95% CI 1.11 to 5.94)	1 RCT (252)	Imprecision	2
	Sustained pain relief	1 week	RR 2.27 (95% CI 1.30 to 3.95)	1 RCT (252)	Imprecision	2

AE=adverse event; CI=confidence interval; RCT=randomized controlled trial; RR=relative risk; N/A=not applicable; SMD=standardized mean difference

SMD>0 indicates the intervention mentioned first in the comparison is better