



# Management of chronic pain secondary to temporomandibular disorders: a systematic review and network meta-analysis of randomised trials

Liang Yao,<sup>1,2</sup> Behnam Sadeghirad,<sup>1,2</sup> Meixuan Li,<sup>3,4</sup> Jing Li,<sup>3,4</sup> Qi Wang,<sup>1</sup> Holly N Crandon,<sup>5,6</sup> Grace Martin,<sup>7</sup> Rebecca Morgan,<sup>1,8</sup> Ivan D Florez,<sup>9,10,11</sup> Birk Stokke Hunskaar,<sup>12</sup> Jeff Wells,<sup>13</sup> Sara Moradi,<sup>1,14</sup> Ying Zhu,<sup>1</sup> Muhammad Muneeb Ahmed,<sup>7</sup> Ya Gao,<sup>3</sup> Liujiao Cao,<sup>15</sup> Kehu Yang,<sup>3,4</sup> Jinhui Tian,<sup>3</sup> Jialing Li,<sup>16</sup> Linda Zhong,<sup>17</sup> Rachel J Couban,<sup>14</sup> Gordon H Guyatt,<sup>1</sup> Thomas Agoritsas,<sup>1,18,19</sup> Jason W Busse<sup>12,14</sup>

For numbered affiliations see end of the article

Correspondence to: K Yang, Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China, 730000;

yangkh@lzu.edu.cn (ORCID 0000-0001-7864-3012)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2023;383:e076226  
http://dx.doi.org/10.1136/bmj-2023-076226

Accepted: 09 November 2023

## ABSTRACT

### OBJECTIVE

We explored the comparative effectiveness of available therapies for chronic pain associated with temporomandibular disorders (TMD).

### DESIGN

Systematic review and network meta-analysis of randomised clinical trials (RCTs).

### DATA SOURCES

MEDLINE, EMBASE, CINAHL, CENTRAL, and SCOPUS were searched to May 2021, and again in January 2023.

### STUDY SELECTION

Interventional RCTs that enrolled patients presenting with chronic pain associated with TMD.

### DATA EXTRACTION AND SYNTHESIS

Pairs of reviewers independently identified eligible studies, extracted data, and assessed risk of bias. We captured all reported patient-important outcomes, including pain relief, physical functioning, emotional functioning, role functioning, social functioning, sleep quality, and adverse events. We conducted frequentist network meta-analyses to summarise the evidence and used the GRADE approach to rate the certainty of evidence and categorise interventions from most to least beneficial.

## RESULTS

233 trials proved eligible for review, of which 153—enrolling 8713 participants and exploring 59 interventions or combinations of interventions—were included in network meta-analyses. All subsequent effects refer to comparisons with placebo or sham procedures. Effects on pain for eight interventions were supported by high to moderate certainty evidence. The three therapies probably most effective for pain relief were cognitive behavioural therapy (CBT) augmented with biofeedback or relaxation therapy (risk difference (RD) for achieving the minimally important difference (MID) in pain relief of 1 cm on a 10 cm visual analogue scale: 36% (95% CI 33 to 39)), therapist-assisted jaw mobilisation (RD 36% (95% CI 31 to 40)), and manual trigger point therapy (RD 32% (29 to 34)). Five interventions were less effective, yet more effective than placebo, showing RDs ranging between 23% and 30%: CBT, supervised postural exercise, supervised jaw exercise and stretching, supervised jaw exercise and stretching with manual trigger point therapy, and usual care (such as home exercises, self stretching, reassurance). Moderate certainty evidence showed four interventions probably improved physical functioning: supervised jaw exercise and stretching (RD for achieving the MID of 5 points on the short form-36 physical component summary score: 43% (95% CI 33 to 51)), manipulation (RD 43% (25 to 56)), acupuncture (RD 42% (33 to 50)), and supervised jaw exercise and mobilisation (RD 36% (19 to 51)). The evidence for pain relief or physical functioning among other interventions, and all evidence for adverse events, was low or very low certainty.

## CONCLUSION

When restricted to moderate or high certainty evidence, interventions that promote coping and encourage movement and activity were found to be most effective for reducing chronic TMD pain.

## REGISTRATION

PROSPERO (CRD42021258567)

## Introduction

Temporomandibular disorders (TMD) are a group of painful conditions affecting the muscles of mastication, temporomandibular joint, and associated structures.<sup>1,2</sup> TMD are the second most common musculoskeletal chronic pain disorder, after low back pain; chronic TMD pain affects 6% to 9% of adults globally,<sup>3</sup> with

## WHAT IS ALREADY KNOWN ON THIS TOPIC

Temporomandibular disorders are the second most common cause of chronic musculoskeletal pain, after low back pain.

Several conservative, pharmacologic, and invasive management options are available to patients, but their comparative effectiveness is uncertain.

## WHAT THIS STUDY ADDS

A BMJ Rapid Recommendation guideline panel including patients, clinical experts, and methodologists defined the scope of our review, informed outcome selection and importance, and interpretation of findings.

Moderate certainty evidence shows that, compared with placebo or sham procedures, cognitive behavioural therapy augmented with biofeedback or relaxation therapy, therapist-assisted jaw mobilisation, and manual trigger point therapy are probably among the most effective interventions for pain relief.

Moderate to high certainty evidence found that five interventions were less effective, yet more effective than placebo/sham procedures for pain relief: cognitive behavioural therapy, supervised postural exercise, supervised jaw exercise and stretching with or without manual trigger point therapy, and usual care (such as home exercises, stretching, reassurance, self massage, and education).

**Box 1: Linked articles in this BMJ Rapid Recommendation cluster**

- Busse JW, Casassus R, Carrasco-Labra A, et al. Management of chronic pain associated with temporomandibular disorders: a clinical practice guideline. *BMJ* 2023;383:e076227 doi:10.1136/bmj-2023-076227
  - Summary of results from the Rapid Recommendation process
- Yao L, Sadeghirad B, Li M, et al. Management of chronic pain secondary to temporomandibular disorders: a systematic review and network meta-analysis of randomised trials. *BMJ* 2023;383:e076226 doi:10.1136/bmj-2023-076226
- MAGICapp (<https://matchit.magicevidence.org/230428dist-tmd>)
  - Expanded version of the results with a multilayered recommendation, evidence summaries, and decision aids for use on all electronic devices

women reporting a higher prevalence than men.<sup>4</sup> TMD have been proposed to include 12 subtypes: myalgia (local myalgia, myofascial pain, myofascial pain with referral), arthralgia, four types of disc displacement disorders, degenerative joint disease, subluxation, headache attributed to TMD, and others.<sup>5</sup>

Several conservative, invasive, and irreversible treatments are available for chronic TMD pain. Conservative therapies include jaw exercise, jaw mobilisation, cognitive behavioural therapy (CBT), self management (such as relaxation strategies), medication (such as anti-inflammatory drugs, analgesics, muscle relaxants), removable oral splints, low level laser therapy, and acupuncture. Invasive therapies include arthrocentesis, joint injection of local anaesthetic with or without steroids, and injection of purported regenerative substances (such as hyaluronic acid). Irreversible treatments include joint surgery, prosthodontics, orthodontics, and irreversible oral splints.

Several network meta-analyses (NMA) have previously summarised the efficacy and safety of interventions for chronic TMD pain but have important limitations.<sup>6-14</sup> These include ranking treatments using the “surface under the cumulative ranking curves” (SUCRA) approach, which ignores the certainty of evidence and the absolute differences between treatment alternatives,<sup>6-14</sup> addressing only certain categories of interventions or subtypes of TMD,<sup>6 7 10 13 14</sup> suboptimal reporting of pooled estimates (such as pooling different instruments that measure a common

domain as the standard mean difference,<sup>7 8 11</sup> only pooling trials that report the same continuous outcome measure<sup>6-14</sup>), focusing on statistical significance without considering whether the magnitude of effect is patient-important,<sup>6-14</sup> and failure to assess the certainty of evidence.<sup>8 13</sup>

We conducted a systematic review and network meta-analysis comparing all available interventions for chronic TMD pain, as a group of conditions, addressing these limitations. This review is part of the *BMJ* Rapid Recommendations project, a collaborative effort from the MAGIC Evidence Ecosystem Foundation (<https://magicevidence.org>) and *The BMJ*.<sup>15</sup> This systematic review informed a parallel guideline published in *The BMJ* and MAGICapp (box 1).<sup>16</sup>

**Methods**

We followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) extension statement for reporting of systematic reviews incorporating NMAs<sup>17</sup> and registered our review protocol with PROSPERO (CRD42021258567).

**Guideline panel involvement**

A *BMJ* Rapid Recommendations guideline panel provided critical oversight for our review, including: (1) defining the study question, (2) categorising interventions, (3) prioritising outcome measures, (4) proposing subgroup analyses, and (5) informing if measures of precision associated with pooled effect estimates were imprecise. The panel included eight clinical experts (two general dentists (AC-L, CP), an orofacial pain specialist (RC), an oral surgeon with expertise in orofacial pain (JD), an oral physician (JMZ), a general internist (MC), and a clinical pharmacologist (CS)), eight methodologists (four of whom were also front-line clinicians), a patient liaison expert, and three patient partners living with chronic TMD pain. All patients received training and support to optimise contributions throughout the guideline development process. The members of the guideline panel led the interpretation of the results based on what they expected the typical values and preferences of patients to be, as well as the variation between patients.

**Patient and public involvement**

Three patients were full members of the guideline panel and contributed to the selection and prioritisation of outcomes, values and preferences assessments, critical feedback to the protocol, and interpretation of findings for the systematic review and the associated *BMJ* Rapid Recommendation.

**Data sources and searches**

An experienced medical librarian (RJC) developed and refined database-specific search strategies (see appendix 1) and searched MEDLINE, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Central Register of Controlled Trials (CENTRAL), and SCOPUS from

**Box 2: Classification of interventions using a minimally contextualised framework<sup>41</sup>**

- Step 1. Choosing a reference intervention (“placebo/sham” was our reference).
- Step 2. Classifying interventions into categories based on comparison with the reference: category 1, not convincingly different than placebo/sham; and category 2 or higher, more effective (or harmful for adverse events) than placebo/sham.
- Step 3. Further classification of category 2 or higher interventions based on comparisons between pairs of interventions. If any intervention proved more effective than another category 2 intervention, that intervention was moved to a higher rated group (category 3). We implemented this same step to differentiate among interventions in category 3 (if there was an intervention in category 3 superior to at least one other, it would move to category 4) until no new groupings resulted.
- Step 4. Separating interventions into two clusters according to certainty of evidence: high or moderate certainty of evidence, and low or very low certainty of evidence.

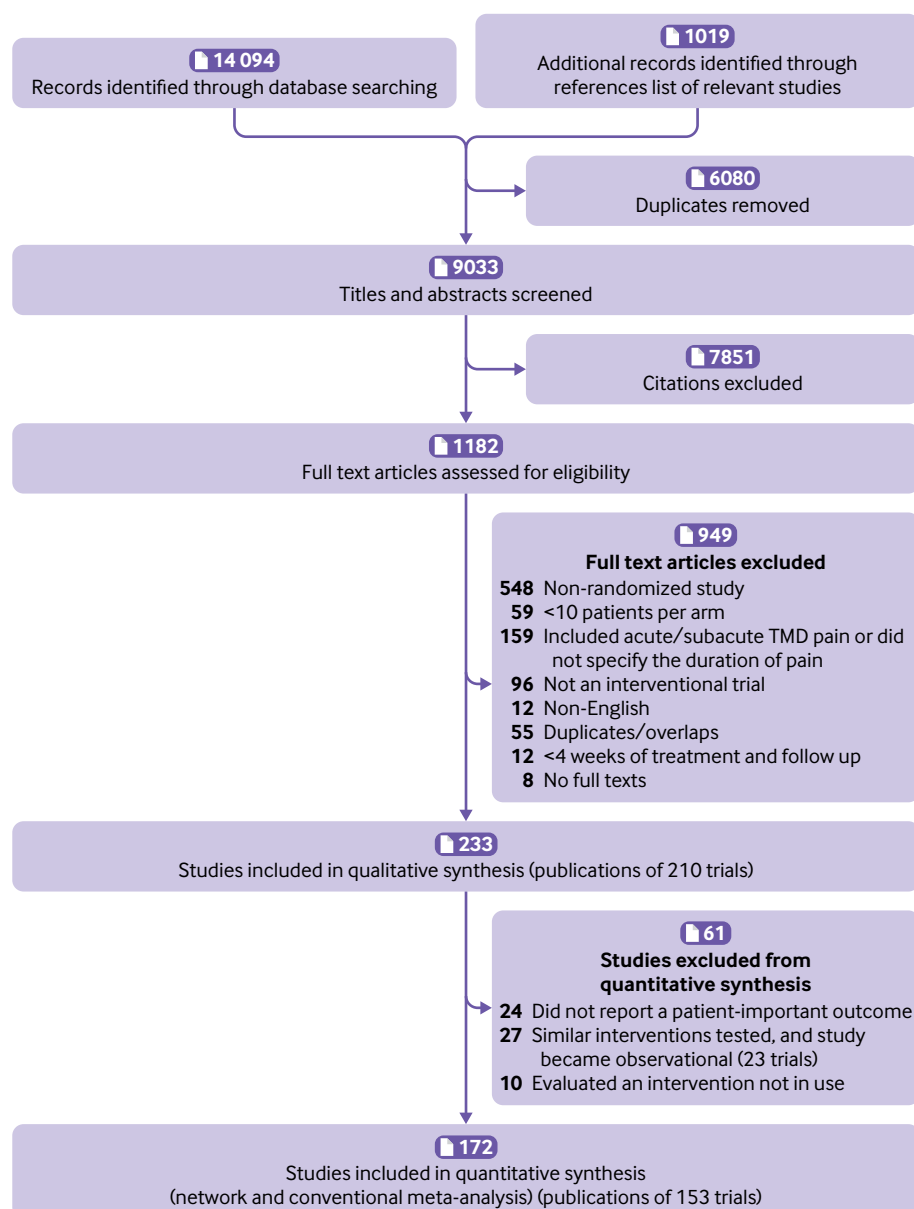


Fig 1 | Study selection process in review of interventions for chronic pain associated with temporomandibular disorders

inception through May 2021. After formulation of the recommendations, we updated our search to January 2023. We reviewed reference lists of included trials and relevant reviews for additional eligible studies.

#### Study selection and eligibility criteria

Pairs of trained reviewers screened titles and abstracts of identified citations independently and in duplicate, using a standardised, pilot-tested form, and assessed full texts of all potentially eligible studies. Disagreements were resolved through discussion to achieve consensus. We used online systematic review software (DistillerSR, Evidence Partners, Ottawa, Canada; <http://systematic-review.net>) to facilitate literature screening.

We included English language, interventional trials that: (1) enrolled adults ( $\geq 18$  years old) living with chronic pain associated with a TMD ( $\geq 3$  months or

defined by the authors as “chronic”); (2) randomised them to an active treatment or an alternative treatment, placebo, sham procedure, or usual care; and (3) included at least 10 participants in each treatment group. All interventions were reviewed and categorised by two or more clinical experts, independently and blinded to study results, and we excluded trials of interventions that our experts advised were not currently in use. We reviewed details of interventions labelled as “physiotherapy” and clustered them according to their reported components, as physiotherapy is a profession and not a single modality.

#### Data extraction

Pairs of reviewers independently abstracted the following data from each eligible article: study characteristics (such as bibliographic information, country of origin, funding source), participant

**Table 1 | Characteristics of included studies for network meta-analysis (n=153)**

Characteristics	No (%) of studies or median (interquartile range)
Sample size	46 (35 to 63)
Age	35 (30 to 39)
Sex (% women)	83 (78 to 91)
Pain duration (months)	44 (13 to 65)
Baseline pain intensity (on 10 cm visual analogue scale)	5.4 (4.3 to 6.6)
Follow up duration (weeks)	12 (5 to 52)
Year of publication (range:1983-2021):	
≤2000	10 (7%)
2001-05	14 (9%)
2006-10	19 (12%)
2011-15	35 (23%)
2016-20	75 (49%)
Country/region:	
Europe	77 (50%)
Asia	28 (18%)
North America	24 (16%)
South America	22 (14%)
Australia	2 (1%)
Diagnosis of temporomandibular disorders (TMD):	
Research Diagnostic Criteria for Temporomandibular Disorders	110 (72%)
Clinical diagnosis and radiographs	14 (9%)
Clinical diagnosis	13 (8%)
Criteria of American Board of Orofacial Pain	7 (5%)
Other	9 (6%)
Pain duration:	
Reported a specific duration (median 42 months, IQR 13 to 65)	47 (31%)
At least 3 months	28 (18%)
At least 6 months	17 (11%)
No duration provided; pain reported as "chronic"	61 (40%)
TMD classifications:	
Myofascial TMD	38 (25%)
TMD with internal derangement of the joint	26 (17%)
TMD with degenerative joint disease	11 (7%)
Mixed patients	12 (8%)
Not specified	66 (43%)
Scales used for measuring pain:	
10 cm visual analogue scale	115 (78%)
11 point numeric rating scale	14 (9%)
Others	19 (13%)
Scales used for measuring physical function:	
Short Form-36	11 (31%)
Jaw Functional Limitation Scale	9 (25%)
Mandibular function impairment questionnaire	3 (8%)
World Health Organization Quality-of-Life Scale	2 (5%)
Others	11 (31%)
Reported funding:	
Non-industry funding	62 (41%)
Unfunded	23 (15%)
No funding statement	68 (44%)

characteristics (such as sample size, age and sex of participants, subtypes of TMD, pain severity, duration of pain), and characteristics of interventions and comparators. We classified countries where trials were conducted as high, upper-middle, lower-middle, or low income as designated by the World Bank.<sup>18</sup> We also extracted data for all patient-important outcomes as guided by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations,<sup>19 20</sup> including pain, physical functioning, emotional functioning, role functioning, social functioning, sleep quality, and adverse events. For trials with different follow-up lengths, we abstracted data from the longest follow-up reported.

### Risk of bias assessment

Pairs of reviewers independently assessed the risk of bias among eligible studies using a modified Cochrane risk of bias instrument that included random sequence generation; allocation concealment; blinding of participants, healthcare providers, and outcome assessor/adjudicator; and incomplete outcome data (≥20% missing data was considered high risk of bias).<sup>21</sup> We rated the risk of bias for each criterion as "low," "probably low," "high," or "probably high." We resolved disagreements between reviewers through discussion. When all the above domains were judged at low or probably low risk, we rated the overall risk of bias as "low," otherwise we rated the overall risk of bias as "high." For interventions in which blinding is

not possible, and when blinding was the only criterion not met, we referred to prior meta-epidemiological studies which showed no systematic difference in estimated treatment effect between trials with and without blinded patients, healthcare providers, or outcome assessors,<sup>22-26</sup> and rated the overall risk of bias as “probably low.”

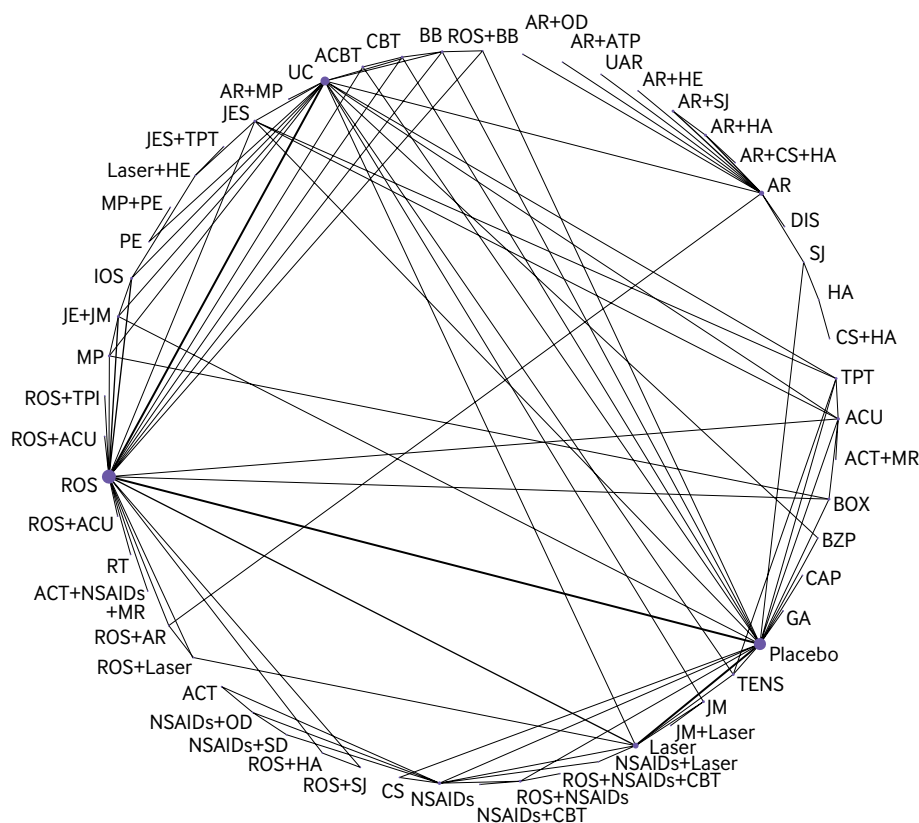
### Data synthesis

We used DerSimonian-Laird random-effects models for meta-analysis of direct comparisons for all patient-important outcomes reported by more than one trial. For pain and function, when studies reported effect estimates using different measurement instruments that captured a common construct, we transformed treatment effects to a common instrument score on a domain-by-domain basis (appendix 2).<sup>27</sup> Specifically, we converted pain intensity to a 10 cm visual analogue scale for pain, and physical functioning to the 100-point 36-item Short Form Survey (SF-36) physical component summary score. We then calculated the weighted mean difference and the associated 95% confidence interval using change scores from baseline to the end of follow-up to address interpatient variability. If authors did not report change scores, we estimated them using the

baseline and end-of-study scores and the associated standard deviations and median correlation coefficient reported by low risk of bias trials. We used methods described in the Cochrane Handbook<sup>28</sup> and by Hozo et al<sup>29</sup> to impute means and standard deviations when the median, range, and sample size were reported, or to impute the standard deviation when the standard error or standard deviation for the differences was not reported.

We pooled dichotomous outcomes (that is, adverse events) as the odds ratio and 95% confidence interval. When at least 10 trials were available for a direct comparison, we assessed small-study effects using Harbord's test for binary outcomes and Egger's test for continuous outcomes.<sup>30 31</sup> Further, when only a single trial was available to inform the effectiveness of an intervention, and reported a large significant effect, we considered this evidence at high risk of bias due to small study effects.

We constructed networks for outcomes in which 10 or more trials contributed data. When networks are sparse, the contrast-based random-effects model may generate non-credible wide confidence intervals for network estimates, even when the direct and indirect estimates are coherent (that is, the confidence interval



**Fig 2 | Network plot for pain relief.** The figure shows lines between treatments that have direct comparisons; thicker lines indicate that more comparison studies have been conducted. This figure is based on data in supplementary table 8. ACBT = augmented cognitive behavioural therapy; ACT = acetaminophen; ACU = acupuncture; AR = arthrocentesis; ATP = antidepressant; BB = beta-blocker; BOX = botulinum toxin; BZP = benzodiazepine; CAP = capsaicin; CBT = cognitive behavioural therapy; CS = cartilage supplement; DIS = discectomy; GA = gabapentin; HA = hyaluronic acid injection; HE = home exercises; IOS = irreversible oral splint; JE = jaw exercise; JES = jaw exercise + stretching; JM = jaw mobilisation; laser = low-level laser therapy; MP = manipulation; MR = muscle relaxant; NSAID = non-steroidal anti-inflammatory drug; OD = opioid; PE = postural exercise; ROS = removable occlusal splint; RT = relaxation therapy; SD = steroid; SJ = steroid injection; TENS = transcutaneous electrical nerve stimulation; TPI = trigger point injection; TPT = trigger point therapy; UAR = ultrasound-guided arthrocentesis; UC = usual care.



of the network estimate is wider than both the estimates of precision associated with the direct and indirect effects).<sup>32</sup> We used a fixed-effect model for pooling in such cases. We used the “design-by-treatment” model

(global test) to assess the coherence assumption for each network.<sup>33</sup> We used the side-splitting method to evaluate local (loop-specific) incoherence in each closed loop of the network as the difference between

Intervention	Pain Relief (Score: 0-10cm, the lower score the better)			Physical Functioning (Score: 0-100mm, the higher score the better)			Adverse Events	
	NMA Estimate, MD (95%CI)	RD for achieving ≥ MID of 1cm, % (95%CI)	COE	NMA Estimate, MD (95%CI)	RD for achieving ≥ MID of 5mm, % (95%CI)	COE	Odds ratio (95%CI)	COE
Augmented cognitive behavioural therapy	-2.62 (-3.03 to 2.20)	36 (33 to 39)	M	10.10 (1.62 to 18.59)	30 (11 to 47)	L		
Jaw mobilization	-2.57 (-3.18 to 1.96)	36 (31 to 40)	M	19.77 (13.47 to 26.08)	49 (37 to 58)	L	1.70 (0.06 to 50.71)	VL
Trigger point therapy	-2.08 (-2.31 to 1.84)	32 (29 to 34)	M	9.73 (4.31 to 15.16)	29 (17 to 40)	L		
Cognitive behavioural therapy	-1.88 (-2.42 to 1.35)	30 (23 to 35)	M	8.25 (2.25 to 14.26)	26 (12 to 39)	L		
Postural exercise	-1.56 (-2.33 to 0.79)	26 (14 to 34)	M	28.73 (20.01 to 37.46)	61 (49 to 67)	L	1.90 (0.02 to 237.71)	VL
Jaw exercise + stretching	-1.55 (-1.99 to 1.11)	26 (20 to 31)	H	16.23 (11.58 to 20.88)	43 (33 to 51)	M		
Usual care	-1.47 (-1.69 to 1.25)	25 (22 to 28)	M	9.13 (3.73 to 14.53)	28 (16 to 39)	L	1.73 (0.11 to 27.25)	VL
Jaw exercise + stretching + trigger point therapy	-1.31 (-1.99 to 0.62)	23 (11 to 31)	M	13.02 (4.68 to 21.35)	36 (18 to 51)	L		
Removeable occlusal splint + steroid injection	-6.11 (-7.42 to 4.80)	44 (43 to 44)	VL					
Removeable occlusal splint + acupuncture	-4.86 (-6.05 to 3.67)	43 (41 to 44)	VL					
Removeable occlusal splint + hyaluronic acid injection	-4.61 (-6.07 to 3.15)	43 (39 to 44)	VL					
Removeable occlusal splint + trigger point injection	-3.85 (-4.63 to 3.07)	42 (39 to 43)	VL					
Arthrocentesis + antidepressant	-4.39 (-6.69 to 2.08)	43 (32 to 44)	VL					
Arthrocentesis + home exercises	-3.83 (-5.48 to 2.17)	42 (33 to 44)	VL					
Arthrocentesis + opioid	-3.67 (-5.35 to 1.99)	41 (31 to 44)	VL					
Arthrocentesis + hyaluronic acid injection	-3.30 (-4.29 to 2.31)	40 (34 to 43)	L					
Manipulation	-3.02 (-3.54 to 2.50)	39 (35 to 41)	L	16.30 (7.77 to 24.83)	43 (25 to 56)	M		
Jaw exercise + mobilization	-2.86 (-3.21 to 2.52)	38 (36 to 40)	L	13.11 (5.42 to 20.81)	36 (19 to 51)	M	1.03 (0.02 to 53.71)	VL
Irreversible oral splint	-2.67 (-3.19 to 2.15)	37 (32 to 40)	L	15.91 (7.41 to 24.42)	42 (24 to 56)	VL		
Removeable occlusal splint + arthrocentesis	-2.67 (-3.53 to 1.81)	37 (29 to 41)	L					
Arthrocentesis + cartilage supplement + hyaluronic acid injection	-3.82 (-6.24 to 1.41)	42 (24 to 44)	VL				1.33 (0.36 to 4.98)	VL
Cartilage supplement + hyaluronic acid injection	-3.37 (-5.18 to 1.55)	40 (26 to 44)	VL					
NSAIDs + low-level laser therapy	-2.76 (-4.10 to 1.43)	37 (24 to 42)	VL					
Manipulation + postural exercise	-2.52 (-3.69 to 1.34)	36 (23 to 41)	L	33.95 (22.99 to 44.92)	65 (54 to 69)	VL	1.16 (0.01 to 165.13)	VL
Removeable occlusal splint + low-level laser therapy	-2.32 (-3.29 to 1.36)	34 (23 to 40)	L					
Removeable occlusal splint + jaw mobilization	-2.14 (-2.84 to 1.44)	32 (24 to 38)	L					
Acupuncture	-2.04 (-2.38 to 1.71)	31 (28 to 35)	L	16.04 (11.60 to 20.48)	42 (33 to 50)	M	1.56 (0.06 to 40.11)	VL
Removeable occlusal splint	-1.81 (-2.01 to 1.60)	29 (26 to 31)	L	11.03 (7.23 to 14.83)	32 (24 to 40)	L	1.34 (0.37 to 4.83)	VL
Transcutaneous electrical nerve stimulation	-1.78 (-2.00 to 1.57)	29 (26 to 31)	L	9.23 (3.82 to 14.65)	28 (16 to 39)	L	0.45 (0.01 to 23.70)	VL
Discectomy	-3.12 (-5.51 to 0.73)	39 (13 to 44)	VL					
Ultrasound-guided arthrocentesis	-3.12 (-5.20 to 1.03)	39 (18 to 44)	VL					
NSAIDs + cognitive behavioural therapy	-2.83 (-4.76 to 0.91)	38 (16 to 43)	L					
Jaw mobilization + low-level laser therapy	-2.22 (-3.68 to 0.76)	33 (14 to 41)	VL					
Arthrocentesis	-2.12 (-3.06 to 1.17)	32 (20 to 39)	L					
NSAIDs + opioid	-2.14 (-2.95 to 1.33)	32 (23 to 38)	VL				2.20 (0.04 to 123.00)	VL
Gabapentin	-2.11 (-3.33 to 0.89)	32 (16 to 40)	VL	10.90 (-3.14 to 24.94)		VL		
Low-level laser therapy + home exercises	-1.75 (-2.33 to 1.17)	28 (20 to 34)	L	26.27 (11.35 to 41.18)	58 (33 to 68)	L		
Botulinum toxin	-1.46 (-1.79 to 1.13)	25 (20 to 29)	L	0.30 (-16.34 to 16.94)		L		
Low-level laser therapy	-1.36 (-1.56 to 1.16)	23 (20 to 26)	L	2.15 (-4.48 to 8.78)		M	1.11 (0.11 to 11.07)	VL
Hyaluronic acid injection	-2.41 (-4.11 to 0.70)	35 (13 to 42)	L				1.09 (0.41 to 2.91)	VL
Arthrocentesis + manipulation	-2.39 (-4.38 to 0.40)	35 (8 to 43)	VL					
Relaxation therapy	-2.01 (-3.34 to 0.67)	31 (12 to 40)	VL					
Removeable occlusal splint + NSAIDs + cognitive behavioural therapy	-1.97 (-3.23 to 0.72)	31 (13 to 40)	VL					
Trigger point injection	-1.68 (-2.80 to 0.55)	27 (10 to 38)	VL					
Arthrocentesis + steroid injection	-1.45 (-2.44 to 0.45)	24 (8 to 35)	L					
Acetaminophen + muscle relaxant	-1.39 (-2.69 to 0.10)	24 (2 to 37)	VL					
Cartilage supplements	-1.06 (-1.56 to 0.56)	19 (10 to 26)	L				1.00 (0.45 to 2.22)	VL
NSAIDs	-0.76 (-1.15 to 0.38)	14 (7 to 20)	L	15.03 (1.94 to 28.13)	40 (12 to 60)	L	2.10 (0.94 to 4.67)	L
Capsaicin cream	-1.21 (-2.82 to 0.40)		L					
Biofeedback	-1.03 (-2.42 to 0.37)		L	6.45 (-31.28 to 44.19)		VL		
Steroid injection	-1.01 (-2.13 to 0.12)		L	3.00 (-6.60 to 12.60)		L	2.22 (0.83 to 5.93)	VL
Removeable occlusal splint + betablocker	-0.95 (-2.58 to 0.69)		L	35.53 (12.83 to 58.23)	66 (36 to 70)	L		
Benzodiazepine	-0.80 (-2.25 to 0.65)		L					
Removeable occlusal splint + NSAIDs	-0.66 (-1.46 to 0.14)		L					
Acetaminophen + NSAIDs + muscle relaxant	-0.66 (-1.35 to 0.03)		VL	-12.97 (-21.1 to 4.84)		VL	14.95 (0.58 to 388.06)	VL
Beta-blocker	-0.39 (-0.78 to 0.00)		L	0.40 (-1.91 to 2.71)		M	2.05 (0.99 to 4.22)	L
NSAIDs + steroid	-0.34 (-1.19 to 0.51)		L				2.20 (0.04 to 123.00)	VL
Acetaminophen	0.20 (-0.61 to 1.00)		L				1.88 (0.03 to 104.70)	VL

Fig 3 | Network meta-analysis results, sorted by GRADE certainty of evidence and effect estimate for the comparisons of active treatments versus placebo/sham procedures, for pain relief, physical function, and adverse events

Certainty	Classifications (the higher level, the better effectiveness)
High (moderate to high)	Category 3 interventions: more effective than one or more category 2 interventions
	Category 2 interventions: more effective than one or more category 1 interventions, but less effective than a category 3 interventions
	Category 1 interventions: more effective than placebo, but less effective than a category 2 interventions
Low (low to very low)	Category 7 interventions: may be more effective than one or more category 6 interventions
	Category 6 interventions: may be more effective than one or more category 5 interventions, but less effective than a category 7 interventions
	Category 5 interventions: may be more effective than one or more category 4 interventions, but less effective than a category 6 interventions
	Category 4 interventions: may be more effective than one or more category 3 interventions, but less effective than a category 5 interventions
	Category 3 interventions: may be more effective than one or more category 2 interventions, but less effective than a category 4 interventions
	Category 2 interventions: may be more effective than placebo, but less effective than a category 3 interventions
	Category 1 interventions: May be among the least effective (for benefits), or may be no more harmful than placebo (for harms)

**Fig 3 | contd Network meta-analysis results, sorted by GRADE certainty of evidence and effect estimate for the comparisons of active treatments versus placebo/sham procedures, for pain relief, physical function, and adverse events**

direct and indirect evidence.<sup>34 35</sup> We performed all analyses in STATA 17.0 MP edition (StataCorp, College Station, TX, USA).

Mean differences are often misunderstood by knowledge users, including clinicians and patients.<sup>36</sup> Therefore, for continuous outcomes, we used the network estimate of treatment effects to model the risk difference for achieving the minimally important difference (MID) to optimise interpretability.<sup>37</sup> We used 1 cm on the 10 cm visual analogue scale as the MID for pain relief,<sup>38</sup> and 5 points on the 100-point SF-36 physical component summary score as the MID for physical function.<sup>39 40</sup> We used the median and standard deviation of the reference treatment (placebo/sham), with the established MID for the outcome in question to model the probability of achieving  $\geq$ MID in the reference treatment. We then used the pooled mean difference to estimate the mean in the treatment group and calculate the probability of achieving  $\geq$ MID in the treatment group. Finally, we used the calculated probabilities for both groups to acquire the risk difference for achieving  $\geq$ MID.

### Categorisation of interventions

We categorised interventions from most to least effective, using a minimally contextualised approach.<sup>41</sup> The minimally contextualised framework is based on two principles: interventions should be grouped in categories, from the most to the least effective or harmful, and judgments that place interventions in such categories should simultaneously consider the estimates of effect and the certainty of evidence (box 2).

For pain relief and functional improvement, we created groups of interventions as follows: (1) category 1, the reference intervention (placebo/sham procedures) and interventions no different from placebo, which we refer to as “among the least effective”; (2) category 2, interventions superior to placebo but inferior to a category 3 intervention; and (3) category 3, interventions that proved superior to at least one category 2 intervention. We used the same approach for adverse events but created groups

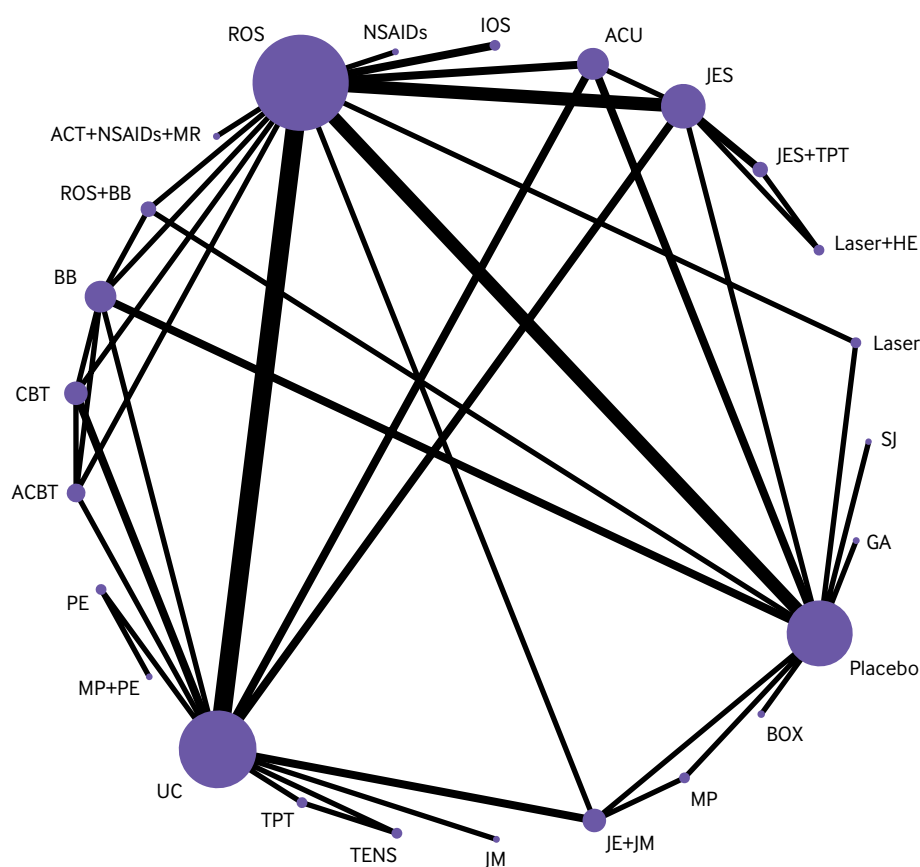
of interventions as follows: (1) no more harmful than placebo; (2) less harmful than a category 3 intervention, but more harmful than placebo; and (3) more harmful than at least one category 2 intervention. We created additional categories for benefits or harms, as needed, using the same approach. For both benefits and harms, we categorised interventions as those supported by moderate or high certainty evidence, and those supported by low or very low certainty evidence relative to placebo/sham procedures.<sup>34 42</sup>

### Subgroup analysis

At the direction of the guideline panel, we explored four a priori subgroup hypotheses to explain variability between trials: (1) subtypes of TMD will show different treatment effects; (2) studies at higher versus lower risk of bias will show larger treatment effects; (3) trials with longer versus shorter follow-up will show smaller treatment effects; and (4) studies enrolling patients receiving disability benefits or engaged in litigation versus not will show smaller treatment effects. We only conducted subgroup analyses if there were two or more studies in each subgroup and used a test of interaction to establish whether subgroups differed significantly from one another. We assessed the credibility of statistically significant subgroup effects (P value for test of interaction  $<0.05$ ) using Instrument to assess the Credibility of Effect Modification Analyse (ICEMAN) criteria.<sup>43</sup>

### Certainty of evidence

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the certainty of evidence for direct, indirect, and network estimates for all outcomes. With this approach, the certainty of direct evidence from randomised trials starts as high but may be rated down for risk of bias, indirectness, imprecision, inconsistency, or small study effects to moderate, low, or very low.<sup>44</sup> Certainty ratings of indirect estimates start at the lowest GRADE rating of the direct comparisons that contributed the most weight to the dominant first-order loop in the



**Fig 4 | Network plot of physical function.** The figure shows lines between treatments that have direct comparisons; thicker lines indicate that more comparison studies have been conducted. This figure is based on data in supplementary table 11. ACBT = augmented cognitive behavioural therapy; ACT = acetaminophen; ACU = acupuncture; BB = beta-blocker; BOX = botulinum toxin; CBT = cognitive behavioural therapy; GA = gabapentin; HA = hyaluronic acid injection; HE = home exercises; IOS = irreversible oral splint; JE = jaw exercise; JES = jaw exercise + stretching; JM = jaw mobilisation; laser = low-level laser therapy; MP = manipulation; MR = muscle relaxant; NSAID = non-steroidal anti-inflammatory drug; PE = postural exercise; ROS = removeable occlusal splint; SJ = steroid injection; TENS = transcutaneous electrical nerve stimulation; TPT = trigger point therapy; UC = usual care.

network, with further consideration of rating down for intransitivity when present.<sup>45-47</sup>

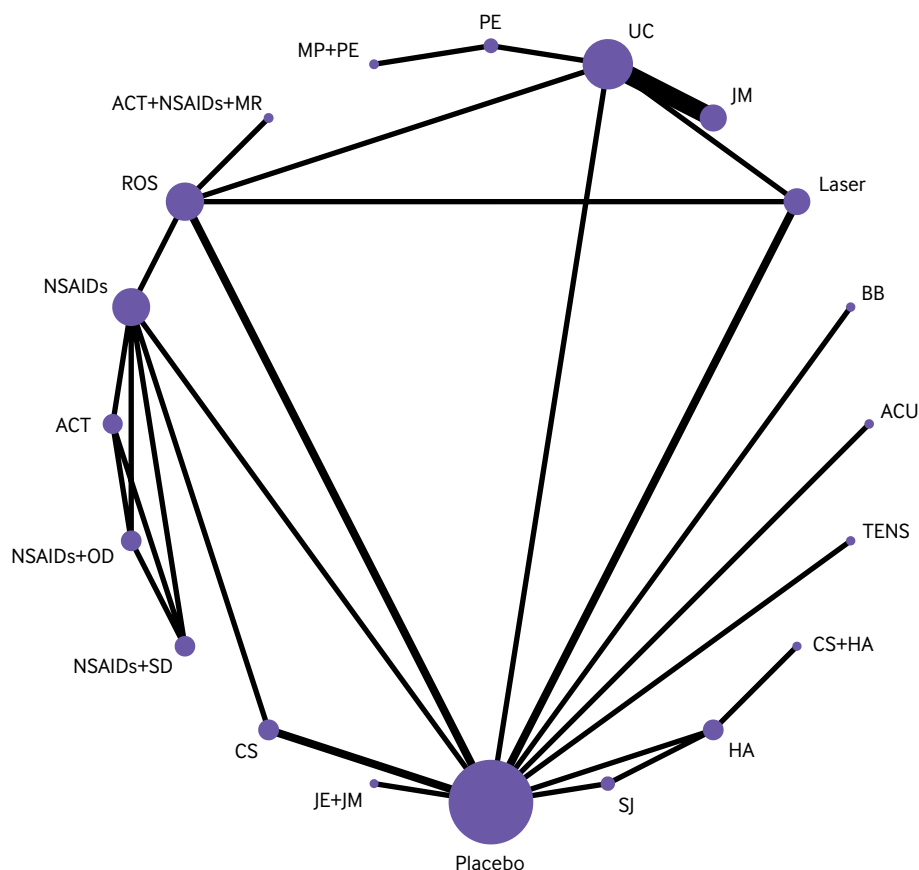
Our assessment of transitivity relied on two fundamental issues: (1) eligible trials are jointly randomisable, and (2) potential effect modifiers are equally distributed between each treatment comparison in the network. We addressed the first issue by exploring the similarity of patient populations in our networks and confirming with our clinical experts that patients across trials were eligible to receive any of interventions considered in the network. We addressed the second issue by generating graphs to explore if the distribution of effect modifiers (that is, age, sex, risk of bias) were similar across comparisons. We were unable to explore subtype of TMD as an effect modifier as most trials enrolled mixed subtypes and reported aggregate results or did not report which subtype(s) were enrolled.

For the certainty of network estimates, we started with the estimate—direct or indirect—that dominated (contribution >50%) the network estimate, or we used the higher certainty of the direct and indirect estimates if they both contributed equally to the network estimate. Discounting imprecision associated with the direct estimate, which was assessed at

the network level, as was incoherence. We judged network estimates as imprecise if the associated 95% confidence interval included decision thresholds<sup>48</sup> chosen by the guideline panel: half the minimally important difference for continuous outcomes and the null value (odds ratio=1) for adverse events. We did not rate down for imprecision if the confidence interval excluded the decision threshold unless the comparison was statistically significant and informed by fewer than 400 observations for continuous outcomes or 300 events for binary outcomes.<sup>49</sup>

When present, we addressed incoherence in the network estimate by using the higher certainty evidence (direct or indirect) instead of the network estimate, or when the certainty of evidence was the same for the direct and indirect evidence, we used the network estimate but rated down the certainty of evidence one level for incoherence. We did not rate down the certainty rating of the network estimate twice if both intransitivity and incoherence were present, as these are related issues. We followed GRADE guidance for communicating our findings<sup>50</sup> and made our results available in the MAGICapp interactive decision support tool MATCH-IT (<https://matchit.magicvidence.org/230428dist-tmd>).





**Fig 5 | Network plot of adverse events.** The figure shows lines between treatments that have direct comparisons; thicker lines indicate that more comparison studies have been conducted. This figure is based on data in supplementary table 14. ACT= acetaminophen; ACU= acupuncture; CS= cartilage supplement; HA= hyaluronic acid injection; JE= jaw exercise; JM= jaw mobilisation; laser= low-level laser therapy; MP= manipulation; MR= muscle relaxant; NSAID= non-steroidal anti-inflammatory drug; OD= opioid; PE= postural exercise; ROS= removeable occlusal splint; SD= steroid; SJ= steroid injection; TENS= transcutaneous electrical nerve stimulation; UC= usual care.

## Results

Our search identified 9033 unique citations, of which 233 (reporting 210 unique trials) proved eligible for review (fig 1). We were unable to acquire the full texts of eight citations for review, and our attempts to contact study authors were unsuccessful (appendix 3). Twenty four trials reported only surrogate outcomes (such as passive jaw opening, pain on palpation), 23 trials compared similar interventions that our clinical experts advised should be collapsed (for example, different types of removable oral splints, two-needle arthrocentesis v single-needle arthrocentesis; none of which showed different treatment effects<sup>6 51</sup>) and thus became observational data, and 10 trials compared interventions that are not typically available in current practice (supplement table 1). Thus, our meta-analyses included 153 studies from 172 references, which enrolled 8713 patients with chronic TMD pain.

Our updated literature search in January 2023, identified 13 additional eligible trials that enrolled a total of 875 patients, four of which only reported treatment effects on surrogate outcomes (supplement table 2). We did not include the results of the nine trials that reported patient-important outcomes in our pooled estimates of treatment effects to maintain

consistency with the evidence used by the guideline panel to formulate recommendations.

## Study characteristics

Studies eligible for our review were parallel trials that evaluated 23 conservative interventions, 15 pharmacological interventions, seven combinations of pharmacological and conservative interventions, 13 surgical interventions with or without adjunct treatments, and placebo/sham procedures (supplement tables 3-5). The median sample size of included trials was 46 (interquartile range 35-63), and the median of the mean age among patients was 35 years (30-39). The median average pain score at baseline was 5.4 cm (4.3-6.6) on a 10cm visual analogue scale, which equates to moderate severity pain,<sup>52</sup> and the median average pain duration when reported was 44 months (13-65). The median follow-up was 12 weeks (5-52).

Most trials provided no details on funding (44%) or declared non-industry funding (41%), with the remainder reporting no funding (15%) (table 1). Half of trials (75/153, 49%) enrolled patients with a specific subtype of TMD (most often myofascial TMD (38/75, 51%)), whereas the remainder (51%) enrolled mixed subtypes or did not report the proportion of TMD subtypes among patients (table 1). No trial explicitly

reported enrolment of veterans, or patients involved in litigation or receiving disability benefits. Almost half of all trials (47%) excluded TMD patients with comorbid mental illness, and only three trials (2%) reported enrolling such patients. Approximately a third of trials explicitly excluded TMD patients with comorbid rheumatoid arthritis (36%) or prior TMJ surgery (32%), and 20% excluded those with comorbid fibromyalgia; only a single trial reported enrolling any patients with these features (supplement table 6).

### Risk of bias

Most trials (133/153, 87%) were at risk for bias for at least one domain; 118 studies (77%) adequately generated their randomisation sequence; 78 (51%) concealed allocation; 65 (42%) blinded patients, 48 (31%) blinded healthcare providers, and 65 (42%) blinded outcome assessors. Most trials (126; 82%) reported <20% missing outcome data. We rated the overall risk as high in 96 (63%) studies, and low or probably low in 57 (37%) studies (supplement table 7).

### Main findings per outcome

#### *Pain relief*

In total, 148 RCTs involving 7867 patients and evaluating 59 interventions reported effects on pain (fig 2). In the 31 direct comparisons with two or more studies available for conventional pairwise meta-analysis, there was substantial heterogeneity ( $I^2 > 70\%$ ) for 11 comparisons (supplement table 8, the expanded GRADE assessment table for pain relief can be found at [https://zenodo.org/record/7607558#Y9\\_P1uyZM-Q](https://zenodo.org/record/7607558#Y9_P1uyZM-Q); the league table for pain relief can be found at <https://zenodo.org/record/7604396#Y93irOyZM-Q>). We found no evidence of global or loop-specific incoherence (supplement table 9).

Compared with placebo, effects on pain for eight interventions were supported by moderate to high certainty evidence. The three therapies probably most effective for pain relief were cognitive behavioural therapy (CBT) augmented with biofeedback or relaxation therapy (modelled risk difference (RD) for achieving the minimally important difference in pain relief of 1 cm on a 10 cm scale 36% (95% CI 33 to 39)), therapist-assisted jaw mobilisation (RD 36% (31 to 40)), and manual trigger point therapy (RD 32% (29 to 34)) (fig 3).

Five interventions were less effective, yet more effective than placebo/sham procedures: cognitive behavioural therapy (RD 30% (23 to 35)), supervised postural exercise (RD 26% (14 to 34)), supervised jaw exercise with stretching (RD 26% (20 to 31)), usual care (such as home exercises, stretching, reassurance, self massage, and education) (RD 25% (22 to 28)), and supervised jaw exercise with stretching and manual trigger point therapy (RD 23% (11 to 31)). The certainty in effects for all other interventions was low or very low (fig 3).

#### *Physical functioning*

Thirty six trials with 2009 patients, evaluating 30 interventions, reported treatment effects on physical functioning; however, four interventions reported in

three studies (that is, arthrocentesis, arthrocentesis with steroid injection, arthrocentesis with hyaluronic acid injection, and discectomy) were disconnected from the network (supplement table 10). As such, our network included 33 RCTs with 1910 patients that evaluated 26 interventions (fig 4). The 33 RCTs included 44 direct comparisons, of which 12 included two or more studies for conventional pairwise meta-analysis and one showed substantial heterogeneity (supplement table 11, the expanded GRADE assessment table for physical functioning can be found at [https://zenodo.org/record/7607558#Y9\\_P1uyZM-Q](https://zenodo.org/record/7607558#Y9_P1uyZM-Q); the league table for physical functioning can be found at <https://zenodo.org/record/7604396#Y93irOyZM-Q>). We found no evidence of global or loop-specific incoherence (supplement table 12).

Moderate certainty evidence showed that, compared with placebo, four interventions probably improved physical functioning: supervised jaw exercise with stretching (RD for achieving the minimally important difference of 5 points on the 100 point short form-36 physical component summary score 43% (95% CI 33 to 51)), manipulation (RD 43% (25 to 56)), acupuncture (RD 42% (33 to 50)), and supervised jaw exercise with mobilisation (RD 36% (19 to 51)). The certainty in effects for physical functioning among other interventions was low or very low (fig 3).

#### *Adverse events*

Adverse events were only reported in 31 studies involving 1987 patients with chronic TMD (fig 5, supplement table 13). Four studies evaluating arthrocentesis with or without co-interventions (arthrocentesis, arthrocentesis plus steroid, arthrocentesis plus opioid, ultrasound guided arthrocentesis, and arthrocentesis plus hyaluronic acid injection), and reporting no adverse events, were disconnected and thus not included in our network (supplement table 14).

Of the 20 direct comparisons that contributed to our network, four had two or more studies for pairwise meta-analysis, and none showed substantial heterogeneity (supplement table 15, the expanded GRADE assessment table for adverse events can be found at [https://zenodo.org/record/7607558#Y9\\_P1uyZM-Q](https://zenodo.org/record/7607558#Y9_P1uyZM-Q); the league table for adverse events can be found at <https://zenodo.org/record/7604396#Y93irOyZM-Q>). We found no evidence of global or loop-specific incoherence (supplement table 16). The certainty in effects on adverse events for all 19 interventions was low or very low (fig 3).

### Additional analysis

We were unable to construct networks for mental functioning, role functioning, social functioning, or sleep quality due to the small numbers of trials reporting these outcomes. A summary of findings for these outcomes (all low certainty) is presented in supplement table 17.

We were unable to explore subgroup effects based on subtypes of TMD or receipt of disability benefits/

engaged in litigation versus not because of insufficient variability among eligible trials. We performed subgroup analysis based on the duration of follow-up and risk of bias and found significant subgroup effects for two comparisons: 1) removeable occlusal splint versus placebo, which found trials at high risk of bias were associated with larger treatment effects for pain relief; and 2) removeable occlusal splint versus usual care, which showed longer follow-up was associated with smaller treatment effects for pain relief (supplement table 18). The credibility of these subgroup effects proved to be moderate (supplement table 19).

There was no evidence of intransitivity based on the distribution of effect modifiers across comparisons. We found no evidence of small study effects for the three comparisons in which there were  $\geq 10$  trials available (supplement table 20).

## Discussion

This network meta-analysis of 153 RCTs that enrolled 8713 participants provided moderate certainty evidence that, compared with placebo/sham procedures, cognitive behavioural therapy (CBT) augmented with biofeedback or relaxation therapy, therapist-assisted jaw mobilisation, and manual trigger point therapy are among the most effective therapies for reducing pain in patients with chronic TMD pain. Moderate to high certainty evidence showed that five interventions were less effective, yet more effective than placebo for pain relief: CBT, supervised postural exercise, supervised jaw exercise and stretching with or without manual trigger point therapy, and usual care (such as home exercises, stretching, reassurance, self massage, and education). Moderate certainty evidence showed that, compared with placebo/sham procedures, supervised jaw exercise with stretching, manipulation, acupuncture, and supervised jaw exercise with mobilisation probably improve physical functioning. The certainty in effects, among other interventions, was low or very low.

## Strengths and limitations

Strengths of our review include a comprehensive synthesis of evidence on the benefits and harms of available pharmacologic and nonpharmacologic interventions for chronic pain associated with TMD. Independent coding of interventions by clinical experts blinded to study results provides reassurances regarding the composition of our network treatment nodes. Further, we used the GRADE approach to evaluate the certainty of evidence and ranked interventions using a minimally contextualised approach that considers both the magnitude of effects and certainty of evidence. Finally, we were able to translate the findings from our NMA into absolute effects, which are more understandable for knowledge users.

Our review also has several limitations. First, despite the large overall number of trials and participants, we found limited direct evidence to inform the effectiveness of several interventions versus placebo/

sham procedures, and the evidence to inform the effectiveness of most interventions proved to be of low or very low certainty, including all evidence to inform adverse events. Second, a key assumption of our review is that treatment effects would be similar across different TMD subtypes, and we were unable to explore for subgroup effects as most trials eligible for our review enrolled mixed subtypes or did not report the proportion of TMD subtypes among patients and reported aggregate results. However, we did not find any evidence of statistical variability across treatment effects in our outcome networks and in the assessments of between-study variances within the closed loops of evidence. Further, diagnostic criteria for TMD subtypes are largely subjective and patients often satisfy criteria for multiple subtypes.<sup>53 54</sup>

Third, because most eligible trials for our reviews excluded patients with chronic TMD pain who also had comorbid mental illness, fibromyalgia, or rheumatoid arthritis, or those who had previously undergone TMD surgery, the generalisability of our findings to these populations is uncertain. Fourth, although litigation and wage replacement benefits may influence treatment effects,<sup>55</sup> data in the included trials were insufficient to form conclusions regarding these issues. Fifth, based on meta-epidemiological studies<sup>22-25</sup> that found little to no effect of blinding on treatment results among randomised trials (including a review of oral health trials<sup>26</sup>), we did not rate down trials for risk of bias for interventions in which blinding is not possible and unblinding was the only risk of bias criterion not met; however, we could not directly test this assumption. Sixth, modelling assumptions for estimating the risk difference of achieving the minimally important difference assume that the standard deviations of outcome measures are the same in both the treatment and control groups, and that change scores in both groups are normally distributed. It is possible these assumptions were not met for some interventions. Seventh, although we planned to report treatment effects on all patient-important outcomes, interventional trials for chronic TMD consistently reported effects only on pain and, to a lesser extent, physical function. Most trials did not report effects on adverse events, and the certainty of evidence for the 19 interventions reporting on adverse events was low to very low. This is particularly concerning for invasive procedures and long term treatment with non-steroidal anti-inflammatory drugs plus opioids, which may be associated with important adverse effects.

## Relation to other studies

There have been nine prior network meta-analyses (NMAs) of interventions for chronic TMD pain.<sup>6-14</sup> None of them, however, assessed all available interventions across all TMD subtypes (supplement table 21). As such, they included only nine to 57 RCTs, versus the 210 unique trials considered in our review. Individual NMAs assessed needling therapies,<sup>10</sup> low level laser therapy,<sup>11</sup> inter-articular injections,<sup>13</sup> pharmacotherapy,<sup>9</sup> and different types of oral splints.<sup>6 12</sup> Of the three NMAs

## Abbreviations

- TMD temporomandibular disorders
- NMA network meta-analysis
- RCT randomised controlled trials
- CENTRAL Cochrane Central Register of Controlled Trials
- CINAHL Cumulative Index to Nursing and Allied Health Literature
- MID minimally important difference
- RD risk difference
- SD standard deviation
- WMD weighted mean difference
- OR odds ratio
- PRISMA Preferred Reporting Items for Systematic reviews and Meta-Analyses
- IMMPACT Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
- VAS visual analogue scale
- IQR interquartile range
- CBT cognitive behavioural therapy
- NSAIDs non-steroidal anti-inflammatory drugs
- SUCRA surface under the cumulative ranking

that considered a broader range of interventions, they limited patient populations to individuals with disc displacement,<sup>14</sup> arthrogenous TMD,<sup>7</sup> or masticatory muscle pain.<sup>8</sup> Findings were highly variable across reviews, and each NMA identified their most promising interventions based on SUCRA rankings. This approach is often misleading, as it considers only the point estimate of effect and not the associated precision or certainty of evidence and ignores absolute differences between treatment alternatives.<sup>56-58</sup>

A 2022 Cochrane review of 12 RCTs assessed treatment effects of psychological therapies for chronic TMD pain and found low certainty evidence from five trials that cognitive behavioural therapy (CBT) may reduce pain intensity at longest follow-up; however, the effect estimate was associated with substantial imprecision.<sup>59</sup> Our NMA has improved the precision associated with the effect of CBT for chronic TMD pain by augmenting direct evidence with indirect evidence.

## Conclusions

In this large NMA of randomised controlled trials, moderate certainty evidence or higher shows interventions that promote coping and encourage movement and activity were most effective for reducing chronic pain associated with TMD. Several interventions commonly in use for chronic TMD pain are supported by only low or very low certainty evidence. The accompanying *BMJ* Rapid Recommendation provides contextualised guidance based on this body of evidence (box 1).<sup>16</sup>

## AUTHOR AFFILIATIONS

<sup>1</sup>Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

<sup>2</sup>Department of Anesthesia, McMaster University, 1280 Main St. West, Hamilton, Ontario, Canada

<sup>3</sup>Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

<sup>4</sup>Health Technology Assessment Centre, Evidence Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou, China.

<sup>5</sup>Michael G. DeGroote Institute for Pain Research and Care, McMaster University, Hamilton, Ontario, Canada

<sup>6</sup>Institute for Management and Innovation, University of Toronto, Mississauga, Ontario, Canada

<sup>7</sup>Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada

<sup>8</sup>School of Medicine, Case Western Reserve University, Cleveland, OH, USA

<sup>9</sup>Department of Pediatrics, University of Antioquia, Medellin, Colombia.

<sup>10</sup>School of Rehabilitation Science, McMaster University, Hamilton, Canada.

<sup>11</sup>Pediatric Intensive Care Unit, Clínica Las Americas, Medellin, Colombia

<sup>12</sup>Institute of Health and Society, University of Oslo, Oslo, Norway

<sup>13</sup>Department of Otolaryngology-Head and Neck Surgery, Cleveland Clinic, Cleveland, OH, USA

<sup>14</sup>Michael G. DeGroote National Pain Centre, McMaster University, Hamilton, Ontario, Canada

<sup>15</sup>West China School of Nursing/West China Hospital, Sichuan University, Chengdu, China

<sup>16</sup>Department of Orthodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, China

<sup>17</sup>School of Biological Sciences, Nanyang Technological University, Singapore

<sup>18</sup>Division of General Internal Medicine, Department of Medicine, University Hospitals of Geneva, Geneva, Switzerland

<sup>19</sup>The MAGIC Evidence Ecosystem Foundation, Oslo, Norway

We thank Naz Torabi of Unity Health Toronto for peer-review of our search strategy; Randi McCabe (McMaster University), David Mock (University of Toronto), Carolyn Palmer (Veterans Affairs Canada), Christine Goertz (Duke University), and Alonso Carrasco-Labra (University of North Carolina) for categorising interventions, blinded to treatment results; and Zijun Li, Mina Ma, Ke Guo, and Minyan Yang of Lanzhou University for reviewing extracted data. Clinical experts on the *BMJ* Rapid Recommendation guideline panel were Rodrigo Casassus (orofacial pain specialist), Alonso Carrasco-Labra (general dentist), David Mock (oral and maxillofacial pathologist), Matteo Coen (general internist), Carolyn Palmer (general dentist), Joanna M Zakrzewska (oral physician), Justin Durham (honorary consultant oral surgeon), and Caroline Samer (clinical pharmacologist). No financial compensation was provided to any of these individuals.

**Contributors:** LY, BS, and MXL contributed equally to the study and are joint first authors. LY, BS, TA, and JWB conceived and designed the study. RJC searched electronic databases for eligible trials. LY, MXL, JL, QW, HC, GM, RM, BSH, JW, SM, YZ, MMA, YG, LC, LDZ, and IDF screened the studies and abstracted data. LY carried out the statistical analysis. YL, BS, and JWB interpreted the data. LY, BS, and JWB drafted the manuscript. All authors critically revised the article for important intellectual content and gave final approval for the article. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding:** This study was supported by a grant from the Chronic Pain Centre of Excellence for Canadian Veterans. JWB is supported, in part, by a Canadian Institutes of Health Research Canada Research Chair in Prevention and Management of Chronic Pain. The funding organisations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form and declare: no financial support from any industry for the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Not required.

**Data sharing:** Details of the characteristics of the included studies were shared in the supplementary materials. The study specific data included in the meta-analysis can be obtained from the first author at yaol12@mcmaster.ca.

**Transparency:** All authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.



**Dissemination to participants and related patient and public communities:** We used MAGICapp decision aids (available at [www.magicapp.org/](http://www.magicapp.org/)) to facilitate conversations between healthcare providers and patients. The MAGICapp decision aids were co-created with people living with chronic pain. We also plan to use social media, the websites of our organisations and pain related associations or societies to disseminate our findings.

- Durham J. Summary of Royal College of Surgeons' (England) clinical guidelines on management of temporomandibular disorders in primary care. *Br Dent J* 2015;218:355-6. doi:10.1038/sj.bdj.2015.194
- Greene CS. Managing the care of patients with temporomandibular disorders: a new guideline for care. *J Am Dent Assoc* 2010;141:1086-8. doi:10.14219/jada.archive.2010.0337
- Macfarlane TV, Glenny AM, Worthington HV. Systematic review of population-based epidemiological studies of oro-facial pain. *J Dent* 2001;29:451-67. doi:10.1016/S0300-5712(01)00041-0
- Sharma S, Ohrbach R. *Definition, aetiology, and epidemiology of temporomandibular disorders*. Handspring Publishing, 2018.
- Schiffman E, Ohrbach R, Truelove E, et al, International RDC/TMD Consortium Network, International association for Dental Research, Orofacial Pain Special Interest Group, International Association for the Study of Pain. Diagnostic criteria for temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the International RDC/TMD Consortium Network\* and Orofacial Pain Special Interest Group†. *J Oral Facial Pain Headache* 2014;28:6-27. doi:10.11607/jop.1151
- Al-Moraissi EA, Farea R, Qasem KA, Al-Wadeai MS, Al-Sabahi ME, Al-Iryani GM. Effectiveness of occlusal splint therapy in the management of temporomandibular disorders: network meta-analysis of randomized controlled trials. *Int J Oral Maxillofac Surg* 2020;49:1042-56. doi:10.1016/j.ijom.2020.01.004
- Al-Moraissi EA, Wolford LM, Ellis E3rd, Neff A. The hierarchy of different treatments for arthrogenous temporomandibular disorders: A network meta-analysis of randomized clinical trials. *J Craniomaxillofac Surg* 2020;48:9-23. doi:10.1016/j.jcms.2019.10.004
- Feng J, Luo M, Ma J, Tian Y, Han X, Bai D. The treatment modalities of masticatory muscle pain: a network meta-analysis. *Medicine (Baltimore)* 2019;98:e17934. doi:10.1097/MD.00000000000017934
- Häggman-Henrikson B, Alstergren P, Davidson T, et al. Pharmacological treatment of oro-facial pain - health technology assessment including a systematic review with network meta-analysis. *J Oral Rehabil* 2017;44:800-26. doi:10.1111/joor.12539
- Al-Moraissi EA, Alradom J, Aladashi O, Goddard G, Christidis N. Needling therapies in the management of myofascial pain of the masticatory muscles: A network meta-analysis of randomised clinical trials. *J Oral Rehabil* 2020;47:910-22. doi:10.1111/joor.12960
- Ren H, Liu J, Liu Y, Yu C, Bao G, Kang H. Comparative effectiveness of low-level laser therapy with different wavelengths and transcutaneous electric nerve stimulation in the treatment of pain caused by temporomandibular disorders: A systematic review and network meta-analysis. *J Oral Rehabil* 2022;49:138-49. doi:10.1111/joor.13230
- Alkhutari AS, Alyahya A, Rodrigues Conti PC, Christidis N, Al-Moraissi EA. Is the therapeutic effect of occlusal stabilization appliances more than just placebo effect in the management of painful temporomandibular disorders? A network meta-analysis of randomized clinical trials. *J Prosthet Dent* 2021;126:24-32. doi:10.1016/j.prosdent.2020.08.015
- Xie Y, Zhao K, Ye G, Yao X, Yu M, Ouyang H. Effectiveness of intra-articular injections of sodium hyaluronate corticosteroids, platelet-rich plasma on temporomandibular joint osteoarthritis: a systematic review and network meta analysis of randomized controlled trials. *J Evid Based Dent Pract* 2022;22:101720. doi:10.1016/j.jebdp.2022.101720
- Li J, Zhang Z, Han N. Diverse therapies for disc displacement of temporomandibular joint: a systematic review and network meta-analysis. *Br J Oral Maxillofac Surg* 2022;60:1012-22. doi:10.1016/j.bjoms.2022.04.004
- Siemieniuk RA, Agoritsas T, Macdonald H, Guyatt GH, Brandt L, Vandvik PO. Introduction to BMJ Rapid Recommendations. *BMJ* 2016;354:i5191. doi:10.1136/bmj.i5191
- Busse JW, Casassus R, Carrasco-Labra A, et al. Management of chronic pain associated with temporomandibular disorders: a clinical practice guideline. *BMJ* 2023;383:e076227. doi:10.1136/bmj-2023-076227
- Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777-84. doi:10.7326/M14-2385
- World Bank. World Bank Country and Lending Groups. 2022. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>.
- Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003;106:337-45. doi:10.1016/j.pain.2003.08.001
- Turk DC, Dworkin RH, Revicki D, et al. Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain. *Pain* 2008;137:276-85. doi:10.1016/j.pain.2007.09.002
- Higgins JP, Altman DG, Gøtzsche PC, et al, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928
- Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ* 2020;368:l6802. doi:10.1136/bmj.l6802
- Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One* 2016;11:e0159267. doi:10.1371/journal.pone.0159267
- Dechartres A, Trinquart L, Faber T, Ravaud P. Empirical evaluation of which trial characteristics are associated with treatment effect estimates. *J Clin Epidemiol* 2016;77:24-37. doi:10.1016/j.jclinepi.2016.04.005
- Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. *Am J Phys Med Rehabil* 2017;96:34-44. doi:10.1097/PHM.0000000000000521
- Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, da Costa BR, Flores-Mir C. Influence of blinding on treatment effect size estimate in randomized controlled trials of oral health interventions. *BMC Med Res Methodol* 2018;18:42. doi:10.1186/s12874-018-0491-0
- Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH. Pooling health-related quality of life outcomes in meta-analysis: a tutorial and review of methods for enhancing interpretability. *Res Synth Methods* 2011;2:188-203. doi:10.1002/jrsm.46
- Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions, version 5.1.0. 2020. <https://training.cochrane.org/handbook/archive/v5.1/>.
- Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005;5:13. doi:10.1186/1471-2288-5-13
- Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 2006;25:3443-57. doi:10.1002/sim.2380
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34. doi:10.1136/bmj.315.7109.629
- Brignardello-Petersen R, Murad MH, Walter SD, et al, GRADE Working Group. GRADE approach to rate the certainty from a network meta-analysis: avoiding spurious judgments of imprecision in sparse networks. *J Clin Epidemiol* 2019;105:60-7. doi:10.1016/j.jclinepi.2018.08.022
- Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012;3:98-110. doi:10.1002/jrsm.1044
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60. doi:10.1136/bmj.327.7414.557
- Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2012;101:447-59. doi:10.1198/016214505000001302
- Johnston BC, Alonso-Coello P, Friedrich JO, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ* 2016;188:25-32. doi:10.1503/cmaj.150430
- Schünemann HJ, Guyatt GH. Commentary--goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res* 2005;40:593-7. doi:10.1111/j.1475-6773.2005.0k375.x
- Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? *Acad Emerg Med* 1998;5:1086-90. doi:10.1111/j.1553-2712.1998.tb02667.x
- Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94:149-58. doi:10.1016/S0304-3959(01)00349-9
- Ward MM, Guthrie LC, Alba M. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care Res (Hoboken)* 2014;66:1783-9. doi:10.1002/acr.22392
- Brignardello-Petersen R, Florez ID, Izcovich A, et al, GRADE working group. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework. *BMJ* 2020;371:m3900. doi:10.1136/bmj.m3900



- 42 Busse JW, Sadeghirad B, Oparin Y, et al. Management of acute pain from non-low back, musculoskeletal injuries: a systematic review and network meta-analysis of randomized trials. *Ann Intern Med* 2020;173:730-8. doi:10.7326/M19-3601
- 43 Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ* 2020;192:E901-6. doi:10.1503/cmaj.200077
- 44 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD
- 45 Puhan MA, Schünemann HJ, Murad MH, et al. GRADE Working Group. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;349:g5630. doi:10.1136/bmj.g5630
- 46 Brignardello-Petersen R, Bonner A, Alexander PE, et al. GRADE Working Group. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol* 2018;93:36-44. doi:10.1016/j.jclinepi.2017.10.005
- 47 Brignardello-Petersen R, Tomlinson G, Florez I, et al. GRADE Working Group. Grading of recommendations assessment, development, and evaluation concept article 5: addressing intransitivity in a network meta-analysis. *J Clin Epidemiol* 2023;160:151-9. doi:10.1016/j.jclinepi.2023.06.010
- 48 Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol* 2011;64:1283-93. doi:10.1016/j.jclinepi.2011.01.012
- 49 Guyatt G, Oxman AD, Kunz R, et al. Corrigendum to GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol* 2011;64:1283-1293. *J Clin Epidemiol* 2021;137:265. doi:10.1016/j.jclinepi.2021.04.014
- 50 Santesso N, Glenton C, Dahm P, et al. GRADE Working Group. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020;119:126-35. doi:10.1016/j.jclinepi.2019.10.014
- 51 Nagori SA, Bansal A, Jose A, Roychoudhury A. Comparison of outcomes with the single-puncture and double-puncture techniques of arthrocentesis of the temporomandibular joint: an updated systematic review and meta-analysis. *J Oral Rehabil* 2021;48:1056-65. doi:10.1111/joor.13228
- 52 Boonstra AM, Schiphorst Preuper HR, Balk GA, Stewart RE. Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. *Pain* 2014;155:2545-50. doi:10.1016/j.pain.2014.09.014
- 53 Kraus SL. Characteristics of 511 patients with temporomandibular disorders referred for physical therapy. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2014;118:432-9. doi:10.1016/j.oooo.2014.06.005
- 54 Lobbezoo-Scholte AM, Lobbezoo F, Steenks MH, De Leeuw JR, Bosman F. Diagnostic subgroups of craniomandibular disorders. Part II: Symptom profiles. *J Orofac Pain* 1995;9:37-43.
- 55 Harris I, Mulford J, Solomon M, van Gelder JM, Young J. Association between compensation status and outcome after surgery: a meta-analysis. *JAMA* 2005;293:1644-52. doi:10.1001/jama.293.13.1644
- 56 Mbuagbaw L, Rochwerf B, Jaeschke R, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. *Syst Rev* 2017;6:79. doi:10.1186/s13643-017-0473-z
- 57 Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol* 2015;15:58. doi:10.1186/s12874-015-0060-8
- 58 Noori A, Sadeghirad B, Wang L, et al. Comparative benefits and harms of individual opioids for chronic non-cancer pain: a systematic review and network meta-analysis of randomised trials. *Br J Anaesth* 2022;129:394-406. doi:10.1016/j.bja.2022.05.031
- 59 Penlington C, Bowes C, Taylor G, et al. Psychological therapies for temporomandibular disorders (TMDs). *Cochrane Database Syst Rev* 2022;8:CD013515.

**Data supplement:** Appendices 1-3; supplementary tables 1-21