

BMJ Open Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis

Carolina Lavazza , Margherita Galli, Alessandra Abenavoli, Alberto Maggiani

To cite: Lavazza C, Galli M, Abenavoli A, *et al.* Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis. *BMJ Open* 2021;**11**:e045106. doi:10.1136/bmjopen-2020-045106

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045106>).

Received 22 September 2020
Revised 15 March 2021
Accepted 16 March 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Research, AIMO, Saronno, Italy

Correspondence to

Carolina Lavazza;
carolina.lavazza@docenti.aimoedu.it

ABSTRACT

Objective To assess the effects and reliability of sham procedures in manual therapy (MT) trials in the treatment of back pain (BP) in order to provide methodological guidance for clinical trial development.

Design Systematic review and meta-analysis.

Methods and analysis Different databases were screened up to 20 August 2020. Randomised controlled trials involving adults affected by BP (cervical and lumbar), acute or chronic, were included. Hand contact sham treatment (ST) was compared with different MT (physiotherapy, chiropractic, osteopathy, massage, kinesiology and reflexology) and to no treatment. Primary outcomes were BP improvement, success of blinding and adverse effect (AE). Secondary outcomes were number of drop-outs. Dichotomous outcomes were analysed using risk ratio (RR), continuous using mean difference (MD), 95% CIs. The minimal clinically important difference was 30 mm changes in pain score.

Results 24 trials were included involving 2019 participants. Very low evidence quality suggests clinically insignificant pain improvement in favour of MT compared with ST (MD 3.86, 95% CI 3.29 to 4.43) and no differences between ST and no treatment (MD -5.84, 95% CI -20.46 to 8.78).

ST reliability shows a high percentage of correct detection by participants (ranged from 46.7% to 83.5%), spinal manipulation being the most recognised technique.

Low quality of evidence suggests that AE and drop-out rates were similar between ST and MT (RR AE=0.84, 95% CI 0.55 to 1.28, RR drop-outs=0.98, 95% CI 0.77 to 1.25). A similar drop-out rate was reported for no treatment (RR=0.82, 95% CI 0.43 to 1.55).

Conclusions MT does not seem to have clinically relevant effect compared with ST. Similar effects were found with no treatment. The heterogeneity of sham MT studies and the very low quality of evidence render uncertain these review findings.

Future trials should develop reliable kinds of ST, similar to active treatment, to ensure participant blinding and to guarantee a proper sample size for the reliable detection of clinically meaningful treatment effects.

PROSPERO registration number CRD42020198301.

Strengths and limitations of this study

- This systematic review and pair-wise meta-analysis summarises existing evidence on the effect, reliability and application of hand contact sham treatment (ST) in manual therapy (MT) randomised controlled trials (RCTs).
- It gives suggestions for researchers on conducting methodical RCT in MT using a reliable sham procedure.
- Settings and practitioner influences on ST effects were not analysed due to lack of data.
- The number of studies included was insufficient to assess the impact of lack of blinding on ST effects.

BACKGROUND

In clinical trials (CT), a placebo is commonly used as a control therapy to evaluate the clinical effectiveness of the treatments tested.¹ Placebo has been defined as ‘an inert substance or sham procedure that is provided to research participants with the aim of making it impossible for them, and usually the researchers themselves, to know who is receiving an active or inactive intervention.’² Placebo interventions are methodological tools used to treat participants in the study arm and the control arm in exactly the same way, except that the study group receives an active substance and the control group does not.

In Europe, its use in pharmacological CT has been regulated by CT Regulation No. 536/2014. According to this regulation, placebo must be treated as an investigatory medical product and as such it has to meet certain standards in order to ensure quality, guarantee patient safety and the reliability of the study results.³

The regulatory aspects of trials involving manual therapies (MT) are very different. Although such studies might be influenced by the type of placebo provided, no clear

guidelines or regulations have been developed to ensure the credibility of trial results and patient safety.

MT is a clinical approach used by different physical therapists and involves hands-on techniques to manipulate, mobilise and massage the body tissues. This type of therapy can help relieve pain and stiffness, promote relaxation of soft-tissues, enhance blood supply to tissues and increase mobility of joint structures.⁴

In MT trials, placebo treatment is often provided in different modalities from trial to trial although the manual techniques or treatments tested are the same. A true placebo does not exist for MT and testing the effectiveness of MT requires a sham intervention. For instance, sham treatment (ST) is commonly administrated as a light touch in the site of pain or as an active treatment in a different site,⁵ with no clear criterion. Such light touch might in fact have a health effect and there is no evidence as to its ineffectiveness. Touch itself could have a positive outcome on health⁶ and active treatments could have an analgesic reflex on pain even if administered elsewhere in the body.⁷

Placebo effect, also called placebo response, is the reported improvement in symptoms among patients that occurs as a result of the placebo administration. Since a placebo has no inherent therapeutic power, it cannot cure the disease but it may contribute to the relief of patients' symptoms such as pain.⁸ Additionally, placebo might be related to an adverse effect (AE) called nocebo. It has been estimated that up to 26% of patients in randomised controlled trials (RCTs) discontinue placebo due to AEs.⁹

It is thought that these psychobiological phenomena may be related to the overall therapeutic context, such as treatment environment, individual patient and clinician factors (eg, beliefs, desire for symptom changes), as well as the patient's expectations of improvement and prior experiences of the treatment.^{10–13}

In pharmacological trials, this overall therapeutic context and its influence on placebo response has been widely studied.¹¹ Less evidence is present for MT trials, where the tactile interaction could be considered as an important characteristic of this therapeutic context.^{14 15} Pharmacological trials avoid the influence of clinicians' beliefs by using a placebo that ensures both patient and clinician blinding to treatment allocation, but, in MT trials, the blinding of clinicians is impossible to achieve. The best alternative in this type of trial is the use of an ST that mimics the active treatment and aims at blinding of participants.

Another important factor that has to be taken into account is that RCTs involving MT usually use patient-reported outcomes (PROs)—such as pain—as primary outcomes. Studies suggested that physical placebo treatments might have a greater effect on these types of outcome compared with pharmacological placebo and that this effect might be a consequence of physical contact.^{1 16 17}

Moreover, especially when subjective PROs outcomes are used, the absence of clinician blinding could also increase the possibility of performance bias.¹⁴

Therefore, a better understanding of sham procedures in manual treatment would be fundamental to define the real difference in efficacy between manual and ST, with a better knowledge of the effect of manual contact on PROs such as pain relief and drop-outs.

The role of placebo—referred to as sham therapy in this review—in MT trials is still very confused and the lack of guidelines allows huge discrepancies in its use in RCTs. Additionally, the reliability of sham procedures in MT trials has been rarely evaluated.

A clear definition of placebo effect could improve trial design, implementing studies with a proper power and sample size, defining clinical relevance of MT and giving more reliability to study results.

The aim of this systematic review with pairwise meta-analyses is to evaluate the use of ST in MT trials in order to analyse the effects, possible harm and the reliability of different kinds of sham procedures provided in RCTs involving MT. A systematic review could help to define ST standards to be applied in CT in order to guarantee methodological quality and patient safety.

Objective

To assess the benefits, potential harm and reliability of ST in MT RCTs in the treatment of back pain—both cervical and lumbar—in order to provide methodological guidance for CT development.

METHODS

This systematic review and meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).¹⁸

Criteria for considering studies for this review

Only RCTs were included in this review. Quasi-randomised trials in which allocation was not strictly random (eg, date of birth or toss of a coin) were excluded. No restrictions were applied to language or setting.

Studies were considered eligible if they included adult participants with acute or chronic back pain including coccyx, lumbar, dorsal and cervical. Trials where pain was related to muscular conditions, articular disorders (such as osteoarthritis) or spinal disc herniation were included.

Trials where musculoskeletal diseases were secondary to other pathologies (eg, amyotrophic lateral sclerosis, fibromyalgia, etc) were excluded.

Trials where pain was related to fracture, surgery, dysmenorrhoea, post partum or pregnancy, headache or dizziness were excluded.

This review involved all types of ST that include hand contact provided by all kinds of physical therapists. Studies where ST was provided by machines (such as inactive ultrasound) were excluded. This choice was based on the fact that many MT used detuned ultrasound as

control. This type of sham was not considered adequate for MT trials where active treatment is provided by hand contact. Therefore, these studies were excluded.

All trials that involved hand contact ST as light touch or a manual treatment in a different site were included.

ST was compared with other MT provided by any type of healthcare provider such as: physiotherapist, chiropractor, osteopath, massage therapist, kinesiologist and reflexologist.

To assess if touch itself could have a positive health effect, ST was also compared with no treatment. Physiotherapeutic exercises were included in the analysis only if associated with manual treatment.

The use of active cointerventions such as oral Nonsteroidal anti-inflammatory drugs (NSAIDs) or other active treatments was accepted if used in all trial arms. Trials with more than two arms of intervention were included, but only data from interested arms were extracted.

Outcomes

Primary outcomes were pain intensity on a validated scale, success in the blinding of participants and AE. Secondary outcomes were number of drop-outs.

Whenever the meta-analysis could not be performed, a narrative summary of the outcomes has been provided. Outcomes were divided into short (≤ 2 months), medium (≤ 4 months) and long term (≥ 6 months). Data were extracted and analysed based on the time closest to these intervals.

Information sources

Search strategy (online supplemental appendix 1) was adapted to the different databases by an experienced information specialist.

RCTs were identified in different databases (up to 20 August 2020): MEDLINE, Embase, CINAHL, SPORT-Discus, PEDro, WHO Clinical Trials Registration Platform, Index to Chiropractic Literature, Cochrane central register of controlled trials, Clinical trials registry and metaRegister of Controlled Trials.

Researchers of unpublished trials, but completed and registered, were contacted by CL to obtain data.

The search in PROSPERO, in the Cochrane Library and in PubMed (clinical queries) was performed to evaluate the presence of ongoing or recently completed systematic reviews. Guidelines from different organisations (eg, National Council for Osteopathic Research, etc) were reviewed and references from relevant publication were analysed.

Data collection and analysis

Search results were screened by two independent reviewers who identified all the potentially eligible trials based on title and abstract. Full texts of all the selected articles were screened first for inclusion. If full text was not available, or the trial was completed but not published, CL contacted the authors in order to obtain the information

needed or used the document delivery service of the 3Bi Biella library.

Uncertainty about the inclusion of a study was discussed by the two reviewers. If no agreement was reached by the two reviewers a third reviewer (AM) was asked for their opinion.

The selection process was recorded and reported through a PRISMA flow diagram.

Data extraction and management

Data extraction was performed by two reviewers with a tested predefined form. Data extracted were related to settings, type of study, participants characteristics (such as localisation and duration of pain, pain score at baseline, previous similar treatment), interventions, outcomes used in the meta-analysis and other relevant data such as difference in ST and active treatment or funding (online supplemental appendix 2).

Risk of bias in individual studies

Bias risk was assessed by CL and agreed by MG using the Cochrane Risk of bias (CRB) tool.¹⁹ This tool was used to assess selection bias, performance bias, attrition bias, reporting bias and other biases.

Each possible risk was evaluated as 'high', 'medium' or 'low' by CL and a revision of the judgements was performed by MG. RevMan V.5.3.5 was used for the graphic representation of each risk. The CRB tool results were then converted to Agency for Healthcare Research and Quality (AHRQ) Standards to assess the quality of the study (good, fair and poor). Trials were judged as good quality when bias risk was judged as low, studies with fair quality were trials where at least one criterion was high risk, while poor-quality studies were trials with two or more criteria with high or unclear risk.

Assessment of reporting biases

Funnel plots were created to explore reporting bias, whenever more than 10 studies were included in the meta-analysis. Furthermore, for each study, an analysis of possible conflicts of interest and funding sources was performed.

Summary measures

Dichotomous outcomes, such as AE (occurred or not), were analysed using risk ratio (RR) with 95% CIs.

Continuous outcomes, such as back pain on Visual Analogue Scale (VAS), were evaluated using mean difference (MD) between ST and the MT/no treatment group with 95% CI and the SD.

The minimal clinically important difference (MCID) between pretreatment and post-treatment was taken as 30 mm changes in 100 mm pain score.^{20–22} These values were used for the interpretation of the clinical significance of the findings.

Success of blinding was reported with a percentage of patients guessing correctly the treatment allocation.

In this review, the unit of analysis was the participant.

Assessment of heterogeneity

The presence of heterogeneity was assessed with a visual inspection of the forest plots and through an inconsistency level test (I^2).

Cochrane Handbook was used for threshold interpretation: heterogeneity was considered as unimportant for values of I^2 between 0% and 40%, as moderate for values between 30% and 60%, as substantial for values between 50% and 90% and considerable for values between 75% and 100%.²³

Synthesis of results

Meta-analysis of pain score, AE and drop-out rates were performed using RevMan V.5.3.5 whenever possible. The meta-analyses compared all kinds of ST with all types of MT and to no treatment. Random effect model was used when a substantial inconsistency was present ($I^2=50\%-90\%$).²⁰ When considerable heterogeneity was present ($I^2>75\%$) and could not be explained by clinical or methodological diversity, the results have been presented narratively.

The statistical significance of measured effects was determined evaluating the p value and 95% CI.

Additional analyses

Different subgroup analyses were planned in the protocol such as on ST type provided (applied locally or in different sites from pain), type of manual technique tested (single or multiple techniques) and localisation of back pain. However, due to the small number of studies included in this review, only a few subgroup analyses were conducted on follow-up periods.

Sensitivity analysis was conducted for the primary outcomes to assess the effects of skewed and imputed data on the effect measure. These analyses are reported as online supplemental appendices.

Summarising results and assessing the quality of the evidence

The quality of evidence for each outcome was evaluated with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach by two independent authors and any disagreement was discussed. The quality for each effect measure was judged as high, moderate, low or very low.¹⁹ The GRADE approach was used to assess the quality of the key outcomes. The software GRADEpro (<https://grade.pro.org>) was used to import data from RevMan V.5.3.5 and to create 'summary of findings tables'.

The following outcomes were chosen to be presented: pain scores at short term, AE and drop-outs.

Patient and public involvement

There was no involvement of patients or public during the outline of this project. The differences noted between therapies tested on primary pain outcome were those clinically meaningful to patients.

RESULTS

Included studies

Table 1 shows a summary of the main characteristics of included studies. 24 studies were included in this review (figure 1), one study had a 2x2 factorial design,²⁴ eight studies had multiple arms.^{25–32} Most of the studies were conducted in physical therapy clinics, in 13 different countries. Three trials did not report in which clinical setting they were conducted.^{29 33 34}

Eight trials were conducted in Europe,^{27 28 30 35–39} five in the USA,^{24 25 31 40 41} three studies in Brazil,^{42–44} one in the UK,²⁶ Egypt,³² Japan⁴⁵ and Australia.⁴⁶

No ongoing or unpublished trials were found.

Population

The included trials randomised a total of 2019 participants, the majority of studies (N=18) were small with a median of 50 participants and a range from 15 to 455.

Most trials included middle aged patients (mean 39.9 range from 18 to 73) with a mean BMI of 21.7 kg/m².

The majority of studies included both genders, with a percentage of male that ranged from 19% to 80%. Two trials included only male,^{38 44} one study included only female participants.⁴²

Sixteen trials enrolled participants with low back pain (LBP), eight included participants with cervical pain.^{26 33 35–37 40–42}

The majority of trials (N=18) included participants with unspecified cause of back pain. Disk herniation was considered in three trials.^{27 30 44}

Duration of symptoms were unassessed in eight trials, nine studies included participants with chronic pain, some included participants with both acute and chronic pain.

Participants with experience of the tested treatment were included in eight trials^{24 29 31 32 35 37 42 43} and excluded in four.^{26 36 39 41} The remaining studies did not provide this information.

Interventions

Interventions deferred for number of sessions and number of techniques applied. Eleven trials used a single therapy session with a single technique performed in eight of those trials. Trials with different therapy sessions ranged from 5^{25 26 30} to 20²⁷ sessions once a week.

Sham treatment

ST was provided by a hand contact on the area of pain in 19 studies, and five studies provided ST in a different area from where the pain was located.^{27 35 43 45 46}

In trials providing spinal manipulation (SM), as inactive treatment the majority of authors used the similar placement of hands on participants without any force applied.^{40–42 44} Two trials used an ST with similar forces applied in different directions.^{25 32} One trial did not specify the inactive manipulation applied.²⁹

In trials that provided multiple techniques in the same treatment session (such as osteopathic treatment, spinal

Table 1 Summary of main characteristics of included studies

Study ID	No of participants	Symptoms duration	Pain localisation	Technique tested (site of application)	Type of sham procedure	Other arms	Follow-up
Antonilos-Campillo, 2014 ³⁵	40	Not reported	Cervical	Soft-tissue (cervical region)	Soft mobilisation of lower limbs	None	No follow-up (outcomes collected after the intervention)
Bialosky, 2014 ²⁵	110	>4 months	Lumbar	Spinal manipulation (SM) (lumbar spine)	Ineffective force applied on lumbar spine	No treatment group	2 weeks
Cleland, 2005 ⁴¹	36	>2 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the intervention)
Eardley, 2013 ²⁶	58	>3 years	Lumbar	Kinesiology (spine)	Protocol of ineffective techniques in the site of pain	No treatment group	7 weeks
Erdogmus, 2007 ²⁷	120	≥1.4 weeks	Lumbar	Physiotherapy (exercises of the spine and articular mobilisation techniques)	Neck massage	No treatment group	1.5 years
Hall, 2006 ⁴⁶	24	Not reported	Lumbar	BLR technique (lower limbs)	Soft-tissue manipulation of the foot	None	24 hours
Haller, 2016 ³⁶	54	>7 months	Cervical	Cranio-sacral therapy (head)	Ineffective touch of head	None	3 months
Hansen, 1993 ²⁸	168	≥18 days	Lumbar	Physiotherapy (exercises of lumbar spine and abdomen, soft-tissues techniques)	Intermittent traction of the spine	Intensive back muscle training	1 year
Hidalgo, 2015	32	Not reported	Lumbar	Articular mobilisation (lumbar spine)	Ineffective mobilisation forces applied on lumbar spine	None	2 weeks
Hoiris, 2004 ²⁹	156	≥2,3 weeks	Lumbar	SM (spine)	Ineffective force applied on spine	Medical treatment	4 weeks
Klein, 2013 ³⁷	61	>1 month and <5 years	Cervical	Strain-counterstrain techniques (cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the intervention)
Kogure, 2015 ⁴⁵	179	>12 months and <10 years	Lumbar	Arthrokinematic Approach-Hakata (AKA-H) (sacro-iliac joint)	Ineffective force applied on sacro-iliac joint	None	6 months

Continued

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

Table 1 Continued

Study ID	No of participants	Symptoms duration	Pain localisation	Technique tested (site of application)	Type of sham procedure	Other arms	Follow-up
Krekoukias, 2017 ³⁰	50	Not reported	Lumbar	Articular mobilisation techniques (lumbar spine)	Hand contact with lumbar skin placement without any movement	Exercise plus Transcutaneous Electrical Nerve Stimulation (TENS)	5 weeks
Lascurain-Aguirrebena, 2018 ³³	40	Not reported	Cervical	Articular mobilisation (Cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the intervention)
Licciardone, 2003 ³¹	91	≥3 months	Lumbar	Osteopathic manual treatment (OMT) (all body)	Protocol of light touch techniques similar to OMT applied to all body	No treatment group	6 months
Licciardone, 2013 ²⁴	455	≥3 months	Lumbar	OMT– (all body)	Protocol of light touch techniques similar to OMT applied to all body	None	8 weeks
Pires, 2015 ⁴²	32	>3 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	72 hours
Quinn, 2008 ³⁹	15	Not reported	Lumbar	Reflexology (foot)	Foot massage with less pressure and in different reflex point (not related to the spine)	None	18 weeks
Selkow (2009) ³⁴	20	1–6 weeks	Lumbar	Muscular energy technique (anterior superior iliac spine and lower limbs)	Practitioner hand positioned as active treatment, but participant rested for 30 s without any active contraction	None	24 hours
Senna, 2011 ³²	93	≥13 months	Lumbar	SM (lumbar spine)	Ineffective force applied on lumbar spine	Maintained SM	10 months
Sillevis, 2010 ⁴⁰	100	≥23 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the intervention)

Continued

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

Table 1 Continued

Study ID	No of participants	Symptoms duration	Pain localisation	Technique tested (site of application)	Type of sham procedure	Other arms	Follow-up
Silva, 2019 ⁴³	28	>3 months	Cervical	Osteopathic visceral treatment (abdomen)	Hand contact on umbilical region without any movement	None	7 days
Veira-Pellez, 2014 ⁴⁴	40	Not reported	Lumbar	SM (lumbar/sacral spine)	Ineffective force applied on lumbar/sacral joints	None	No follow-up (outcomes collected after the intervention)
Younes, 2017 ³⁸	17	<3 months	Lumbar	OMT (all spine)	Placebo mimed active treatment with an ineffective force applied.	None	7 days

mobilisation and physiotherapy) the ST was administered with different techniques that mimed active treatments using light touch or light tractions.

Only one trial compared one single sham technique with both single active technique and multiple treatment techniques. In this case only data of the first arm were extracted.³⁷

Manual and controls treatments

Different manual treatments were provided:

- SM/chiropractic (7 studies, 567 participants).
- Osteopathy (5 trials, 645 participants).
- Kinesiology (1 trial, 58 participants).
- Articular mobilisations (6 trials, 445 participants).
- Muscular release (5 trials, 304 participants).

Four trials with multiple arms compared ST to no intervention (379 participants)^{25–27 31} and one to muscle relaxant group (156 participants).²⁹

The manual treatment was generally applied in the area of pain, some trials used techniques additionally in other areas. Just one trial using reflexology provided both MT and sham in a different zone.³⁹

Characteristics of the practitioner who administrated treatments were provided by 16 trials. Trials involved physiotherapists (N=8), physical therapists (N=4), osteopaths (N=3) and osteopathic students (N=1). Only seven studies provided information on years of practice experience of the physicians involved ranging from 6 to 17 years.^{30 33 35–37 40 42 44} The gender of practitioners was indicated in only three trials.^{26 30 37}

Risk of bias in included studies

Figure 2 shows risks of bias.

Blinding of participants and assessors will be described due to the nature of this review.

According to AHRQ standards of CRB tool,¹⁹ the majority of trials were judged as poor quality (N=22). Good quality was conferred on only two studies.^{36 45}

The random sequence and allocation concealment were adequately reported in 71% and 63% of trials, respectively.

The lack of blinding of participants was the most common bias and was judged as high risk in 38% of studies, while 38% were considered as unclear risk.

The reasons for this judgement were mainly related to trials involving SMs. These studies used a technique which can be easily recognised by patients as active treatment for the popping sound emitted by joints. Additionally, these trials involved participants who could have already received this type of treatment, making the masking of technique almost impossible.

Blinding of outcomes was evaluated mainly as unclear risk in 46% of trials. Only two trials reported the strategies adopted to guarantee assessor blinding.^{28 32}

Incomplete outcome data were the least common bias risk with 80% of trials judged as low risk. Reporting bias was evaluated as unclear in 55% of trials where registration number and trial protocol were not reported or found.

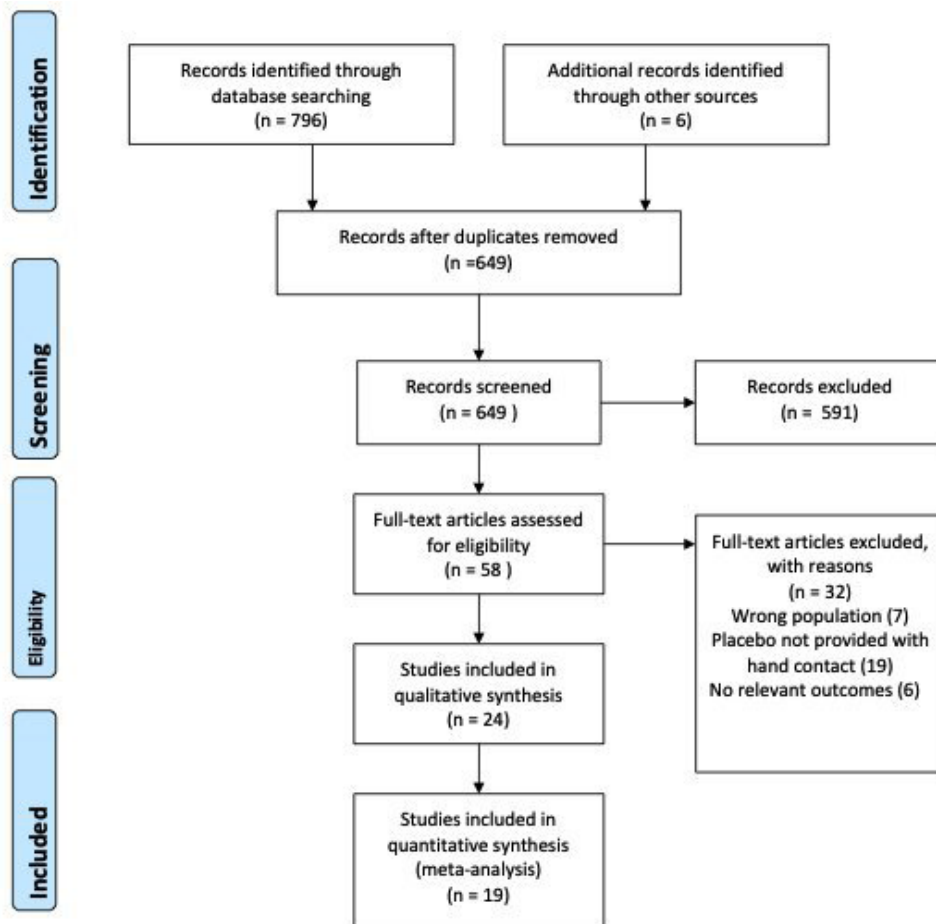


Figure 1 PRISMA flow diagram. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Arnold-Campillo PJ 2014	?	?	?	?	?	?	?
Blakosky J 2014	?	?	?	?	?	?	?
Cleland JA 2005	?	?	?	?	?	?	?
Fardley S 2013	?	?	?	?	?	?	?
Erdogmus C 2007	?	?	?	?	?	?	?
Haller H 2016	?	?	?	?	?	?	?
Hall T 2004	?	?	?	?	?	?	?
Hansen F 1993	?	?	?	?	?	?	?
Hidalgo B 2015	?	?	?	?	?	?	?
Hoiris K 2004	?	?	?	?	?	?	?
Klein R 2013	?	?	?	?	?	?	?
Kogure A 2015	?	?	?	?	?	?	?
Krekoulas G 2017	?	?	?	?	?	?	?
Lascurain-Aguirrebeña I 2018	?	?	?	?	?	?	?
Liccardone J 2013	?	?	?	?	?	?	?
Liccardone J 2003	?	?	?	?	?	?	?
Pires FP 2015	?	?	?	?	?	?	?
Quinn F 2008	?	?	?	?	?	?	?
Sellow NM 2009	?	?	?	?	?	?	?
Senna MK 2011	?	?	?	?	?	?	?
Silveira R 2010	?	?	?	?	?	?	?
Silva A 2018	?	?	?	?	?	?	?
Veira-Pérez F 2014	?	?	?	?	?	?	?
Younes M 2017	?	?	?	?	?	?	?

Figure 2 Risk of bias summary. Review authors' judgements about each risk of bias item for each included study.

Other bias occurred was generally considered as high risk for baseline differences of the population in 30% of trials.

EFFECTS OF INTERVENTION

Table 2 summaries treatment effects and GRADE quality of the evidence for all comparisons.

ST versus other MTs

Pain

The following outcomes on back pain are presented with a 100 mm VAS, 0–100; higher scores refer to worse pain. Trials using a 10 mm scales were converted to 100 mm scores.

The comparison between ST and MT was performed in 17 studies. One trial used a different scale and data were obtained with a conversion formula.²⁷ Data from seven studies could not be extracted.

The meta-analysis at short term showed substantial heterogeneity levels using a random effects model. To further investigate inconsistency levels, a sensitivity analysis excluding two trials was performed. One trial used a different validated scale,²⁵ while the other was suspected of publication bias.³⁰ This thought was verified with a funnel plot, which showed an asymmetric

Table 2 Summary of findings of treatment effects and certainty of the evidence (GRADE) included for all comparisons

1. Sham treatment (ST) compared with manual therapies (MT)

Patient or population: back pain Intervention: ST Comparison: MT		Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
Outcomes		Risk with MT	Risk with ST				
Pain improvement assessed with: VAS score§ Scale from: 0 to 100			MD 3.86 higher (3.29 lower to 4.43 higher)	–	805 (15 RCTs)	⊕○○○ VERY LOW†‡	A small effect, not clinically relevant, in pain improvement was detected in favour of MT. This analysis excluded two trials (one suspected of publication bias, one used a different scale) which increased heterogeneity levels but did not affect overall efficacy meaningfully.
Adverse events assessed with: no of AE occurred	144 per 1.000	121 per 1.000 (79 to 184)		RR 0.84 (0.55 to 1.28)	531 (6 RCTs)	⊕⊕○○ LOW†	Pooled data from six studies did not show any difference in AE occurrence between ST and MT.
Drop-outs rate assessed with: no of participants that leaved the study	174 per 1.000	171 per 1.000 (134 to 218)		RR 0.98 (0.77 to 1.25)	1238 (11 RCTs)	⊕⊕○○ LOW†	Pooled data from 11 trials did not show difference in drop-out rate between ST and MT.

2. ST compared with no treatment

Patient or population: back pain Intervention: ST Comparison: No treatment		Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
Outcomes		Risk with no treatment	Risk with ST				
Pain improvement assessed with: VAS score Scale from: 0 to 100			MD 5.84 lower (20.46 lower to 8.78 higher)	–	177 (3 RCTs)	⊕○○○ VERY LOW†‡¶	Pooled data from three trials, highly inconsistent, showed no differences between ST and no treatment group in pain improvement.
Drop-outs rate assessed with: no of participants that leaved the study	150 per 1.000	123 per 1.000 (65 to 233)		RR 0.82 (0.43 to 1.55)	225 (4 RCTs)	⊕○○○ VERY LOW†¶	Very low quality of evidence suggests no differences in drop-out rate between ST and no treatment.

GRADE Working Group grades of evidence.

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

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

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

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

*The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

†The majority of trials were judged as poor quality according to AHRQ standards.

‡Most of the studies were small trial

§Heterogeneity levels at 85%.

¶Number of participants <400.

AE, adverse effect; GRADE, Grading of Recommendations Assessment, Development and Evaluation; MD, Mean difference; RCT, randomised controlled trials; RR, Risk ratio; VAS, Visual Analogue Scale.

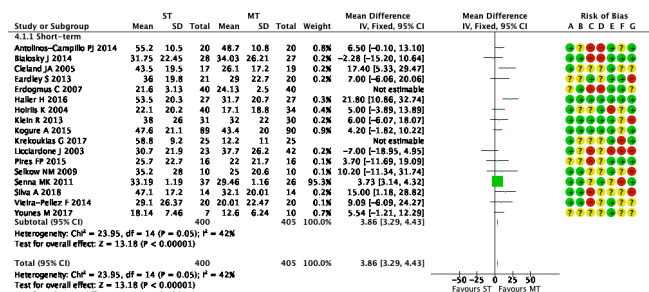


Figure 3 Forest plot of comparison ST versus MT in back pain outcome at short term. MT, manual therapy; ST, sham treatment.

distribution with the inclusion of these two studies (online supplemental appendix 3). This sensitivity analysis did not influence overall effectiveness results, but inconsistency levels decreased considerably at short term. It can be deduced that a possible cause of heterogeneity was found (full analysis in online supplemental appendix 4).

The sensitivity analysis using a fixed model at short term showed a slight difference, not clinically meaningful, between ST and MT in favour of MT on pain outcome (MD 3.86, 95% CI 3.29 to 4.43, 805 participants, $I^2=42\%$, $p<0.0001$, very low quality of evidence downgraded two levels for very serious risk of bias and imprecision) (figure 3).

Comparisons between ST and MT at medium and long term could not be performed due to substantial levels of heterogeneity found using a random effects model. The heterogeneity levels were not explainable by clinical or methodological diversities within trials (medium-term $I^2=91\%$ $P<0.0001$, long-term $I^2=81\%$ $p=0.005$) (online supplemental appendix 4.1).

Success of blinding

Success of blinding was evaluated in five trials; one did not report the results.³⁰

Patients were asked to assess if they understood their treatment allocations. Due to the type of data extracted (percentage of correct guessing) meta-analysis was not performed and results are reported descriptively.

Two trials compared ST with SM, these trials showed a correct perception of treatment allocation that ranged from 63.5%²⁵ to 83.5%.²⁹ In this last study, patients were considered eligible if they had already received SM.

One trial compared ST to an articular mobilisation technique. 54.5% participants correctly guessed the treatment allocation.⁴⁶

Participants of one study that compared ST to reflexology had the lowest percentage of correct detection of allocation (46.7%). Participants in this trials did not know about the type of treatment tested.³⁹

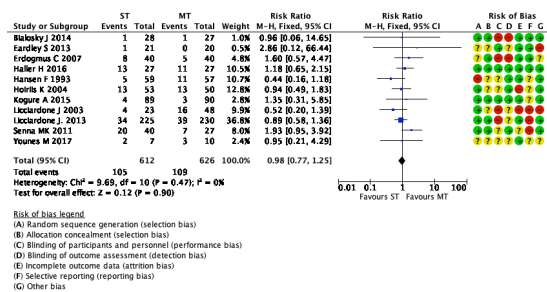


Figure 4 Forest plot of comparison ST versus MT in number of drop-outs outcome. MT, manual therapy; ST, sham treatment.

Drop-outs

Pooled data from 11 trials at the last follow-up suggested no difference in drop-outs rate between ST and MT at the end of the trials (105/612 compared with 109/626; RR 0.98, 95% CI 0.77 to 1.25; 1238 participants, $I^2=0\%$, $p=0.90$; low quality of evidence downgraded two levels for high risk of bias) (figure 4).

Adverse effects

AEs were generally under-reported, six trials were included in the meta-analysis.^{26–28 36 37 45}

Two trials reported AE overall occurrence without specified event rates in the groups.^{24 32}

AEs were predominantly minor and lasted for 2/3 days after treatment, in the majority of trials transient worse pain, tiredness, muscle weakness and transient headache were reported.^{26 36 37 45}

Senna and Machaly reported the most common AEs were local discomfort and tiredness but no serious complications were noted.³²

Haller *et al* reported two patients dropping out from the trial for recurrent headache after treatments, both Haller *et al* and Klein *et al* reported dizziness of one patient.^{36 37}

Licciardone *et al* reported 27% of patients with AE, 2% had serious AE not related to study interventions.²⁴

Overall results showed no clear difference in AE occurrence between ST and MT (32/267 compared with 38/264; RR 0.84, 95% CI 0.55 to 1.28; 531 participants, $I^2=26\%$, $p=0.42$; low quality of evidence downgraded two levels for inconsistency) (figure 5). Senna and Licciardone were excluded from analysis because they did not provide separate data for each group.

Sham versus no treatment

Pain

Four studies compared ST to no intervention, three were included in random effect meta-analysis at short term.^{25–27 29} Data from one trial could not be extracted.³¹

Pooled data showed the presence of significant heterogeneity, therefore results are reported narratively: two trials showed no difference between ST and no treatment on pain outcome, while Eardley *et al* showed an effect in favour of ST (pooled data from three trials: MD -5.84, 95% CI from -20.46 to 8.78, 252 participants, $I^2=85\%$, $p=0.43$). The exclusion of Erdogmus *et al* (that used a

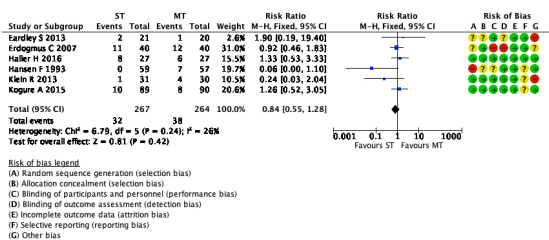


Figure 5 Forest plot of comparison ST versus MT in number of adverse events outcome at short term. M-H, Mantel-Haenszel; MT, manual therapy; ST, sham treatment.

different scale) did not affect the results of effectiveness neither decreased levels of heterogeneity (MD -10.83, 95% CI -32.44 to 10.79, $I^2 = 84\%$, $p = 0.33$) (online supplemental appendix 5).

Drop-outs

No differences were shown in the fixed-effect meta-analysis on drop-out rate between ST and no intervention in four trials (14/112 compared with 17/113; RR 0.82, 95% CI 0.43 to 1.55; 225 participants, $I^2 = 0\%$, $p = 0.54$; very low quality of evidence downgraded two levels for very serious risk of bias and imprecision) (figure 6).

Adverse effects

Of the four studies comparing ST to no intervention, only two reported AE.

One, Eardley *et al* did not evaluate AE occurred in the no treatment group while Erdogmus *et al* reported that 10/40 in the no intervention group and 11/40 in ST group turned to other therapies for complains.

DISCUSSION

Results show a small, not clinically meaningful effect in favour of MT for short-term pain relief compared with ST. However, the quality of evidence is very low, suggesting that the true effect may be different from the estimated effect. Substantial levels of heterogeneity within the four studies analysed showed no differences between ST and no treatment in pain reduction.

Success of blinding was reported in four trials that compared ST to MT, with a high percentage of correct detection of treatment allocations by participants.

AEs were generally under-reported, with a similar rate of occurrence between sham and MT accompanying low levels of heterogeneity. Only one study reported AE in its

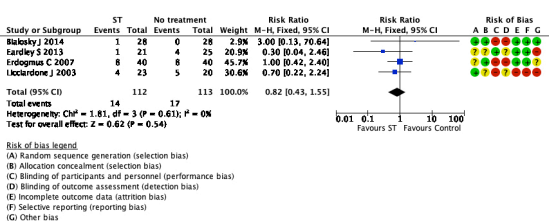


Figure 6 Forest plot of comparison ST versus no treatment in number of drop-outs outcome. M-H, Mantel-Haenszel; ST, sham treatment.

no treatment group with no significant difference from ST.

SM techniques were the treatment most evaluated ($N = 7$). These techniques are highly recognisable by patients for a popping sound emitted by the column during their performance.⁴⁷ The fact that participants enrolled in these trials were eligible despite having already received SM, threatens the validity of blinding. This thought is strengthened by the high percentage of participants who recognised treatment allocation in this kind of trial (from 63.5% to 83.5%).^{25 29} Additionally, five trials applied ST in a different site compared with pain and active treatment. This might have had important influences on sham therapy reliability and consequently to study results.

Lack of blinding seemed not to be related to drop-outs rate, although both these data were reported only in two trials Bialosky *et al* and Hoiriis *et al* showed high percentages of correct treatment allocation detection by participants but drop-out rate between sham and MT group did not differ.^{25 29} These results seem to be in conflict, nevertheless, participants could have wanted to remain in the trial for several other reasons such as the setting or the attraction of being evaluated by an expert clinician free. This possibility is reinforced by the fact that a similar drop-out rate was reached in the comparison sham versus no treatment. These data suggest that drop-out rate might not be a dependable outcome for assessing reliability of ST.

Another factor that seemed to put blinding validity at risk was the use of a single technique. Single techniques were generally more difficult to mask, negatively affecting the validity of blinding of participants. The majority of trials judged as at high or unclear risk of performance bias used a single technique evaluating its effects on pain soon after its performance, or its effect after different sessions.

When compared with no intervention, ST showed no effect. Only one study of the four included in the meta-analysis showed a statistically significant effect in favour of ST. This study was the only one judged at low risk of performance bias because researches tried to mask ST performing techniques very similar to MT and excluding participants that already received the treatment tested.²⁶ This trial was the one that showed a marked effect on pain (MD -21.7, 95% CI -33.5 to -9.9, 42 participants) (online supplemental appendix 5). Other studies included in this comparison, judged at high risk of performance bias, showed no effect of ST. These results suggest that lack of blinding could have had an impact on this comparison.

This review included generally small trials. Only 14 of 24 studies performed a sample size calculation but just two of these considered MCID in this computation. The MCID is the measure of smallest change of PROs that patients perceive as important, beneficial or harmful. MCID is useful for clinicians to interpret the findings of trials and apply them in clinical practice and to their decision making.⁴⁸ An adequate sample size calculation, using

MCID especially in trials with PROs, is fundamental to assess the number of participants needed to detect clinically relevant treatment effects. Oversized trials, which expose too many people to unnecessary therapies, or underpowered trials, which may not achieve significant results, should be avoided.^{49–51}

Comparison with other studies

Similar findings were found in other reviews conducted on LBP. Ruddock *et al* included studies where SM was compared with what authors called ‘an effective ST’, namely a credible sham manipulation that physically mimics the SM. Pooled data from four trials showed a very small and not clinically meaningful effect in favour of MT.⁵²

Rubinstein *et al*⁵³ compared SM and mobilisation techniques to recommended, non-recommended therapies and to ST. Their findings showed that 5/47 studies included attempted to blind patients to the assigned intervention by providing an ST. Of these five trials, two were judged at unclear risk of participants blinding. The authors also questioned the need for additional studies on this argument, as during the update of their review they found recent small pragmatic studies with high risk of bias. We agree with Rubinstein *et al* that recent studies included in this review did not show a higher quality of evidence. The development of RCT with similar characteristic will probably not add any proof of evidence on MT and ST effectiveness.⁵³

Limitations

This review aimed to compare different kinds of sham therapy with different kinds of MT and no intervention. The nature of this comparison needed an NMA, but this analysis could not be performed due to the small number of trials using hand contact ST. The decision to include only this kind of sham therapy was mainly due to the intention of analysing the effect of manual interaction between practitioner and patients, which is suspected of leading to an amplified placebo effect.⁵⁴ Additionally, the use of machine placebo trials in the same meta-analysis could have increased diversity within included trials due to the possible enhanced presence of biases such as performance and consequently detection ones.

Although the population differed—some trials analysed cervical, others lumbar pain with different aetiologies and different symptoms duration—this factor did not affect the meta-analysis performed, as highlighted by the low heterogeneity found in the primary outcome.

As already suggested by other authors,¹ placebo effect might be influenced by chronic pain, nevertheless, in this review, this analysis could not be performed due to the range of pain duration in trials included (from acute to chronic in the same trial).

Data concerning settings and operators were insufficient to evaluate the influence of these two factors on sham therapy response. Experience of practitioners was

considered in data extraction but insufficient information was provided by authors to draw any hypothesis.

Another limit was in not considering non-objective outcomes as primary outcome for meta-analysis. Nevertheless, most of the trials included did not evaluate an objective outcome and the few studies which analysed this type of outcome used different kinds of scales not easily comparable in a meta-analysis.

Pairwise comparison on pain outcome between sham and MT showed slightly higher effects of MT in trials where blinding was ensured. A linear regression analysis was planned to assess the impact of blinding on meta-analysis results. Due to the small number of trials, this analysis could not be performed. This trend follows what has been already suggested by other studies.⁵⁵ However, trials with bigger sample size are needed to assess a real correlation between these two factors.

Another limit of this study is that risk of bias was assessed by one author (CL) and agreed by another (MG). This aspect could have been improved if both authors had worked independently on bias risk assessment and then discussed any discrepancy.

Implications for practitioners

In some clinical contexts, MT could be difficult to apply; for example, some patients may present hyperalgesia to tactile stimuli. Defrin *et al* suggested that tactile allodynia might be present in 60% of patients with chronic LBP associated with radicular pain.⁵⁶

In this kind of patient the use of MT could be excessively painful, and any MT that triggers pain should be avoided.⁵⁷ ST—and therefore a possible placebo effect—could represent a valid alternative to MT in the multidisciplinary approach to back pain, promoting pain relief without increasing the possibility of AE occurrence.

This thought is strengthened by our findings: ST was found to be equally safe to MT without increasing the risk of AE occurrence when compared with no intervention. Furthermore, when blinding was guaranteed, ST showed a statistically significant effect on pain reduction in chronic LBP patients compared with no treatment.

ST could be seen as an ‘affective touch’, which it is suggested creates a pleasant therapeutic experience promoting affiliative behaviours and pain improvement.^{58 59}

Nevertheless, due to the low quality of the studies included in this review, further studies are needed to verify the possible role of ST among patients where MT is not well tolerated.

Implications for research

In MT trials, a true placebo is impossible to achieve so trials should implement strategies to guarantee patient and assessor blinding, for example, avoiding the inclusion of participants who already received the active treatment and avoiding single technique performance which are more difficult to mask. Plans to avoid performance

bias, such as giving similar treatment with similar localisation have to be implemented.

Moreover, the evaluation of the success of blinding should be considered as, at least, secondary outcome.

Researchers should pay particular attention to sample size calculation using the MCID. This difference is fundamental both for research and patients. MCID indicates patients' values and preferences and can help clinicians improve interpretation and promote the understanding of the importance of intervention effects in RCTs.

National Institute for Health and Care Excellence guidelines for LBP suggest the use of MT only as 'a part of a treatment package including exercise, with or without psychological therapy'.⁶⁰ Therefore, the development of future CT should imitate the real multidisciplinary clinical context to assess the external validity of future findings.

Future researches should also evaluate the real effects of ST comparing it both with active treatment and with the no intervention groups. Only with this kind of design could the real placebo effect in MT be defined.

CONCLUSIONS

This review aimed to evaluate ST effect in MT trials. MT showed higher efficacy than ST, but when blinding was ensured the effects of ST and MT were larger. Nevertheless, these findings were not clinically meaningful and the very low quality of the included studies might undermine the reliability of this reviews' results.

The use of ST and its application in MT study is very controversial. Future trials should focus on developing a reliable kind of sham procedure similar to the active treatment, to ensure participants blinding and to guarantee a proper sample size for the detection of reliable, clinically relevant, treatment effects.

Acknowledgements We thank Valentina Rotondi, statistician, for all her support during the statistical analysis and data interpretation. We thank Roberta Maoret, librarian, for helping us in the search strategy implementation.

Contributors CL conceived the idea of this review and designed the study with the contribution of MG who also helped in literature search and in the interpretation of study findings. CL and MG revised studies, performed data extraction and analysis and wrote this review. AA and AM provided clinical and technical support, reviewed the manuscript and helped in publication and with the clinical interpretation of study findings. CL is the guarantor of this paper. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. Patients and public were not involved in this project.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available on reasonable request. Data may be obtained from a third party and are not publicly available. All data relevant to the study are included in the article or uploaded as online supplemental information. Details of the characteristics of the included studies and data extracted are available from the corresponding author at carolina.lavazza@docenti.aimoedu.it. Extra data can

be accessed via the Dryad data repository at <http://datadryad.org/> with the doi: 10.5061/dryad.v9s4mw6tb.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID ID

Carolina Lavazza <http://orcid.org/0000-0002-9907-6819>

REFERENCES

- Meissner K, Fässler M, Rücker G, *et al.* Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. *JAMA Intern Med* 2013;173:1941–51.
- Sciences CfoOm. *International ethical guidelines for health-related research involving humans*. 4th edn. Geneva, 2016.
- Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing directive 2001/20/EC 2014.
- Korthals-de Bos IBC, Hoving JL, van Tulder MW, *et al.* Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. *BMJ* 2003;326:1–6.
- Kawchuk GN, Haugen R, Fritz J. A true blind for subjects who receive spinal manipulation therapy. *Arch Phys Med Rehabil* 2009;90:366–8.
- Mancini F, Beaumont A-L, Hu L, *et al.* Touch inhibits subcortical and cortical nociceptive responses. *Pain* 2015;156:1936–44.
- So PS, Jiang Y, Qin Y. Touch therapies for pain relief in adults. *Cochrane Database Syst Rev* 2008;CD006535.
- Cai L, He L. Placebo effects and the molecular biological components involved. *Gen Psychiatr* 2019;32:1–10.
- Kaptchuk TJ, Miller FG. Placebo effects in medicine. *N Engl J Med* 2015;373:8–9.
- Finniss DG, Kaptchuk TJ, Miller F, *et al.* Placebo effects: biological, clinical and ethical advances. *Lancet* 2010;375:686.
- Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol* 2008;59:565–90.
- Miller CB, Espie CA, Epstein DR, *et al.* The evidence base of sleep restriction therapy for treating insomnia disorder. *Sleep Med Rev* 2014;18:415–24.
- Colloca L, Miller FG. The nocebo effect and its relevance for clinical practice. *Psychosom Med* 2011;73:598–603.
- Koes BW. How to evaluate manual therapy: value and pitfalls of randomized clinical trials. *Man Ther* 2004;9:183–4.
- Chaibi A, Šaltytė Benth J, Bjørn Russell M. Validation of placebo in a manual therapy randomized controlled trial. *Sci Rep* 2015;5:1–8.
- Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2004;2.
- Kaptchuk TJ, Stason WB, Davis RB, *et al.* Sham device V inert pill: randomised controlled trial of two placebo treatments. *BMJ* 2006;332:391–7.
- Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
- Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions* version 5.1.0 (updated March 2011), 2011. Available: https://handbook-5-1.cochrane.org/index.htm#chapter_8/8_assessing_risk_of_bias_in_included_studies.htm
- Pool JUM, Ostelo RWJG, Hoving JL, *et al.* Minimal clinically important change of the neck disability index and the numerical rating scale for patients with neck pain. *Spine* 2007;32:3047–51.

- 21 Ostelo RWJG, Deyo RA, Stratford P, *et al.* Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 2008;33:90–4.
- 22 Farrar JT, Young JP, LaMoreaux L, *et al.* Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94:149–58.
- 23 Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011).9.5.2 Identifying and measuring heterogeneity, 2011. Available: https://handbook-5-1.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm
- 24 Licciardone JC, Minotti DE, Gatchel RJ, *et al.* Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. *Ann Fam Med* 2013;11:122–9.
- 25 Bialosky JE, George SZ, Horn ME, *et al.* Spinal manipulative therapy-specific changes in pain sensitivity in individuals with low back pain (NCT01168999). *J Pain* 2014;15:136–48.
- 26 Eardley S, Brien S, Little P, *et al.* Professional kinesiology practice for chronic low back pain: single-blind, randomised controlled pilot study. *Forsch Komplementmed* 2013;20:180–8.
- 27 Erdogmus CB, Resch K-L, Sabitzer R, *et al.* Physiotherapy-based rehabilitation following disc herniation operation: results of a randomized clinical trial. *Spine* 2007;32:2041–9.
- 28 Hansen FR, Bendix T, Skov P, *et al.* Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain. A randomized, observer-blind trial. *Spine* 1993;18:98–108.
- 29 Hoiriis KT, Pfleger B, McDuffie FC, *et al.* A randomized clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain. *J Manipulative Physiol Ther* 2004;27:388–98.
- 30 Krekorkias G, Gelalis ID, Xenakis T, *et al.* Spinal mobilization vs conventional physiotherapy in the management of chronic low back pain due to spinal disk degeneration: a randomized controlled trial. *J Man Manip Ther* 2017;25:66–73.
- 31 Licciardone JC, Stoll ST, Fulda KG, *et al.* Osteopathic manipulative treatment for chronic low back pain: a randomized controlled trial. *Spine* 2003;28:1355–62.
- 32 Senna MK, Machaly SA. Does maintained spinal manipulation therapy for chronic nonspecific low back pain result in better long-term outcome? *Spine* 2011;36:1427–37.
- 33 Lascurain-Aguirrebeña I, Newham DJ, Casado-Zumeta X, *et al.* Immediate effects of cervical mobilisations on global perceived effect, movement associated pain and neck kinematics in patients with non-specific neck pain. A double blind placebo randomised controlled trial. *Musculoskelet Sci Pract* 2018;38:83–90.
- 34 Selkow NM, Grindstaff TL, Cross KM, *et al.* Short-term effect of muscle energy technique on pain in individuals with non-specific lumbopelvic pain: a pilot study. *J Man Manip Ther* 2009;17:14E–18.
- 35 Antolinos-Campillo PJ, Oliva-Pascual-Vaca A, Rodríguez-Blanco C, *et al.* Short-term changes in median nerve neural tension after a suboccipital muscle inhibition technique in subjects with cervical whiplash: a randomised controlled trial. *Physiotherapy* 2014;100:249–55.
- 36 Haller H, Lauche R, Cramer H, *et al.* Craniosacral therapy for the treatment of chronic neck pain: a randomized sham-controlled trial. *Clin J Pain* 2016;32:441–9.
- 37 Klein R, Bareis A, Schneider A, *et al.* Strain-counterstrain to treat restrictions of the mobility of the cervical spine in patients with neck pain: a sham-controlled randomized trial. *Complement Ther Med* 2013;21:1–7.
- 38 Younes M, Nowakowski K, Didier-Laurent B, *et al.* Effect of spinal manipulative treatment on cardiovascular autonomic control in patients with acute low back pain. *Chiropr Man Therap* 2017;25:33.
- 39 Quinn F, Hughes CM, Baxter GD. Reflexology in the management of low back pain: a pilot randomised controlled trial. *Complement Ther Med* 2008;16:3–8.
- 40 Sillevs R, Cleland J, Hellman M, *et al.* Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. *J Man Manip Ther* 2010;18:181–90.
- 41 Cleland JA, Childs JD, McRae M, *et al.* Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. *Man Ther* 2005;10:127–35.
- 42 Pires PF, Packer AC, Dibai-Filho AV, *et al.* Immediate and short-term effects of upper thoracic manipulation on myoelectric activity of sternocleidomastoid muscles in young women with chronic neck pain: a randomized blind clinical trial. *J Manipulative Physiol Ther* 2015;38:555–63.
- 43 Silva ACdeO, Biasotto-Gonzalez DA, Oliveira FHM, *et al.* Effect of osteopathic visceral manipulation on pain, cervical range of motion, and upper trapezius muscle activity in patients with chronic nonspecific neck pain and functional dyspepsia: a randomized, double-blind, placebo-controlled pilot study. *Evid Based Complement Alternat Med* 2018;2018:1–9.
- 44 Vieira-Pellenz F, Oliva-Pascual-Vaca A, Rodríguez-Blanco C, *et al.* Short-term effect of spinal manipulation on pain perception, spinal mobility, and full height recovery in male subjects with degenerative disk disease: a randomized controlled trial. *Arch Phys Med Rehabil* 2014;95:1613–9.
- 45 Kogure A, Kotani K, Katada S, *et al.* A randomized, single-blind, placebo-controlled study on the efficacy of the arthrokinematic approach-hakata method in patients with chronic nonspecific low back pain. *PLoS One* 2015;10:1–16.
- 46 Hall T, Hardt S, Schäfer A, *et al.* Mulligan bent leg raise technique-a preliminary randomized trial of immediate effects after a single intervention. *Man Ther* 2006;11:130–5.
- 47 Dunning J, Mourad F, Barbero M, *et al.* Bilateral and multiple cavitation sounds during upper cervical thrust manipulation. *BMC Musculoskelet Disord* 2013;14:24.
- 48 Johnston BC, Ebrahim S, Carrasco-Labra A, *et al.* Minimally important difference estimates and methods: a protocol. *BMJ Open* 2015;5:e007953–5.
- 49 Charles P, Giraudeau B, Dechartres A, *et al.* Reporting of sample size calculation in randomised controlled trials: review. *BMJ* 2009;338:b1732.
- 50 Moher D, Hopewell S, Schulz KF, *et al.* Consort 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg* 2012;10:28–55.
- 51 Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS One* 2013;8:e59202.
- 52 Ruddock JK, Sallis H, Ness A, *et al.* Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. *J Chiropr Med* 2016;15:165–83.
- 53 Rubinstein SM, de Zoete A, van Middelkoop M, *et al.* Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2019;364:l689.
- 54 Monroe CM. The effects of therapeutic touch on pain. *J Holist Nurs* 2009;27:85–92.
- 55 Rubinstein SM, van Eekelen R, Oosterhuis T, *et al.* The risk of bias and sample size of trials of spinal manipulative therapy for low back and neck pain: analysis and recommendations. *J Manipulative Physiol Ther* 2014;37:523–41.
- 56 Defrin R, Devor M, Brill S. Tactile allodynia in patients with lumbar radicular pain (sciatica). *Pain* 2014;155:2551–9.
- 57 Nijs J, Van Houdenhove B. From acute musculoskeletal pain to chronic widespread pain and fibromyalgia: application of pain neurophysiology in manual therapy practice. *Man Ther* 2009;14:3–12.
- 58 Geri T, Viceconti A, Minacci M, *et al.* Musculoskeletal science and practice. manual therapy: exploiting the role of human touch 2019;44.
- 59 Walker SC, Trotter PD, Swaney WT, *et al.* C-tactile afferents: cutaneous mediators of oxytocin release during affiliative tactile interactions? *Neuropeptides* 2017;64:27–38.
- 60 (NICE) NifHaCE. Low back pain and sciatica in over 16s: assessment and management, 2020. Available: <https://www.nice.org.uk/guidance/ng59/chapter/Recommendations11>