

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<u>http://bmjopen.bmj.com</u>).

If you have any questions on BMJ Open's open peer review process please email <u>info.bmjopen@bmj.com</u>

BMJ Open

# **BMJ Open**

# Placebo effect in manual therapy trials on back pain patients: a systematic review and pair-wise meta-analysis

Journal:	BMJ Open
Manuscript ID	bmjopen-2020-045106
Article Type:	Original research
Date Submitted by the Author:	22-Sep-2020
Complete List of Authors:	Lavazza, Carolina; AIMO, Research Galli, Margherita; AIMO, Research Abenavoli, Alessandra; AIMO, Research Maggiani, Alberto; AIMO, Research
Keywords:	COMPLEMENTARY MEDICINE, STATISTICS & RESEARCH METHODS, Back pain < ORTHOPAEDIC & TRAUMA SURGERY

SCHOLARONE<sup>™</sup> Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our <u>licence</u>.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which <u>Creative Commons</u> licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

reliez oni

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

avazza (0000-0002-9907-6819), a Galli (0000-0003-4050-3131), A (0000-0001-8941-2308), A (0000-0002-7980-9405) nt of Research AIMO institute, antuario 7, 21047 Saronno, Italy, avazza associate professor, nt of Research AIMO institute,
antuario 7, 21047 Saronno, Italy, avazza associate professor, nt of Research AIMO institute,
antuario 7, 21047 Saronno, Italy, a Galli associate professor, nt of Research AIMO institute, antuario 7, 21047 Saronno, Italy, oli associate professor, Department ch AIMO institute, Piazzale 7, 21047 Saronno, Italy, A Maggian
dence to: vazza@docenti.aimoedu.it

#### Copyright/license for publication:

Carolina Lavazza The Corresponding Author of this article contained within the original manuscript which includes any diagrams & photographs within and any related or stand alone film submitted ("the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at:

http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyrightopen-access-and-permission-reuse.

Please tick one or more boxes as appropriate:

×

- I am the sole author of the Contribution.
- I am one author signing on behalf of all co-owners of the Contribution.
- П The Contribution has been made in the course of my employment and I am signing as authorised by my employer.
  - I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown\* acting in the course of my employment
- I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown acting in the course of my employment
- I am not an employee of the UK Crown acting in the course of my employment but some/all of my coauthors are.\*

\*Such authors should consult guidance and if necessary return any completed form.

Licence: The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Conflict of interest: I have read and understood the BMJ Group policy on declaration of interests and declare the following interests: none

Ethical approval: not required.

Transparency: The manuscript's guarantor affirms that this 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.

Funding: None.

**Competing interests**: No funding were provided for this project and authors did not received support from any organisation for the submitted work. Authors have no financial relationships with any organisations that might have an interest in this project and have no other relationships or activities that could appear to have influenced this review.

in ha

1	
2	
3	Acronyms
4	
5	AE= adverse effects
6	
7	BP= back pain
8	CI= confidence intervals
9	CT= clinical trial
10	MA = meta-analysis
11	
12	MCID= minimal clinically important difference
13	MD=mean difference
14	MT= Manual Therapies
15	OMT = osteopathic manipulative treatment
16	PROs = patient-reported outcomes
17	RCT= randomised controlled trials
18	
19	RR= risk ratio
20	SM= Spinal Manipulation
21	
22	
23 24	
24 25	
25 26	
20 27	
27	
28 29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49 50	
50	
51	
52	
53	
54 57	
55 56	
56	
57	
58	
59	

# ABSTRACT

#### **Objective**:

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

#### Methods and analysis:

Different databases were screened up to 20 August 2020. RCT involving adults affected by BP, acute or chronic, were included.

Hand contact placebo was compared to different MT (physiotherapy, chiropractic, osteopathy, massage, kinesiology and reflexology) and to control. Primary outcomes were BP improvement, placebo reliability and adverse effect (AE). Secondary outcomes were number of dropouts. Dichotomous outcomes were analysed using risk ratio (RR), continuous using mean difference (MD), 95% confidence intervals (CI). The minimal clinically important difference was 30 mm changes in pain score.

#### **Results**:

24 trials were included with 2,019 participants. Very low evidence quality suggests clinically insignificant pain improvement in favour of MT compared to placebo (MD 3.86, 95% CI 3.29 to 4.43) and no differences between placebo and control (MD -6.04, 95% CI -16.68 to 4.59). Placebo reliability shows a high percentage of correct detection by participants (ranged from 46.7% to 83.5%), spinal manipulation being the most recognized technique.

Low quality of evidence suggests that AE and dropout rates were similar between placebo and MT (RR AE=0.84, 95% CI 0.55 to 1.28, RR dropouts= 0.98, 95% CI 0.77 to 1.25). A similar dropout rate in control (RR=0.79, 95% 0.51 to 1.23).

#### **Conclusions:**

Comparison of placebo and MT shows a small, clinically meaningless effect in pain improvement. Similar effects were found with control. The heterogeneousness of placebo MT studies and the very low quality of evidence render uncertain these review findings. Future trials should develop reliable kinds of placebo, similar to active treatment, to ensure participant-blinding and to guarantee proper sample size for the reliable detection of clinically meaningful treatment effects.

#### PROSPERO register: CRD42020198301

https://www.crd.york.ac.uk/prospero/display\_record.php?RecordID=198301

#### 

# Strengths and limitations of this study

#### Strengths

This systematic review and pair-wise meta-analysis:

- summarises existing evidence on the effectiveness, reliability and application of hand contact placebo in MT randomised controlled trials;
- gives suggestions for researchers on conducting methodical RCT in MT using a reliable placebo.

#### Limitation

- This study did not include a comparison with machine provided placebo, its aim focused on hand contact placebo
- Insufficient number of studies were included to conduct a network meta-analysis

#### Summary of findings:

1. Placebo compared to Manual Therapies (MT)

Outcomes		ed absolute (95% CI)	Relative effect	Nº of participants	Certainty of the evidence	Comments	
Outcomes	Risk with MT	Risk with Placebo	(95% CI)	(studies)	(GRADE)	Comments	
Pain improvement assessed with: VAS score Scale from: 0 to 100		MD <b>3.86</b> higher (3.29 higher to 4.43 higher)	-	805 (15 RCTs)	⊕⊖⊖⊖ VERY LOW a,b	A small effect, not clinically relevant, in pain improvement wa detected in favour of MT. This analysis excluded two trials (one suspected of publication bias, one used a different scale) which increased heterogeneity levels bu did not affect overall efficacy meaningfully.	
Adverse events assessed with: number of AE occurred	144 per 1.000	<b>121 per</b> <b>1.000</b> (79 to 184)	<b>RR 0.84</b> (0.55 to 1.28)	531 (6 RCTs)	⊕⊕⊖⊖ Low ª	Pooled data from 6 studies did nor show any difference in AE occurrence between placebo and MT.	
Dropouts rate assessed with: number of participants that leaved the study	174 per 1.000	<b>171 per</b> <b>1.000</b> (134 to 218)	<b>RR 0.98</b> (0.77 to 1.25)	1238 (11 RCTs)		Pooled data from 11 trials did not show difference in dropout rate between placebo and MT.	
2. Placebo comp		ntrol			0		
Patient or population: Intervention: Placebo Comparison: control	back pain						
Outcomes		ed absolute (95% CI)	Relative effect	Nº of participants	Certainty of the evidence	Comments	
	Risk with control	Risk with placebo	(95% CI)	(studies)	(GRADE)		
		MD <b>6.04</b>			$\oplus \bigcirc \bigcirc \bigcirc$	Pooled data from four trials, highl inconsistent, showed no	
Pain improvement assessed with: VAS score Scale from: 0 to 100		lower (16.68 lower to 4.59 higher)	-	251 (4 RCTs)	VERY LOW a,c,d	differences between placebo and control group in pain improvement.	

**Cl:** Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

- 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

#### Explanations

- a. The majority of trials were judged as poor quality according to AHRQ standards.
- b. Most of the studies were small trial.
- c. Heterogeneity levels at 80%.
- d. Number of participants < 400

to peer teries only

### Background

In Clinical Trials (CT), placebo is commonly used as control therapy to evaluate 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)

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 (IMP) and as such it has to follow different standards in order to ensure quality, guarantee patient safety and the reliability of study results. (3)

Regulatory aspects of trials involving Manual Therapies (MT) are very different. Although these 18 types of studies might be influenced by the type of placebo provided, no clear guidelines or 20 regulations have been developed to ensure credibility of trial results and patient safety. In MT trials, placebo treatment is often provided in different modalities from trial to trial although 22 manual techniques or treatments tested are the same. For instance, placebo treatment is commonly administrated as a light touch in the site of pain or as an active treatment in a different 24 site (4), with no clear criterion. Furthermore, in these studies, placebo reliability has been rarely 26 evaluated. An analysis of its credibility results could help to understand better the participants' point of view, assessing which kind of placebo seems more similar to the active treatment 28 provided. 29 30

Placebo effect, also called placebo response, is the reported improvement in symptoms among patients in randomized controlled trials (RCT). Since a placebo has no inherent therapeutic power, it rarely cures the disease but it may contribute to the relief of patients' symptoms such as pain.(5) Additionally, placebo might be related to an adverse effect called nocebo. It has been estimated that up to 26% of patients in RCTs discontinue placebo due to adverse effects. (6)

37 It is thought that these psychobiological phenomena may be related to the overall therapeutic 38 context, such as treatment environment, individual patient and clinician factors (e.g. beliefs, desire 39 for symptom changes), as well as the patient's expectations of improvement and prior experiences 40 of the treatment. (7-10) 41

42 In pharmacological trials this overall therapeutic context and its influence on placebo response has 43 been widely studied. (8) Less evidence is present for MT trials, where other important 44 characteristics should be considered as part of this therapeutic context such as the tactile 45 interaction between patient and practitioner and clinician beliefs. As a matter of fact, touch might 46 47 have a positive health effect (11) and placebo might be influenced by the same therapist beliefs 48 which are actively providing the inactive treatment. (12, 13)

49 Another important factor that has to be taken into account is that RCTs involving MT usually use 50 patient-reported outcomes (PROs) - such as pain - as primary outcomes. Studies suggested that 51 52 physical placebo treatments might have a greater effect on these types of outcome compared to 53 pharmacological placebo and that this effect might be consequent to the physical contact. (1, 14, 54 15) 55

Therefore, a better understanding of placebo response in manual treatment would be fundamental to define the real difference in efficacy between active and inactive treatment, with a better knowledge of the effect of manual contact on PROs such as pain relief and dropouts.

59 60

9

56

57 58

1 2 3

4 5 6

7 8

9

10

11

12

13 14

15

16

17

19

21

23

25

27

31

32

33

34

The role of placebo in MT trials is still very confused and the lack of guidelines allows huge discrepancies in its use in RCTs. 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 pair-wise meta-analyses (MA) is to evaluate the use of placebo in MT trials in order to analyse the effects, possible harm and the reliability of different kinds of sham treatments provided in RCTs involving MT. A systematic review could help to define placebo 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 placebo in manual therapy (MT) randomized controlled trials in the treatment of back pain (BP) in order to provide methodological guidance for clinical trial development.

# Methods

This systematic review and meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews (PRISMA)(16).

The protocol registration was performed in PROSPERO (<u>http://www.crd.york.ac.uk/PROSPERO/</u>) and review registration number is *CRD42020198301*.

#### Criteria for considering studies for this review

Only randomised controlled studies (RCTs) were included in this review. Quasi-randomised trials in which allocation was not strictly random (e.g. 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 is related to muscular conditions, articular disorders (such as osteoarthritis) or spinal disc herniation were included. Trials where musculoskeletal diseases were secondary to other pathologies (e.g. 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 placebo that include hand contact, studies where placebo was
   provided by machines (such as inactive ultrasound) were excluded.
- All trials that involved hand contact placebo as light touch or a manual treatment in a different site were included.
- Placebo was compared to other manual therapies such as: physiotherapy, chiropractic,
   osteopathy, massage, kinesiology and reflexology and to control.
- Physiotherapeutic exercises were included in the analysis only if associated with manual treatment.

The use of active co-interventions such as oral 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 of interested arms were extracted.

#### Outcomes

Primary outcomes were pain intensity on a validated scale, reliability of placebo and adverse effect. Secondary outcomes were number of dropouts.

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

#### Information sources

Search strategy (*Appendix 1*) was adapted to the different databased by an experienced information specialist.

RCTs were identified in different databases (up to 20 August 2020): MEDLINE, Embase, CINAHL, SPORTDiscus, PEDro, World Health Organization Clinical Trials Registration Platform, Index to Chiropractic Literature, Cochrane central register of controlled trials (CENTRAL), Clinical trials registry and metaRegister of Controlled Trials (mRCT).

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 on-going or recently completed systematic reviews. Guidelines from different organizations (e.g. National Council for Osteopathic Research etc.) were reviewed and references from relevant publication were analysed.

#### Data collection and analysis

Searches 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 firstly 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 the document delivery service of the 3Bi Biella library.

Uncertainty about the inclusion of a study were 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 pre-defined form. Data extracted were related to settings, type of study, participants characteristics (such as localization 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 placebo and active treatment or funding. (*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 (27). 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 judgments was performed by *MG*. RevMan 5.3.5 was used for the graphic representation of each risk. The CRB tool results were then converted to 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 criteria was high risk, poor quality studies instead 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 adverse events (occurred or not), were analysed using risk ratio (RR) with 95% confidence intervals (CI).

Continuous outcomes, such as back pain on VAS scale, were evaluated using mean difference (MD) between placebo and the MT/control group with 95% CI and the standard deviation (SD). The minimal clinically important difference (MCID) between pre- and post-treatment was taken as 30 mm changes in 100 mm pain score. (17-19) These values were used for the interpretation of the clinical significance of the findings.

Placebo reliability 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<sup>2</sup>).

Cochrane Handbook was used for threshold interpretation: heterogeneity was considered as unimportant for value of I<sup>2</sup> between 0% and 40%,, as moderate for values between 30% and 60%, as substantial for values between 50% an 90% and considerable for values between 75% to 100%. (20)

#### Synthesis of results

Meta-analysis of pain score, AE and dropout rates were performed using RevMan 5.3.5 whenever possible. The meta-analyses compared all kinds of placebo with all types of manual therapies and to control. Random-effect model was used when a substantial inconsistency was present (I<sup>2</sup>= 50-90%). (20) When considerable heterogeneity was present (I<sup>2</sup>>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 placebo type provided (applied locally or in different sites from pain), type of manual technique tested (single or multiple techniques) and localization 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 appendices.

#### Summarizing results and assessing the quality of the evidence

The quality of evidence for each outcome was evaluated with the 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 (21). The GRADE approach was used to assess the quality of the key outcomes. The software GRADEpro (https://gradepro.org) was used to import data from RevMan 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 dropouts.

#### Patient and Public Involvement

There was no involvement of patients or public during the outline of this project.

to occurrence of the terms of terms of

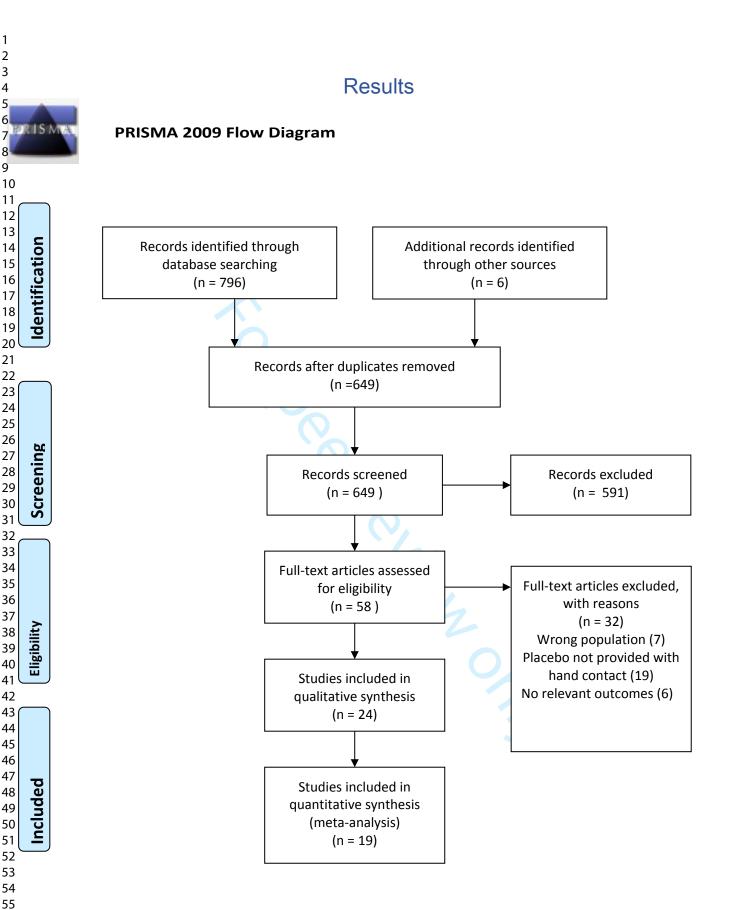


Figure 1: PRISMA flow diagram

### **Included studies**

All 24 studies included in this review were RCT. One study had a 2x2 factorial design, (22) eight studies had multiple arms. (23-30) Only arms involving the interested treatments were included in analyses.

Most of the studies were conducted in physical therapy clinics, in 13 different countries. Three trials did not report where their were conducted. (27, 31, 32)

Eight trials were conducted in Europe , (25, 26, 28, 33-37) five in the United States, (22, 23, 29, 38, 39) three studies in Brazil, (40-42) one in UK, (24) Egypt, (30) Japan (43) and Australia. (44) No ongoing or unpublished trials were found.

#### Population

The included trials randomized a total of 2,019 participants, the majority of studies 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/m<sup>2</sup>.

The majority of studies included both genders, with a percentage of male that ranged from 19% to 80% and a percentage of women that ranged from 20% to 82%. Two trials included only male, (36, 42) one study included only female participants. (40)

16 trials enrolled participants with low back pain (LBP), nine included participants with cervical pain (CP).

The majority of trials (N=15) included participants with unspecified cause of back pain. Disk herniation was considered in three trials, (25, 28, 42) other three studies included participants with mechanical pain (described as pain exacerbated by movement). (23, 24, 39).

Duration of symptoms were unassessed in eight trials, most of the studies included participants with chronic pain (N=9), some included participants with both acute and chronic pain.

Participants with experience of the tested treatment were included in 8 trials (22, 27, 29, 30, 33, 35, 40, 41) and excluded in four. (24, 34, 37, 39) Remaining studies did not provide this information.

#### Interventions

Interventions deferred for number of sessions and number of techniques applied. Most of the trials used a single therapy session (N=11) with a single technique performed (N=8). Trials with different therapy sessions ranged from 5 (23, 24, 28) to 20 (25) sessions once a week.

#### Placebo

Placebo was provided with a hand contact on the area of pain in 19 studies, five studies provided placebo in a different area from where the pain was located. (25, 33, 41, 43, 44)
In trials providing spinal manipulation, as inactive treatment the majority of authors used the similar placement of hands on participants without any force applied. (38-40, 42)
Two trials used a placebo with similar forces applied in different directions. (23, 30)
one trial did not specify the inactive manipulation applied(27).
In trials that provided multiple techniques in the same treatment session (such as osteopathic treatment, spinal mobilization and physiotherapy) the placebo was administrated with different techniques that mimed active treatments using light touch or light tractions.

Only one trial compared one single placebo technique with both single active technique and multiple treatment techniques. In this case only data of the first arm were extracted. (35)

#### Active and controls treatments

Different active treatments were provided:

- Physiotherapy (2 trials, 288 participants)
- Spinal manipulation (SM)/chiropractic (7 studies, 567 participants)
- Osteopathy (5 trials, 645 participants)
- Kinesiology (one trial, 58 participants)
- Articular mobilizations (5 trials, 325 participants)
- Muscular release (4 trials, 136 participants)

Five trials with multiple arms compared placebo to control group (343 participants).

The active treatment was generally applied in the area of pain, some trials used techniques additionally in other areas. Just one trial using reflexology provided both active and inactive manipulation in a different zone. (37)

Characteristics of practitioner who administrated treatments were provided by 16 trials. Most of the trials involved physiotherapists (N=8), generally defined physical therapists (N=4), osteopaths (N=3) and students (N=1). Only seven studies provided information on years of practice experience of physicians involved that ranged from 6 to 17 years. (28, 31, 33-35, 38, 40, 42). Information of their gender was provided only in three trials. (24, 28, 35)

#### 

Study ID	N° of participants	Symptoms duration	Pain localization	Technique tested (site of application)	Type of placebo	Other arms	Follow-up
Antonilos- Campillo PJ 2014	40	Not reported	Cervical	Soft-tissue (cervical region)	Soft mobilization of lower limbs	None	No follow-up (outcomes collected after the interventio
Bialosky J 2014	110	> 4 months	Lumbar	Spinal Manipulation (SM) <i>(lumbar</i> <i>spine)</i>	Ineffective force applied on lumbar spine	Control group	2 weeks
Cleland JA 2005	36	>2 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the interventio
Eardley S 2013	58	> 3 years	Lumbar	Kinesiology <i>(spine)</i>	Protocol of ineffective techniques in the site of pain	Control group	7 weeks
Erdogmus S 2007	120	$\geq$ 1.4 weeks	Lumbar	Physiotherapy (spine)	Neck massage	Control group	1.5 years
Hall T 2004	24	Not reported	Lumbar	BLR technique (lower limbs)	Soft-tissue manipulation of the foot	None	24 hours
Haller H 2016	54	> 7 months	Cervical	Cranio-sacral therapy (head)	Ineffective touch of head	None	3 months
Hansen F 1993	168	$\geq$ 18 days	Lumbar	Physiotherapy (lumbar spine and abdomen)	Intermittent traction of the spine	Intensive back muscle training	1 year
Hidalgo B 2015	32	Not reported	Lumbar	Articular mobilization (lumbar spine)	Ineffective mobilization forces applied on lumbar spine	None	2 weeks
Hoiriis K 2004	156	$\geq$ 2,3 weeks	Lumbar	SM (spine)	Ineffective force applied on spine	Medical treatment	4 weeks
Klein R 2003	61	>1 month and <5 years	Cervical	Strain- counterstain techniques (cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the interventic
Kogure A 2015	179	> 12 months and < 10 years	Lumbar	АКА-Н (sacro- iliac joint)	Ineffective force applied on sacro-iliac joint	None	6 months
Krekoukias G 2017	50	Not reported	Lumbar	Articular mobilization techniques (lumbar spine)	Hand contact with lumbar skin placement without any movement	Exercise plus TENS	5 weeks
Lascurain- Aguirrebena I 2018	40	Not reported	Cervical	Articular mobilization (Cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the interventic
Licciardone J 2003	91	$\geq$ 3 months	Lumbar	Osteopathic manual treatment (OMT) <i>(all body)</i>	Protocol of light touch techniques similar to OMT	Control group	6 months

					applied to all body		
Licciardone J 2013	455	$\geq$ 3 months	Lumbar	Osteopathic manual treatment (OMT) – (all body)	Protocol of light touch techniques similar to OMT applied to all body	None	8 weeks
Pires FP 2015	32	> 3 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	72 hours
Quinn F 2008	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 M 2009	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 seconds without any active contraction	None	24 hours
Senna MK 2011	93	$\geq$ 13 months	Lumbar	SM (lumbar spine)	Ineffective force applied on lumbar spine	Maintained SM	10 months
Sillevis R 2010	100	$\geq$ 23 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the interventio
Silva A 2019	28	>3 months	Cervical	Osteopathic visceral treatment (abdomen)	Hand contact on umbilical region without any movement	None	7 days
Veira-Pellez F 2014	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 M 2017	17	< 3 months	Lumbar	OMT (all spine)	Placebo mimed active treatment with an ineffective force applied.	None	7 days

#### Risk of bias in included studies

- Figure 2 shows risks of bias judged by two authors.
  Blinding of participants and assessors will be described due to the nature of this review.
  According to AHRQ standards of CRB tool, (45) the majority of trials were judged with poor quality (N=22). Good quality was conferred on only two studies. (34, 43)
  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 judgment were mainly related to trials involving spinal manipulations.
  These studies used a technique which can be easily recognized 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. (26, 30)
- Incomplete outcome data was the least common bias risk with 80% of trials judged as low risk.
   Reporting bias was evaluated unclear in 55% of trials where registration number and trial
   protocol were not reported or found.
  - Other bias occurred was generally considered at high risk for baseline differences of the population in 30% of trials.

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

# Effects of intervention

#### Placebo versus other manual therapies

#### Pain

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

The comparison between placebo and MT was performed in 17 studies. One trial used a different scale and data were obtained with a conversion formula.(25) Data from seven studies could not be extracted.

The meta-analysis at short-term showed substantial heterogeneity levels using a randomeffects 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 (28). This thought was verified with a funnel plot, which showed an asymmetric distribution with the inclusion of these two studies (*Appendix 3*). This sensitivity analysis did not influence overall effectiveness results but inconsistency levels decreased considerably at short-term. It can be deducted that a possible cause of heterogeneity was found (*Full analysis in appendix 4*)

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

#### Figure 3: Forest plot of comparison Placebo versus MT in back pain outcome at short-term.

Comparisons between placebo 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) (*Appendix 4.1*)

 Reliability of placebo was evaluated in five trials; one did not report the results. (28) 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 placebo with SM, these trials showed a correct perception of treatment allocation that ranged from 63.5% (23) to 83.5%. (27) In this last study patients were considered eligible if they already received SM.

One trial compared placebo to an articular mobilization technique. 54.5% participants correctly
 guessed treatment allocation. (44)

Participants of one study that compared placebo to reflexology had the lowest perception of the correct detection of allocation (46.7%). Participants in this trials were naïve to the type of treatment tested. (37)

#### Dropouts

Pooled data from 11 trials at the last follow-up suggested no difference in dropouts rate between placebo and MT at the end of the trials(105/612 compared to 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)* 

Figure 4: Forest plot of comparison Placebo versus MT in number of dropouts outcome

#### Adverse effects

Adverse effects were generally under-reported, six trials were included in the metaanalysis.(24-26, 34, 35, 43)

Two trials reported AE overall occurrence without specified event rates in the groups. (22, 30). AE were predominantly minor and lasted for two/three days after treatment, in the majority of trials transient worse pain, tiredness, muscle weakness and transient headache were reported.(24, 34, 35, 43)

Senna M 2011 reported the most common AE were local discomfort and tiredness but no serious complications were noted. (30)

Haller H 2016 reported two patients dropping out from the trial for recurrent headache after treatments, both Haller H and Klein R 2013 reported dizziness of one patient.

Licciardone J 2013 reported 27% of patients with AE, 2% had serious AE not related to study interventions. (22)

Overall results showed no clear difference in AE occurrence between placebo and MT (32/267 compared to 38/264; RR 0.84, 95% CI 0.55 to 1.28; 531 participants, I<sup>2</sup>=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.

# *Figure 5:* Forest plot of comparison Placebo versus MT in number of adverse events outcome at short-term

#### Placebo versus control

#### Pain

Five studies compared placebo to control, four were included in random-effect meta-analysis at short-term. (23-25, 27) Data from one trial could not be extracted. (29) Pooled data showed the presence of significant heterogeneity, therefore results are reported narratively: three trials showed no difference between placebo and control on pain outcome, while Eardley S. 2013 showed an effect in favour of placebo (pooled data from 4 trials: MD - 6.04, 95%CI from -16.68 to 4.59, 252 participants, I<sup>2</sup>= 80%, P=0.27). The exclusion of Erdogmus S 2013 (that used a different scale) did not affect the results of effectiveness but decreased levels of heterogeneity (MD -9.72, 95%CI -19.94 to 0.51, I<sup>2</sup>= 69%, P=0.12) (*Appendix 5*)

#### Dropouts

No differences were showed in the fixed-effect meta-analysis on dropout rate between placebo and control in five trials (27/165 compared to 34/166; RR 0.79, 95% CI 0.51 to 1.23; 331 participants,  $I^2=0\%$ , P=0.30; very low quality of evidence downgraded two levels for very serious risk of bias and imprecision). *(Figure 6)* 

Figure 6: Forest plot of comparison Placebo versus MT in number of dropouts outcome

#### Adverse effects

Of the six studies reporting AE, only two compared placebo and control. One, Eardley S 2013, did not evaluate the AE occurred in control group while Erdogmus C 2007 reported that 10/40 in the control group and 11/40 in placebo group turned to other therapies for complains.

4 5 6

7 8

9

10

11

12

13 14

15

16

21

22

23

24 25

26

31

32

33

35 36

37

38

39 40

41 42

43

44

45

46 47

48

49

50

51 52 53

54

55

56

## Discussion

In the treatment of back pain very low quality of evidence suggests a slight improvement of pain, not clinically meaningful, in favour of MT at short-term. Substantial levels of heterogeneity within the four studies analysed, showed no differences between placebo and control in pain reduction.

Reliability of placebo was reported in four trials that compared placebo to MT, with high percentage of correct detection of treatment allocations by participants.

AE were generally under-reported, with a similar rate of occurrence between placebo and MT accompanying low levels of heterogeneity. Only one study reported AE in control group with no significant difference from placebo.

SM techniques were the treatment most evaluated (N=7). These techniques are highly recognizable by patients for a popping sound emitted by the column during their performance. (46) 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 recognized treatment allocation in this kind of trial (from 63.5% to 83.5%)(23, 27).

27 Additionally, five trials applied placebo treatment in a different site compared to pain and 28 active treatment. This might have had important influences on study results. 29 30

Reliability placebo seemed not to be related to dropouts rate, although both these data were reported only in two trials. Bialosky J and Hoiriis K showed high percentages of correct treatment allocation detection by participants but dropout rate between placebo and MT group did not differ. These results seem to be in conflict, nevertheless, participants could have 34 wanted to remain in the trial for several other reasons such as settings or being evaluated by an expert clinician free. This possibility is reinforced by the fact that a similar dropout rate was reached in the comparison placebo versus control. These data suggest that dropout rate might not be a dependable outcome for assessing reliability of placebo.

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. (47)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. (48-50)

Our results are similar to other reviews findings, notwithstanding that these reviews did not consider the difference between kinds of placebo provided (hand contact or machines) and evaluated the effect of a singular type of MT (such as SM or OMT) compared to placebo. (51, 52)

#### Limitations

This review aimed to compare different kinds of placebo with different kinds of MT and control. 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 placebo. 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.(11) Additionally, the use of machine placebo trials in the same meta-analysis could have increased diversity within trials included 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 metaanalysis 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 placebo 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 metaanalysis. 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

Pair-wise comparison on pain outcome between placebo 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. (53) However trials with bigger sample size are needed to assess a real correlation between these two factors.

#### Implications for practice and research

There is very low quality of evidence that placebo compared to MT might be less effective and equally safe in the treatment of patients affected by back pain. Future studies should improve their methodological properties to ensure patients safety and to guarantee reliability of study results.

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.

Trials should also implement strategies to guarantee patients and assessors blinding, for
 example avoiding the inclusion of participants who already received the active treatment. Plans
 to avoid performance bias, such as giving similar treatment with similar localization have to be
 implemented. Moreover, the evaluation of the reliability of blinding should be considered as, at
 least, secondary outcome.

Future researches should also evaluate the real effects of placebo comparing it both with active treatment and to control groups. Only with this kind of design the real placebo effect in MT could be defined.

Trials should also focus on including participants with similar characteristics such as duration of symptoms (acute or chronic pain).

The majority of studies included in this review used a single technique treatment (N=11), however the clinical relevance of demonstrating the effectiveness of a singular technique is not clear. In a clinical context, most manual treatments usually involve different kinds of techniques in the same treatment session, so trials that evaluate the effectiveness of a type of treatment should include a routine of techniques in order to be more similar to clinical approaches. Studies should also consider using objectives end points, not patient-reported or observerreported, with a longer period of follow-up.

# Conclusions

This review aimed to evaluate placebo effect in MT trials. Although MT showed higher efficacy than placebo, 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 placebo and its application in MT study is very controversial. Future trials should focus on developing a reliable kind of placebo, 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.

#### **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.

**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.

**Data sharing:** Details of the characteristics of the included studies and data extracted are available from the corresponding author at <u>carolina.lavazza@docenti.aimoedu.it</u>.

References

2	
3	
4	
5	
6	
6 7 8	
, 0	
0	
9	
10	
11	
12	
12	
15	
13 14 15	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35 36 37	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
22	

1

#### 1. Meissner K, Fässler MRG, Kleijnen J, Hróbjartsson A, Schneider A, Linde K. Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. . JAMA internal medicine. 2013;173(1941-1951). Sciences CflOoM. International Ethical Guidelines for Health-Related Research Involving 2. Humans, Fourth Edition. Geneva2016. REGULATION (EU) No 536/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL 3. of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC, (2014). 4. Kawchuk GN, Haugen R, Fritz J. A true blind for subjects who receive spinal manipulation therapy. Archives of physical medicine and rehabilitation. 2009;90(2):366-8. Cai L, He L. Placebo effects and the molecular biological components involved. General 5. psychiatry. 2019;32(5):1-10. Kaptchuk TJ, Miller FG Placebo effects in medicine. N Engl J Med. 2015;373(1):8-9. 6. Finniss DG, Kaptchuk TJ, Miller F., Benedetti F. Placebo effects: biological, clinical and 7. ethical advances. . Lancet. 2010;375:686. 8. 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 FG, Kaptchuk TJ. The power of context: reconceptualizing the placebo effect. 9. Journal of the Royal Society of Medicine. 2008;101:222-5. 10. Colloca L, Miller FG. The nocebo effect and its relevance for clinical practice. Psychosomatic medicine. 2011;73:598-603. Monroe CM. The effects of therapeutic touch on pain. Journal of Holistic Nursing. 11. 2009;27(2):85-92. Chaibi A, Benth JŠ, Russell MB. Validation of placebo in a manual therapy randomized 12. controlled trial. Scientific reports. 2015;5(1):1-8. 13. Koes BW. How to evaluate manual therapy: value and pitfalls of randomized clinical trials. Manual therapy. 2004:183-4. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. Cochrane 14. database of systematic reviews. 2004;2. Kaptchuk TJ, Stason WB, Davis RB, Legedza AR, Schnyer RN, Kerr CE et al. Sham device v 15. inert pill: randomised controlled trial of two placebo treatments. BMJ. 2006;332:391-7. 16. PRISMA. Transparent reporting of systematic reviews and meta-analyses http://prismastatement.org/prismastatement/flowdiagram.aspx2009 Pool JJ, Ostelo RW, Hoving JL, Bouter LM, De Vet HC. Minimal clinically important change 17. of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. . Spine. 2007;32(26):3047-51. 18. Ostelo RW, Stratford P, Waddell G, Croft P, Von Korff M. Henrica C. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. . Spine. 2008;33(1):90-4. Farrar JT, Young Jr JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes 19. in chronic pain intensity measured on an 11-point numerical pain rating scale. . Pain. 2001;94(2):149-58.

2	
3	20. Library C. Identifying and measuring heterogeneity
4	http://handbook.cochrane.org/chapter 9/9 5 2 identifying and measuring heterogeneity.ht
5	m2011 [
6	21. Training C. GRADE approach to evaluating the quality of evidence: a pathway
7 8	
9	http://training.cochrane.org/path/grade-approach-evaluating-quality-evidence-pathway2013 [
10	22. Licciardone JC, Minotti DE, Gatchel RJ, Kearns CM, Singh KP. Osteopathic manual
11	treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. The
12	Annals of Family Medicine. 2013;11(2):122-9.
13	23. Bialosky JE, George SZ, Horn ME, Price DD, Staud R., Robinson ME. Spinal manipulative
14	therapy-specific changes in pain sensitivity in individuals with low back pain (NCT01168999).
15 16	Journal of Pain. 2014;15(2):136-48.
10	24. Eardley S, Brien S, Little P, Prescott P, Lewith G. Professional kinesiology practice for
18	chronic low back pain: single-blind, randomised controlled pilot study Complementary
19	Medicine Research. 2013;20(3):180-8.
20	25. Erdogmus CB, Resch KL, Sabitzer R, Müller H, Nuhr M, Schöggl A et al. Physiotherapy-
21	based rehabilitation following disc herniation operation: results of a randomized clinical trial.
22	Spine. 2007;32(19):2041-9.
23 24	26. Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L et al. Intensive, dynamic
24	
26	back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back
27	pain. A randomized, observer-blind trial. Spine. 1993;18(1):98-108.
28	27. Hoiriis KT, Pfleger B, McDuffie FC, Cotsonis G, Elsangak O, Hinson R et al. A randomized
29	clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain.
30	Journal of manipulative and physiological therapeutics. 2004;27(6):388-98.
31 32	28. Krekoukias G, Gelalis ID, Xenakis T, Gioftsos G, Dimitriadis Z, Sakellari V. Spinal
33	mobilization vs conventional physiotherapy in the management of chronic low back pain due to
34	spinal disk degeneration: a randomized controlled trial. Journal of Manual and Manipulative
35	Therapy. 2017;25(2):66-73.
36	29. Licciardone JC, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W et al. Osteopathic
37	manipulative treatment for chronic low back pain: A randomized controlled trial. Spine.
38 39	2003;28(13):1355-62.
39 40	30. Senna MK, Machaly SA. Does maintained spinal manipulation therapy for chronic
41	nonspecific low back pain result in better long-term outcome? Spine. 2011;36(18).
42	31. Lascurain-Aguirrebeña I, Newham DJ, Casado-Zumeta X, Lertxundi A, Critchley DJ.
43	Immediate effects of cervical mobilisations on global perceived effect, movement associated
44	
45	pain and neck kinematics in patients with non-specific neck pain. A double blind placebo
46 47	randomised controlled trial. Musculoskeletal Science and Practice. 2018;38:83-90.
47	32. Selkow NM, Grindstaff TL, Cross KM, Pugh K, Hertel J, Saliba S. Short-term effect of
49	muscle energy technique on pain in individuals with non-specific lumbopelvic pain: a pilot
50	study Journal of manual & manipulative therapy. 2009;19(1):14E-8E.
51	33. Antolinos-Campillo PJ, Oliva-Pascual-Vaca A, Rodríguez-Blanco C, Heredia-Rizo AM, Espí-
52	López G. V., Ricard F. Short-term changes in median nerve neural tension after a suboccipital
53	muscle inhibition technique in subjects with cervical whiplash: a randomised controlled trial.
54 55	Physiotherapy. 2014;100(3):249-55.
56	34. Haller H, Lauche R, Cramer H, Rampp T, Saha FJ, Ostermann T et al. Craniosacral therapy
57	for the treatment of chronic neck pain: a randomized sham-controlled trial. The Clinical journal
58	of pain. 2016;32(5):441-9.
59	
60	
	20
	28

**BMJ** Open

35. Klein R, Bareis A, Schneider A, Linde K. Strain–counterstrain to treat restrictions of the mobility of the cervical spine in patients with neck pain—A sham-controlled randomized trial. Complementary Therapies in Medicine. 2013;21(1):1–7.

36. Younes M, Nowakowski K, Didier-Laurent B, Gombert M, Cottin F. Effect of spinal manipulative treatment on cardiovascular autonomic control in patients with acute low back pain. Chiropractic and Manual Therapies. 2017;25(1).

37. Quinn F, Hughes CM, Baxter GD. Reflexology in the management of low back pain: A pilot randomised controlled trial. Complementary Therapies in Medicine. 2008;16(1):3-8.

38. Sillevis R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. . Journal of Manual & Manipulative Therapy *,*. 2010;18.

39. Cleland JA, Childs MJD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. *Manual therapy*. 2005;10(2):127-35.

40. Pires PF, Packer AC, Dibai-Filho AV, Rodrigues-Bigaton D. 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. Journal of manipulative and physiological therapeutics. 2015;38(8).

41. Silva AC, Biasotto-Gonzalez DA, Oliveira FHM, Andrade AO, Gomes CAFDP, Lanza FDC 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. . Evidence-Based Complementary and Alternative Medicine. 2018;2018:1-9.

42. Vieira-Pellenz F, Oliva-Pascual-Vaca Á, Rodriguez-Blanco C, Heredia-Rizo AM, Ricard F, Almazán-Campos G. 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. Archives of physical medicine and rehabilitation. 2014;95(9):1613-9.

43. Kogure A, Kotani K, Katada S, Takagi H, Kamikozuru M, Isaji T et al. A randomized, singleblind, placebo-controlled study on the efficacy of the arthrokinematic approach-hakata method in patients with chronic nonspecific low back pain. PloS one. 2015;10(12):1-16.

44. Hall T, Hardt S, Schafer A, Wallin L. Mulligan bent leg raise technique-a preliminary randomized trial of immediate effects after a single intervention. Manual Therapy. 2006;11(2):130-5.

45. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. <u>https://handbook-5-</u> <u>1.cochrane.org/index.htm#chapter 8/8 assessing risk of bias in included studies.htm</u>: The Cochrane Collaboration; 2011.

46. Dunning J, Mourad F, Barbero M, Leoni D, Cescon C, Butts R. Bilateral and multiple cavitation sounds during upper cervical thrust manipulation. BMC musculoskeletal disorders. 2013;14(1):24.

47. Johnston BC, Ebrahim S, Carrasco-Labra A, Furukawa TA, Patrick DL, Crawford MW et al. Minimally important difference estimates and methods: a protocol. BMJ Open. 2015;5:1-5.

48. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials. 338. 2009;b1732.

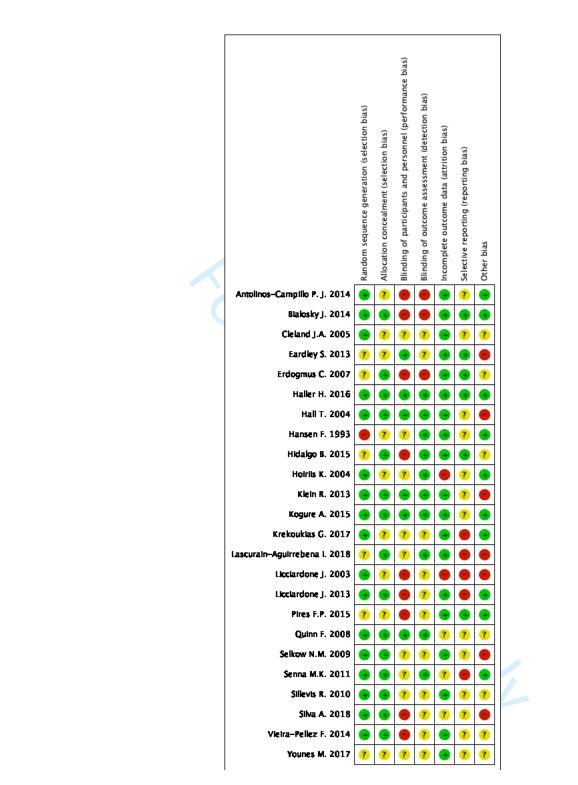
49. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. . International journal of surgery. 2012;10(1):28-55.

50. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PloS one. 2013;8(3).

51. Ruddock JK, Sallis H, Ness A, Perry RE. Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. Journal of chiropractic medicine. 2016;15(3):165-83.

52. Rubinstein SM, De Zoete A, Van Middelkoop M, Assendelft WJ, De Boer MR, Van Tulder MW. 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.

Rubinstein SM, van Eekelen R, Oosterhuis T, de Boer MR, Ostelo RW, van Tulder MW. Ec. e size o, mendations. 53. The risk of bias and sample size of trials of spinal manipulative therapy for low back and neck pain: analysis and recommendations. Journal of manipulative and physiological therapeutics. 2014;37(8):523-41.



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

2											
3		F	lacebo			МТ			Mean Difference	Mean Difference	<b>Risk of Bias</b>
4	Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% CI	ABCDEFG
	4.1.1 Short-term	•									
5	Antolinos-Campillo PJ 2014 Bialosky J 2014	55.2	10.5 22.45	20 28	48.7 34.03	10.8 26.21	20 27	0.8%	6.50 [-0.10, 13.10] -2.28 [-15.20, 10.64]		• ? • • • ? •
6	Cleland JA 2005	43.5	19.5	17	26.1	17.2	19	0.2%	17.40 [5.33, 29.47]		
7	Eardley S 2013	36	19.8	21	29	22.7	20	0.2%	7.00 [-6.06, 20.06]		224244
	Erdogmus C 2007	21.6	3.13	40	24.13	2.5	40	0.0%	-2.53 [-3.77, -1.29]		<b>? • • • •</b> • • 7
8	Haller H 2016	53.5	20.3	27	31.7	20.7	27	0.3%	21.80 [10.86, 32.74]	<del></del>	<b></b>
9	Hoirits K 2004	22.1	20.2	40	17.1	16.6	34	0.4%	5.00 [-3.89, 13.89]		• ? ? • • ? •
	Kieln R 2013	36	26 21.1	31	32 43.4	22 20	30	0.2%	6.00 [-6.07, 18.07]		
10	Kogure A 2015 Krekouklas G 2017	47.6 58.8	9.2	89 25	43.4	11	90 25	0.9%	4.20 [-1.82, 10.22] 46.60 [40.98, 52.22]	T	
11	Licclardone J 2003	30.7	21.9	23	37.7	26.2	42	0.2%	-7.00 [-18.95, 4.95]		
12	Pires FP 2015	25.7	22.7	16	22	21.7	16	0.1%	3.70 [-11.69, 19.09]		220200
	Selkow NM 2009	35.2	28	10	25	20.6	10		10.20 [-11.34, 31.74]	<u>+</u>	🗣 🗣 🖓 🗣 🥐 🛑
13	Senna MK 2011	33.19	1.19		29.46	1.16	26	95.3 <b>%</b>	3.73 [3.14, 4.32]		•••?••
14	Silva A 2018	47.1	17.2	14		20.01	14	0.2%	15.00 [1.18, 28.82]		
15	Vieira-Pellez F 2014 Younes M 2017	29.1 18.14	26.37 7.46	20	20.01 12.6	6.24	20 10	0.1% 0.7%	9.09 [-6.09, 24.27] 5.54 [-1.21, 12.29]		
	Subtotal (95% CI)	10.14	7.40	400	12.0	9.24		100.0%	3.86 [3.29, 4.43]		
16	Heterogeneity: Chi <sup>2</sup> = 23.95	. df = 14	(P = 0.0)		- 42%			20010/1	5100 [0120] 110]	'	
17	Test for overall effect: $Z = 1$										
18	Total (95% CI)			400			405	100.0%	3.86 [3.29, 4.43]	1	
19	Heterogeneity: $Ch^2 = 23.95$ Test for overall effect: $Z = 1$				42%					-50 -25 0 25 50	_
20	Test for subgroup difference			,1,						Favours Placebo Favours MT	
	Risk of bias legend										
21	(A) Random sequence gener	ation (sel	ection bi	as)							
22	(B) Allocation concealment (s			_							
23	(C) Blinding of participants a				ce bias)						
	<ul> <li>(D) Blinding of outcome asse</li> <li>(E) Incomplete outcome data</li> </ul>			Dias)							
24	(F) Selective reporting (repor										
25	(G) Other bias										
26											

Figure 3: Forest plot of comparison Placebo versus MT in back pain outcome at short-term.

Study or Subgroup         Events         Total         Events         Total         Weight         M-H, Fixed, 95% Cl         M-H, Fixed, 95% Cl         A B C D E F           Bialosky J 2014         1         28         1         27         1.0%         0.96 [0.06, 14.65]         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7 </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>									
Bialosky j 2014       1       28       1       27       1.0%       0.96 [0.06, 14.65]         Eardley S 2013       1       21       0       20       0.5%       2.86 [0.12, 66.44]         Erdogmus C 2007       8       40       5       40       4.8%       1.60 [0.57, 4.47]       7       7       7         Haller H 2016       13       27       11       27       10.5%       1.18 [0.65, 2.15]       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7		Place	bo	мт			Risk Ratio	Risk Ratio	Risk of Bias
Eardley S 2013       1       21       0       20 $0.5\%$ $2.86$ $[0.12, 66.44]$ Erdogmus C 2007       8       40       5       40 $4.8\%$ $1.60$ $[0.57, 4.47]$ Haller H 2016       13       27       11       27 $10.5\%$ $1.18$ $[0.65, 2.15]$ Hansen F 1993       5       59       11       57 $10.7\%$ $0.44$ $[0.49, 1.83]$ Holrits K 2004       13       53       13       50 $12.8\%$ $0.94$ $[0.49, 1.83]$ Kogure A 2015       4       89       3       90 $2.8\%$ $1.35$ $[0.31, 5.85]$ Licclardone J 2003       4       23       16       48 $9.9\%$ $0.52$ $[0.20, 1.39]$ Licclardone J. 2013       34       225       39       230 $36.8\%$ $0.89$ $[0.58, 1.36]$ Senna MK 2011       20       40       7       27 $8.0\%$ $1.93$ $[0.95, 3.92]$ 7       7         Younes M 2017       2       7 $3$ $102$ $4.8\%$ $0.95$ $[0.21, 4.29]$ 7 $7$ $7$	Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI	M-H, Fixed, 95% Cl	ABCDEF
Erdogmus C 2007       6       40       5       40       4.8%       1.60       [0.57, 4.47]         Haller H 2016       13       27       11       27       10.5%       1.18       [0.65, 2.15]         Hansen F 1993       5       59       11       57       10.7%       0.44       [0.16, 1.18]       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7	Blakosky j 2014	1	28	1	27	1.0%	0.96 [0.06, 14.65]		
Haller H 2016       13       27       11       27       10.5%       1.18       [0.65, 2.15]         Hansen F 1993       5       59       11       57       10.7%       0.44       [0.16, 1.18]       97       9       9         Hoirits K 2004       13       53       13       50       12.8%       0.94       [0.49, 1.83]       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97       97	Eardley S 2013	1	21	0	20	0.5%	2.86 [0.12, 66.44]		??
Hansen F 1993       5       59       11       57       10.7%       0.44 [0.16, 1.18]         Holriis K 2004       13       53       13       50       12.8%       0.94 [0.49, 1.83]         Kogure A 2015       4       69       3       90       2.8%       1.35 [0.31, 5.85]         Loclardone J 2003       4       23       16       48       9.9%       0.52 [0.20, 1.39]       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7	Erdogmus C 2007	6	40	5	40	4.6%	1.60 [0.57, 4.47]		?
Hoirits K 2004       13       53       13       50       12.8% $0.94$ [ $0.49$ , $1.83$ ]       - $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? $0.7$ ? <td>Haller H 2016</td> <td>13</td> <td>27</td> <td>11</td> <td>27</td> <td>10.5%</td> <td>1.18 [0.65, 2.15]</td> <td></td> <td></td>	Haller H 2016	13	27	11	27	10.5%	1.18 [0.65, 2.15]		
Kogure A 2015       4       89       3       90       2.8%       1.35 [0.31, 5.85]         Licclardone J 2003       4       23       16       48       9.9%       0.52 [0.20, 1.39]         Licclardone J. 2013       34       225       39       230       36.8%       0.89 [0.58, 1.36] $\bullet$ $\bullet$ $\circ$	Hansen F 1993	5	59	11	57	10.7%	0.44 [0.16, 1.18]		
Licclardone j 2003       4       23       16       48       9.9% $0.52 [0.20, 1.39]$ Licclardone J. 2013       34       225       39       230       36.8% $0.89 [0.58, 1.36]$ Senna MK 2011       20       40       7       27 $8.0\%$ $1.93 [0.95, 3.92]$ Younes M 2017       2       7       3       10 $2.4\%$ $0.95 [0.21, 4.29]$ Total (95% Cl)       612       626       100.0% $0.98 [0.77, 1.25]$ Heterogeneity: Chi <sup>2</sup> = 9.69, df = 10 (P = 0.47); i <sup>2</sup> = 0%       0.01       0.1       1       10	Hoiriis K 2004	13	53	13	50	12.6%	0.94 [0.49, 1.83]		
Licctardone J. 2013 $34$ $225$ $39$ $230$ $36.8\%$ $0.89$ $[0.58, 1.36]$ Senna MK 2011 $20$ $40$ $7$ $27$ $8.0\%$ $1.93$ $[0.95, 3.92]$ Younes M 2017 $2$ $7$ $3$ $10$ $2.4\%$ $0.95$ $[0.21, 4.29]$ Total (95% Cl)       612       626 $100.0\%$ $0.98$ $[0.77, 1.25]$ Total events       105 $109$ Heterogenetty: Ch <sup>2</sup> = 9.69, df = 10 (P = 0.47); l <sup>2</sup> = 0% $0.01$ $0.1$ $1$ $10$ $100$	Kogure A 2015	4	69	3	90	2.6%	1.35 [0.31, 5.85]		
Senna MK 2011       20       40       7       27       8.0%       1.93 [0.95, 3.92]         Younes M 2017       2       7       3       10       2.4%       0.95 [0.21, 4.29]         Total (95% Cl)       612       626       100.0%       0.98 [0.77, 1.25] $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$ $?$	Licclardone J 2003	4	23	16	46	9.9%	0.52 [0.20, 1.39]		
Younes M 2017       2       7       3       10       2.4%       0.95 $[0.21, 4.29]$ 7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7       7 <th7< th="">       7       7</th7<>	Licclardone J. 2013	34	225	39	230	36.8X	0.69 [0.56, 1.36]		••••?••
Total (95% Cl)       612       626       100.0%       0.98 [0.77, 1.25]         Total events       105       109         Heterogenetity: $Chi^2 = 9.69$ , $df = 10$ ( $P = 0.47$ ); $i^2 = 0\%$ 0.01       0.1       1       10       100	Senna MK 2011	20	40	7	27	8.0%	1.93 [0.95, 3.92]		
Total events       105       109         Heterogeneity: $Chi^2 = 9.69$ , $df = 10$ (P = 0.47); $l^2 = 0\%$ 0.01       0.1       1       10       100	Younes M 2017	2	7	3	10	2.4%	0.95 [0.21, 4.29]		222242
Heterogeneity: $Chi^2 = 9.69$ , $df = 10$ (P = 0.47); $l^2 = 0\%$ 0.01 0.1 1 10 100 0.01 0.1 1 10 100	Total (95% CI)		612		626	100.0%	0.98 [0.77, 1.25]	•	
	Total events	105		109					
Test for overall energy Favours MT					);	)%		0.01 0.1 1 10 100 Favours Placebo Favours MT	j.
					s)				
(A) Random sequence generation (selection bias)	,,								
(B) Allocation concealment (selection bias)						e bias)			
<ul> <li>(B) Allocation concealment (selection bias)</li> <li>(C) Blinding of participants and personnel (performance bias)</li> </ul>	(D) Blinding of outcor	ne assessn	nent (d	etection l	oias)				

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias) (G) Other bias

Figure 4: Forest plot of comparison Placebo versus MT in number of dropouts outcome 

1						
2 3		Placebo	мт	Risk Ratio	Risk Ratio	Risk of Bias
4	Study or Subgroup Eardley \$ 2013	Events Total I 2 21	Events Total Weight 1 20 2.6%		M–H, Fixed, 95% Cl	
5	Erdogmus C 2007	11 40	12 40 31.0%	0.92 [0.46, 1.83]	+	200000
6	Haller H 2016 Hansen F 1993	8 27 0 59	6 27 15.5% 7 57 19.7%	1.33 [0.53, 3.33] 0.06 [0.00, 1.10]		
7	Klein R 2013 Kogure A 2015	1 31 10 69	4 30 10.5% 6 90 20.6%	0.24 [0.03, 2.04] 1.26 [0.52, 3.05]		
8 9	Total (95% CI)	267	264 100.0%	0.84 [0.55, 1.28]		
10	Total events	32	38	0.04 [0.55, 1.20]		
11	Heterogeneity: Chi <sup>2</sup> = Test for overall effect	• 6.79, df = 5 (P = t: Z = 0.81 (P = 0.4	0.24);		0.001 0.1 1 10 1000 Favours Placebo Favours MT	
12	<u>Risk of bias legend</u>					
13	(A) Random sequence (B) Allocation conceal					
14 15	(C) Blinding of partici	pants and personn	el (performance bias)			
16	<ul> <li>(D) Blinding of outcom</li> <li>(E) Incomplete outcom</li> </ul>					
17	<ul> <li>(F) Selective reporting</li> <li>(G) Other bias</li> </ul>	g (reporting bias)				
18						
19	Figure F. F.	ot plat af a		aba warawa 🎙	IT in number of and a	rea avanta autoarra a l
20	-	st plot of C	omparison Plac	ebo versus M	n in number of adve	rse events outcome at
21 22	short-term					
22						
24						
25						
26						
27						
28 29						
30						
31						
32						
33						
34 35						
36						
37						
38						
39						
40 41						
41						
43						
44						
44 45 46						
46 47						
47 48						
49						
50						
51						
52 53						
54						
54 55						
56						
57						
58 50						
59 60						
00						

	Place	bo	Cont	rol		Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI	M-H, Fixed, 95% CI	ABCDEFG
Białosky j 2014	1	28	0	28	1.4%	3.00 [0.13, 70.64]		
Eardley S 2013	1	21	4	25	10.6%	0.30 [0.04, 2.46]		?? . ?
Erdogmus C 2007	8	40	6	40	23.2%	1.00 [0.42, 2.40]	<b>_</b>	? • • • • • ?
Hoirils K 2004	13	53	17	53	49.3X	0.76 [0.41, 1.41]		
Licclardone J 2003	4	23	5	20	15.5%	0.70 [0.22, 2.24]		<b>♀</b> ? ●? ●●€
Total (95% CI)		165		166	100.0%	0.79 [0.51, 1.23]	•	
Total events	27		34				-	
Heterogeneity: Chi <sup>2</sup> =	1.84, df	= 4 (P	= 0.77);	$l^2 = 0\%$	1			
Test for overall effect:	z = 1.03	(P = 0	).30)				0.01 0.1 1 10 1 Favours Placebo Favours Contr	00 <sup>°</sup>

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Figure 6: Forest plot of comparison Placebo versus MT in number of dropouts outcome

Medline
1. Mesh descriptor: [Back Pain] explode all trees
2. dorsalgia/
3. backache
4.(neck OR cervical) adj1 pain → Mesh
5. exp Brachial Plexus Neuropaties
6. exp Lumbar Plexus Neuropaties
7. Neck Pain/
8. neckache
9. Torticollis/
10. whiplash.mp
11. cervicodynia.mp
12. spondylitis/ OR spondylosis/ OR spondylolysis/ OR spondylolysthesis
13.(lumbar OR dorsal OR neck OR cervical OR sciatica) adj2 (pain OR ache)
14. (lumbar OR dorsal OR neck OR cervical) adj2 (discitis OR disc adj 1 herniation OR disc adj1
herniation)
15. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
16. (PLACEBO* or MOCK* or SHAM* or FAKE* or VEHICLE* or DUMM* or ATTENTION* CONTR
or PSEUDO* TREAT* or UN?SPECIFIC* or NON?SPECIFIC* OR simulat\$ treatment OR inert age
17. Chiropractic/
18. Manipulation, Chiropractic/
19. chiropract\$.tw.
20. (manual adj2 therap\$).mp
21. spinal manipulation.mp. or Manipulation, Spinal/ 🚫
22. osteopath\$.tw.
23. Osteopathic Medicine/
24. Physical Therapy Modalities/ or "Physical Therapy (Specialty)"/ or physical therap\$.tw. or
physiotherap\$.tw.
25. myotherapy.mp
26. shiatsu.mp
27.exp Therapeutic Touch/
28. exp Massage/
29. (neuromuscular adj therapy).mp
30. 17 OR 18 OR 19 OR 20 OR OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29
31. pain
32. range of motion
33. ROM
34. 31 OR 32 OR 33
35. Clinical Trial/
36. Randomized Controlled Trial/
37. controlled clinical trial/
38. exp RANDOMIZATION/
39. PLACEBO/
40. (random\$ adj2 allocat\$).tw.
41. single blind\$.tw.

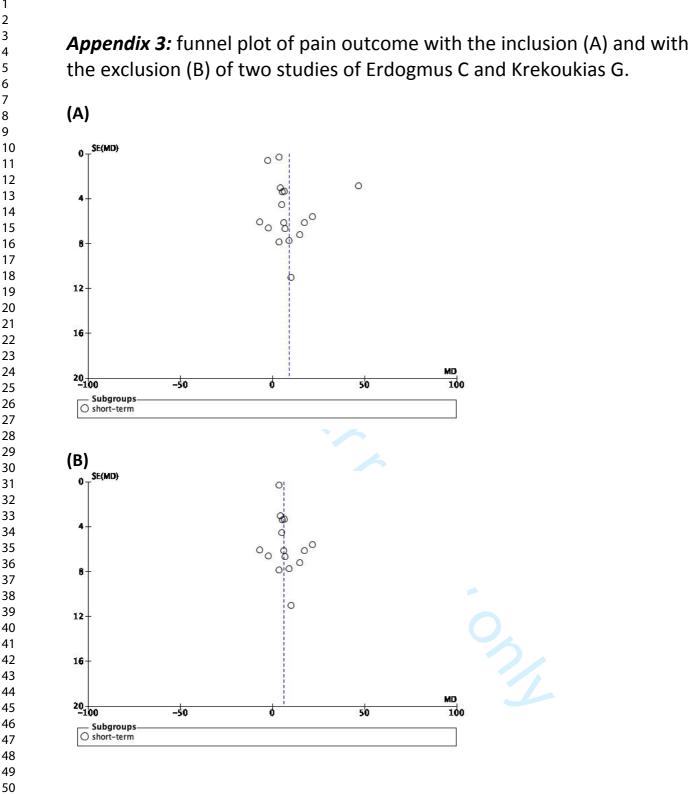
3 4	42. double blind\$.tw.
5	43. placebo\$.tw.
6	44. 35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43
7	45. animals/
8 9	46. humans/
9 10	47. 46 NOT 45
11	48. 15 AND 16 AND 30 AND 44 AND 47
12	
13	
14	
15 16	
10	
18	
19	
20	
21	
22 23	
23 24	
25	
26	
27	
28	
29 30	
31	
32	
33	
34	
35 36	
30 37	
38	
39	
40	
41 42	
42 43	
44	
45	
46	
47	
48 49	
49 50	
51	
52	
53	
54	
55 56	
56 57	
58	
59	
60	

# Appendix 2: data extraction form

Methods	Trial Design
	Settings
Deuticipente	Total mushes of soutising sta
Participants	Total number of participants:
	Age:
	Gender(M/F):
	BMI:
1	Activity:
	Duration of the symptoms:
	<b>Location of pain (</b> one-sided, double-sided, central, cervical, dorsal or lumbar):
	<b>Cause of pain:</b> (e.g. disc herniation, contractures, aspecific pain)
	Previous experience of the treatment provided: Y/N/ N/A
	Inclusion and exclusion Criteria:
	VAS:
	Practitioner characteristics: (years of experience, gender)
Interventions	Placebo:
	Comparator:
Outcomes	Outcomes used in the meta-analysis:
	Length of follow-up:
Notes	Difference between Placebo and active treatment:
	Placebo check for reliability:

Adverse event:
Lost to follow-up:
Funding source:

to per terien on



Appendix 4: forest plot of comparison pain outcome placebo vs manual therapies with the inclusion of two trials (Erdogmus C and Krekoukias G) at short, medium and long-term.

3											
4		Р	lacebo			мт			Mean Difference	Mean Difference	<b>Risk of Bias</b>
5	Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
6	3.1.1 Short-term Antolinos-Campillo P. J. 2014	55.2	10.5	20	48.7	10.8	20	4.6%	6.50 [-0.10, 13.10]	-	
0 7	Białosky J. 2014	31.75		10000	34.03		27		-2.28 [-15.20, 10.64]		
	Cleland J.A. 2005	43.5	19.5	17	26.1	17.2	19	2.4%	17.40 [5.33, 29.47]	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	
8	Eardley S. 2013	36	19.8	21	29	22.7	20	2.2%	7.00 [-6.06, 20.06]	a diama	2292999
9	Erdogmus C. 2007 Haller H. 2016	21.6 53.5	3.13 20.3	40 27	24.13 31.7	2.5	40 27	7.1%	-2.53 [-3.77, -1.29] 21.80 [10.86, 32.74]	· · · · · · · · · · · · · · · · · · ·	?
10	Hoirlis K. 2004	22.1	20.2	40	17.1	18.6	34	3.5%	5.00 [-3.89, 13.89]		
	Klein R. 2013	36	26	31	32	22	30	2.4%	6.00 [-6.07, 18.07]		
11	Kogure A. 2015	47.6	21.1	69	43.4	20	90	4.9%	4.20 [-1.82, 10.22]	+	0000070
12	Krekouklas G. 2017 Licclardone J. 2003	58.8 30.7	9.2 21.9	25 23	12.2	11 26.2	25 42	5.1% 2.5%	46.60 [40.98, 52.22] -7.00 [-18.95, 4.95]		
13	Pires F.P. 2015	25.7	22.7	16	22	21.7	16	1.7%	3.70 [-11.69, 19.09]		2202000
14	Selkow N.M. 2009	35.2	28	10	25	20.6	10		10.20 [-11.34, 31.74]		
	Senna M.K. 2011	33.19	1.19	0.5020	29.46	1.16	26	7.2%	3.73 [3.14, 4.32]	•	••?•?•
15	Silva A. 2018 Vieira-Pellez F. 2014		17.2	14 20		20.01	14 20	2.0%	15.00 [1.18, 28.82]	5 BA 20	
16	Younes M. 2017	18.14		20	12.6	6.24	10	4.5%	9.09 [-6.09, 24.27] 5.54 [-1.21, 12.29]		2222022
17	Subtotal (95% CI)	10.1.1		465			470	57.8%	8.96 [4.45, 13.47]	•	
18	Heterogeneity: Tau <sup>2</sup> = 62.49;			f = 16	(P < 0.0	00001);	r <sup>2</sup> = 95	×			
	Test for overall effect: Z = 3.8	9 (P < 0.0	001)								
19	3.1.2 Medium-term										
20	Erdogmus C. 2007	18.35	3.03	40	17.11	5.48	40	6.9%	1.24 [-0.70, 3.16]	· · · · ·	? 🗣 🛛 🗣 🗣 ?
21	Haller H. 2016	47.8	19.3	27	31.6	19	27	3.0%	16.20 [5.98, 26.42]		0000000
22	Kogure A. 2015 Senna M.K. 2011	46.3 35.16	22.2	89	36 35.16	17.4	90 26	5.0% 7.2%	10.30 [4.45, 16.15] 0.00 [-0.64, 0.64]		
	Subtotal (95% CI)	JJ.14	1.69	193	33.10	1.60	183	22.0%	3.93 [0.29, 7.57]	•	
23	Heterogeneity: Tau <sup>2</sup> = 9.14; C			3 (P <	0.0001	); $f^2 = 6$	37%			-	
24	Test for overall effect: $Z = 2.1$	2 (P = 0.0)	)3)								
25	3.1.3 Long-term										
26	Erdogmus C. 2007	16.85	3.71		21.05	3.58	40	7.0%	-4.20 [-5.80, -2.60]		?
27	Kogure A. 2015 Licclardone J. 2003	45.5 24.5	22 21.1	89 19	31.2 31.6	18.8	26 32	3.6% 2.4%	14.30 [5.75, 22.85] -7.10 [-19.36, 5.16]		
28	Senna M.K. 2011	36.8	1.39		35.53	2.13	26	7.1%	1.27 [0.34, 2.20]	·	
	Subtotal (95% CI)			185			124	20.1%	0.70 [-4.32, 5.71]	+	
29	Heterogeneity: Tau <sup>2</sup> = 18.02;			= 3 (P	< 0.000	)01); ř	93%				
30	Test for overall effect: Z = 0.2	/ (r = 0.7	(9)								
31	Total (95% CI)			843				100.0%	6.35 [4.02, 8.67]	•	
32	Heterogeneity: Tau <sup>2</sup> = 19.51;			f = 24	(P < 0.0)	00001);	r <sup>2</sup> = 95	×		-50-25 0 25 50	
	Test for overall effect: Z = 5.3 Test for subgroup differences:			2 /8 -	0.051	r _ cc	04			Favours Placebo Favours MT	
33	Risk of bias legend	- Q.	v-, ui -	- <b>-</b> 11 -							
34	(A) Random sequence generat			)							
35	(B) Allocation concealment (sel										
36	<ul> <li>(C) Blinding of participants and</li> <li>(D) Blinding of outcome assess</li> </ul>				bias)						
	(E) Incomplete outcome data (a										
37	(F) Selective reporting (reportin	ng bias)									
38	(G) Other bias										
39											
40											
41											
42											
43											

# Appendix 4.1: Sensitivity analysis with the exclusion of Ergogmus C and Krekoukias G at short, medium and long-term

	Study or Subgroup	Mean	lacebo	Total	Mean	MT	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
ŀ	3.1.1 Short-term	mean	50	Total	mean	50	Total	weight	14, Randoni, 55% Ci		ADCDETG
;	Antolinos-Campillo P. J. 2014	55.2	10.5	20	48.7	10.8	20	5.6%	6.50 [-0.10, 13.10]	-	9? 009? 9
	Blakosky J. 2014		22.45		34.03		27		-2.28 [-15.20, 10.64]	107 100 100	
	Cleland J.A. 2005	43.5		17			19	2.2%	17.40 [5.33, 29.47]	2	
	Eardley S. 2013	36		21	Acres 1200		20	2.0%	7.00 [-6.06, 20.06]	201 12 12 12 12	22929999
	Erdogmus C. 2007 Haller H. 2016	21.6 53.5	3.13 20.3	27	24.13 31.7	2.5	40 27	0.0%			
	Holrits K. 2004	22.1					34	3.7%	5.00 [-3.89, 13.89]		
	Kieln R. 2013	38	26	31		22	30	2.2%	6.00 [-6.07, 18.07]		
)	Kogure A. 2015	47.6	21.1	89		20	90	6.2%	4.20 [-1.82, 10.22]		0000070
	Krekouklas G. 2017	58.8	9.2			11	25	0.0%			
	Licelardone J. 2003	30.7		23			42			0.000	9707000
2	Pires F.P. 2015 Selkow N.M. 2009	25.7 35.2	22.7	16 10	22 25	21.7 20.6	16 10	1.5%	3.70 [-11.69, 19.09] 10.20 [-11.34, 31.74]	2.5- <sup>2</sup> 2	220200
	Senna M.K. 2011	33.19		0.000	29.46		26	15.0%	3.73 [3.14, 4.32]		
	Silva A. 2016	47.1		14		20.01	14	1.6%	15.00 [1.18, 28.82]		
•	Vieira-Pellez F. 2014		26.37	0.77.02	20.01		20	1.5%	9.09 [-6.09, 24.27]		
;	Younes M. 2017		7.46	7		6.24	10	5.4%	5.54 [-1.21, 12.29]		2222922
	Subtotal (95% CI)			400			405	54.8%	6.09 [3.23, 8.96]	•	
	Heterogeneity: Tau <sup>2</sup> = 9.37; C			• 14 (P	= 0.05)	; l <sup>2</sup> = 42	2%				
,	Test for overall effect: $Z = 4.1$	6 (P < 0.0	0001)								
	3.1.2 Medium-term										
	Erdogmus C. 2007	18.35	3.03	40	17.11	5.48	40	0.0%	1.24 [-0.70, 3.16]		? • • • • • ?
	Haller H. 2016	47.8		27		19	27	3.0%	16.20 [5.98, 26.42]		0000000
	Kogure A. 2015	46.3	22.2	0.2.2.2	36	17.4	90	6.5%	10.30 [4.45, 16.15]	+	9999979
	Senna M.K. 2011 Subtotal (95% CI)	35.16	1.25	153	35.16	1.28	26 143	15.0% 24.4%	0.00 [-0.64, 0.64] 7.93 [-1.92, 17.79]		99797
	Heterogeneity: Tau <sup>2</sup> = 65.39;	$Cht^2 = 21$	26. df		< 0.000	11): F =		24.470	1.55 [ 1.52, 11.15]		
	Test for overall effect: Z = 1.5										
		•									
	3.1.3 Long-term	10.00	3 71	40	21.05	3 6 8	40	0.08	4 20 1 5 80 2 501		7
	Erdogmus C. 2007 Kogure A. 2015	16.85 45.5	3.71 22		21.05 31.2		40 26	0.0%	-4.20 [-5.80, -2.60] 14.30 [5.75, 22.85]		
	Licclardone J. 2003	24.5		19	E. (1) (1) (2)		32	2.2%			
	Senna M.K. 2011		1.39	0.00759	35.53		26	14.7%	1.27 [0.34, 2.20]	•	
	Subtotal (95% CI)			145			84	20.8%	3.33 [-6.29, 12.95]	•	
	Heterogeneity: Tau <sup>2</sup> = 56.21; Test for overall effect: Z = 0.6			= 2 (P	= 0.005	i);	31%				
	rest for overall effect. L = 0.0	u (r — v									
	Total (95% CI)			698				100.0%	4.97 [3.01, 6.93]	• • • •	
	Heterogeneity: Tau <sup>2</sup> = 6.56; C Test for overall effect: Z = 4.9	1	1	= 20 (	P < 0.00	)001); r	= 657	3		-50-25 0 25 50	
	Test for subgroup differences:			2 (P	0 80)	$f^2 = 0.94$				Favours Placebo Favours MT	
	Risk of bias legend	- v.		- ()	0.00)						
	(A) Random sequence generation			5)							
	(B) Allocation concealment (sel				12.12						
	(C) Blinding of participants and				e bias)						
	<ul> <li>(D) Blinding of outcome assess</li> <li>(E) Incomplete outcome data (a)</li> </ul>			las)							
	(F) Selective reporting (reportin		(13)								
	, , ,	-									

(G) Other bias

# Appendix 5: forest plot of comparison Placebo versus control in back pain outcome at short-term with exclusion of Erdogmus C 2007 trial

	F	lacebo		(	Control			Mean Difference	Mean Difference	<b>Risk of Bias</b>
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Blaksky J. 2014	31.75	22.45	27	31.39	26.21	28	31.3%	0.36 [-12.52, 13.24]	-	
Eardley S. 2013	36	19.8	21	57.7	19.2	21	33.2%	-21.70 [-33.50, -9.90]		2292999
Erdogmus C. 2007	21.6	10.65	40	19.25	10.96	40	0.0%	2.35 [-2.39, 7.09]		?
Hoirils K. 2004	24.4	22.2	34	31.8	24	40	35.4%	-7.40 [-17.94, 3.14]	-	977997
Total (95% CI)			82			89	100.0%	-9.72 [-21.90, 2.47]	•	
Heterogeneity: Tau <sup>2</sup> -	- 80.12;	$Cht^2 = 6$	6.50, d	f = 2 (P	= 0.04	$f^2 = 6$	9%		-100 -50 0 50 100	ł
Test for overall effect	: Z = 1.5	6 (P = (	).12)						Favours Placebo Favours Control	
Risk of bias legend										
(A) Random sequence	e generat	tion (sele	ection b	ias)						
(B) Allocation conceal	ment (sel	ection b	ias)							
(C) Blinding of particip	pants and	d person	nnel (pe	rforman	ce bias)					
(D) Blinding of outcom	ne assess	sment (d	etection	n bias)						

(E) Incomplete outcome data (attrition bias) (F) Selective reporting (reporting bias)

(G) Other bias



# PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #					
TITLE								
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Page 1					
ABSTRACT								
Structured summary	Structured summary 2 Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criter participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.							
INTRODUCTION								
Rationale	3	Describe the rationale for the review in the context of what is already known.	Pages 9- 10					
Objectives	bjectives 4 Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).							
METHODS								
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Pages 6 and 10					
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Pages 10-11					
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Page 11					
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix 1					
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Page 12					
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Page 12					
) Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Page 12- 13					
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Page 12					
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	Pages 12-13					

Page 45 of 45

ERLIS MR

2 3

# PRISMA 2009 Checklist

Synthesis of results		Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I <sup>2</sup> ) for each meta-analysis.	Page 13
		Page 1 of 2	
Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Pages 12- 13
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Page 13
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Page 15
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Pages 16- 19
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Page 20, figure 2
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Pages 21- 23
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Pages 7-8
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Pages 20- figures 2,3,4,5,6
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Appendices 4-6
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Pages 7-8, 24
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Page 25
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Page 26
FUNDING		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

BMJ Open

BMJ Open



# PRISMA 2009 Checklist

5 4 5 6	Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Page 2
7 8	From: Moher D, Liberati A, Tetzl doi:10.1371/journal.pmed1000097	aff J, Altm	nan DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS N	Med 6(6): e1000097.
9	uoi. 10. 137 1/journal.pmeu 1000097		For more information, visit: www.prisma-statement.org.	
10			Page 2 of 2	
11				
12				
13				
14 15				
16				
17				
18				
19				
20 21				
22				
23				
24				
25				
26 27				
27				
29				
30			For more information, visit: www.prisma-statement.org. Page 2 of 2	
31				
32 33				
33 34				
35				
36				
37				
38 39				
39 40				
41				
42				
43				
44			For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	
45 46			To peer review only integr/binjopen.onlj.com/site/about/guidennes.html	
40				

**BMJ** Open

# **BMJ Open**

#### Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pair-wise metaanalysis

Journal:	BMJ Open
Manuscript ID	bmjopen-2020-045106.R1
Article Type:	Original research
Date Submitted by the Author:	04-Feb-2021
Complete List of Authors:	Lavazza, Carolina; AIMO, Research Galli, Margherita; AIMO, Research Abenavoli, Alessandra; AIMO, Research Maggiani, Alberto; AIMO, Research
<b>Primary Subject Heading</b> :	Complementary medicine
Secondary Subject Heading:	Research methods
Keywords:	COMPLEMENTARY MEDICINE, STATISTICS & RESEARCH METHODS, Back pain < ORTHOPAEDIC & TRAUMA SURGERY





I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our <u>licence</u>.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which <u>Creative Commons</u> licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

reliez oni

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Address for each author Separate each part of address with a comma (don't use a separate line for each part) Please include postcodes Authors' names and positions Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Carolina Lavazza (0000-0002-9907-6819), Margherita Galli (0000-0003-4050-3131), A Abenavoli (0000-0001-8941-2308), A Maggiani (0000-0002-7980-9405) Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Carolina Lavazza associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor, Department of Research AIMO institute,
Address for each author Separate each part of address with a comma (don't use a separate line for each part) Please include postcodes Authors' names and positions Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Carolina Lavazza associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor,
Separate each part of address with a comma (don't use a separate line for each part) Please include postcodes <b>Authors' names and positions</b> Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Piazzale Santuario 7, 21047 Saronno, Italy, Carolina Lavazza associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor,
(don't use a separate line for each part) Please include postcodes <b>Authors' names and positions</b> Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Carolina Lavazza associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor,
Please include postcodes <b>Authors' names and positions</b> Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor,
Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name	Margherita Galli associate professor,
give one job position for each author, (under the relevant address) on separate line from author's name	-
the relevant address) on separate line from author's name	Department of Research AIMO institute,
author's name	
	Piazzale Santuario 7, 21047 Saronno, Italy,
	A Abenavoli associate professor, Departmen
	of Research AIMO institute, Piazzale
Q	Santuario 7, 21047 Saronno, Italy, A Maggiar
Corresponding author	professor
corresponding aution	Correspondence to:
	carolina.lavazza@docenti.aimoedu.it

# Copyright/license for publication:

Carolina Lavazza The Corresponding Author of this article contained within the original manuscript which includes any diagrams & photographs within and any related or stand alone film submitted ("the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at:

http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyrightopen-access-and-permission-reuse.

Please tick one or more boxes as appropriate:

×

- I am the sole author of the Contribution.
- I am one author signing on behalf of all co-owners of the Contribution.
- П The Contribution has been made in the course of my employment and I am signing as authorised by my employer.
  - I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown\* acting in the course of my employment
- I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown acting in the course of my employment
- I am not an employee of the UK Crown acting in the course of my employment but some/all of my coauthors are.\*

\*Such authors should consult guidance and if necessary return any completed form.

Licence: The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Conflict of interest: I have read and understood the BMJ Group policy on declaration of interests and declare the following interests: none

Ethical approval: not required.

Transparency: The manuscript's guarantor affirms that this 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.

Funding: None.

**Competing interests**: No funding were provided for this project and authors did not received support from any organisation for the submitted work. Authors have no financial relationships with any organisations that might have an interest in this project and have no other relationships or activities that could appear to have influenced this review.

in the second seco

1	
2	
3	Acronyms
4	•
5	AE= adverse effects
6	
7	BP= back pain
8 9	CI= confidence intervals
9 10	CRB= Cochrane risk of bias
10	CT= clinical trial
12	MA = meta-analysis
13	MCID= minimal clinically important difference
14	MD=mean difference
15	
16	MT= manual therapies
17	OMT = osteopathic manipulative treatment
18	PROs = patient-reported outcomes
19	RCT= randomised controlled trials
20	RR= risk ratio
21	SM= spinal manipulation
22	ST= sham treatment
23	
24	
25 26	
20 27	
27	
29	
30	RR= risk ratio SM= spinal manipulation ST= sham treatment
31	
32	
33	
34	
35	
36	
37	
38	
39 40	
40 41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53 54	
54 55	
55 56	
57	
58	

# 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. RCT involving adults affected by BP (cervical and lumbar), acute or chronic, were included.

Hand contact sham treatment (ST) was compared to different MT (physiotherapy, chiropractic, osteopathy, massage, kinesiology and reflexology) and to control. Primary outcomes were BP improvement, success of blinding and adverse effect (AE). Secondary outcomes were number of dropouts. Dichotomous outcomes were analysed using risk ratio (RR), continuous using mean difference (MD), 95% confidence intervals (CI). The minimal clinically important difference was 30 mm changes in pain score.

# **Results:**

24 trials were included with 2,019 participants. Very low evidence quality suggests clinically insignificant pain improvement in favour of MT compared to ST (MD 3.86, 95% CI 3.29 to 4.43) and no differences between ST and control (MD -6.04, 95% CI -16.68 to 4.59).

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

Low quality of evidence suggests that AE and dropout rates were similar between ST and MT (RR AE=0.84, 95% CI 0.55 to 1.28, RR dropouts= 0.98, 95% CI 0.77 to 1.25). A similar dropout rate in control (RR=0.79, 95% 0.51 to 1.23).

# **Conclusions:**

Comparison of ST and MT shows a small, clinically meaningless effect in pain improvement. Similar effects were found with control. The heterogeneousness 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 proper sample size for the reliable detection of clinically meaningful treatment effects.

# PROSPERO register: CRD42020198301

https://www.crd.york.ac.uk/prospero/display\_record.php?RecordID=198301

# Strengths and limitations of this study

#### Strengths

This systematic review and pair-wise meta-analysis:

- summarises existing evidence on the effectiveness, reliability and application of hand contact sham treatment in MT randomised controlled trials;
- gives suggestions for researchers on conducting methodical RCT in MT using a reliable sham procedure.

#### Limitation

- This study did not include a comparison with machine provided placebo, its aim focused on hand contact sham treatment
- Insufficient number of studies were included to conduct a network meta-analysis

# Background

In Clinical Trials (CT), placebo is commonly used as a control therapy to evaluate 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)

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 (IMP)
 and as such it has to follow different standards in order to ensure quality, guarantee patient
 safety and the reliability of study results.(3)

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 to relieve pain and stiffness, promote relaxation of soft-tissues, enhancing blood supply to tissues and increase mobility of joint structures. (4)

27 In MT trials, placebo treatment is often provided in different modalities from trial to trial 28 although the manual techniques or treatments tested are the same. For instance, placebo 29 treatment is commonly administrated as a light touch in the site of pain or as an active 30 treatment in a different site (5), with no clear criterion. Such light touch might in fact have a 31 health effect and there is no evidence as to its ineffectiveness. Touch itself could have a positive 32 33 outcome on health (6) and active treatments could have an analgesic reflex on pain even if 34 administered anywhere in the body.(7) 35

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 rarely cures 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 called nocebo. It has been estimated that up to 26% of patients in randomized control trials (RCTs) discontinue placebo due to adverse effects.(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 (e.g. beliefs,
 desire for symptom changes), as well as the patient's expectations of improvement and prior
 experiences of the treatment. (10-13)

49 In pharmacological trials this overall therapeutic context and its influence on placebo response 50 has been widely studied. (11) Less evidence is present for MT trials, where other important 51 characteristics should be considered as part of this therapeutic context such as the tactile 52 interaction between patient and practitioner and clinician beliefs. (14, 15) Pharmacological 53 trials avoid the influence of clinicians' beliefs by using a placebo that ensures both patients and 54 55 clinicians blinding to treatment allocation, but, in MT trials, the blinding of clinicians is almost 56 impossible to achieve. The best alternative in this type of trial is the use of a sham treatment 57 that mimics the active treatment and aims to ensure at least the blinding of participants. 58

7

59 60

1 2 3

8

9

10 11

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 to pharmacological placebo and that this effect might be a consequence of physical contact.

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

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

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 pair-wise meta-analyses (MA) is to evaluate the use of placebo in MT trials in order to analyse the effects, possible harm and the reliability of different kinds of sham treatments provided in RCTs involving MT. A systematic review could help to define placebo 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 sham treatment in manual therapy (MT) randomized controlled trials in the treatment of back pain (BP) both cervical and lumbar in order to provide methodological guidance for clinical trial development.

# Methods

This systematic review and meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews (PRISMA)(16).

The protocol registration was performed in PROSPERO (<u>http://www.crd.york.ac.uk/PROSPERO/</u>) and review registration number is **CRD42020198301.** 

# Criteria for considering studies for this review

Only randomised controlled studies (RCTs) were included in this review. Quasi-randomised trials in which allocation was not strictly random (e.g. 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 is related to muscular conditions, articular disorders (such as osteoarthritis) or spinal disc herniation were included. Trials where musculoskeletal diseases were secondary to other pathologies (e.g. 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 placebo that include hand contact provided by all kinds of physical therapists. Studies where placebo was provided by machines (such as inactive ultrasound) 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 to other manual therapies such as: physiotherapy, chiropractic, osteopathy, massage, kinesiology and reflexology and to control.

To assess if touch itself could have a positive health effect, ST was also compared to control. Physiotherapeutic exercises were included in the analysis only if associated with manual treatment.

The use of active co-interventions such as oral 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 of interested arms were extracted.

## Outcomes

Primary outcomes were pain intensity on a validated scale, success of blinding of and adverse effect. Secondary outcomes were number of dropouts.

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

#### Information sources

Search strategy (Appendix 1) was adapted to the different databased by an experienced information specialist.

RCTs were identified in different databases (up to 20 August 2020): MEDLINE, Embase, CINAHL, SPORTDiscus, PEDro, World Health Organization Clinical Trials Registration Platform, Index to Chiropractic Literature, Cochrane central register of controlled trials (CENTRAL), Clinical trials registry and metaRegister of Controlled Trials (mRCT).

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 on-going or recently completed systematic reviews. Guidelines from different organizations (e.g. National Council for Osteopathic Research etc.) were reviewed and references from relevant publication were analysed.

# Data collection and analysis

Searches 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 firstly 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 the document delivery service of the 3Bi Biella library.

Uncertainty about the inclusion of a study were 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 pre-defined form. Data extracted were related to settings, type of study, participants characteristics (such as localization 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. (*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 (27). 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 judgments was performed by *MG*. RevMan 5.3.5 was used for the graphic representation of each risk. The CRB tool results were then converted to 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, poor quality studies instead 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 adverse events (occurred or not), were analysed using risk ratio (RR) with 95% confidence intervals (CI).

Continuous outcomes, such as back pain on VAS scale, were evaluated using mean difference (MD) between ST and the MT/control group with 95% CI and the standard deviation (SD). The minimal clinically important difference (MCID) between pre- and post-treatment was taken as 30 mm changes in 100 mm pain score. (17-19)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<sup>2</sup>).

Cochrane Handbook was used for threshold interpretation: heterogeneity was considered as unimportant for value of  $I^2$  between 0% and 40%,, as moderate for values between 30% and 60%, as substantial for values between 50% an 90% and considerable for values between 75% to 100%. (20)

# Synthesis of results

Meta-analysis of pain score, AE and dropout rates were performed using RevMan 5.3.5 whenever possible. The meta-analyses compared all kinds of ST with all types of manual therapies and to control. Random-effect model was used when a substantial inconsistency was present (I<sup>2</sup>= 50-90%). (20) When considerable heterogeneity was present (I<sup>2</sup>>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 localization 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 appendices.

# Summarizing results and assessing the quality of the evidence

The quality of evidence for each outcome was evaluated with the 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.(21) The GRADE approach was used to assess the quality of the key outcomes. The software GRADEpro (https://gradepro.org) was used to import data from RevMan 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 dropouts.

# **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

Figure 1: PRISMA flow diagram

## Included studies

Table 1 shows a summary of main characteristics of included studies.

24 studies were included in this review (*Figure 1*), one study had a 2x2 factorial design, (22) eight studies had multiple arms. (23-30)Most of the studies were conducted in physical therapy clinics, in 13 different countries. Three trials did not report in which clinical setting their were conducted. (27, 31, 32)

Eight trials were conducted in Europe ,(25, 26, 28, 33-37) five in the United States,(22, 23, 29, 38, 39) three studies in Brazil,(40-42) one in UK, (24) Egypt, (30) Japan (43) and Australia. (44) No ongoing or unpublished trials were found.

# Population

The included trials randomized a total of 2,019 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/m<sup>2</sup>.

The majority of studies included both genders, with a percentage of male that ranged from 19% to 80%. Two trials included only male, (36, 42) one study included only female participants. (40)

16 trials enrolled participants with low back pain (LBP), nine included participants with cervical pain (CP).

The majority of trials (N=18) included participants with unspecified cause of back pain. Disk herniation was considered in three trials. (25, 28, 42)

Duration of symptoms were unassessed in eight trials, most of the studies included participants with chronic pain (N=9), some included participants with both acute and chronic pain.

Participants with experience of the tested treatment were included in 8 trials(22, 27, 29, 30, 33, 35, 40, 41) and excluded in four. (24, 34, 37, 39) Remaining studies did not provide this information.

# Interventions

Interventions deferred for number of sessions and number of techniques applied. Generally the trials used a single therapy session (N=11) with a single technique performed (N=8). Trials with different therapy sessions ranged from 5(23, 24, 28) to 20 (25) sessions once a week.

# Sham treatment

ST was provided with a hand contact on the area of pain in 19 studies, five studies provided ST in a different area from where the pain was located. (25, 33, 41, 43, 44)

In trials providing spinal manipulation, as inactive treatment the majority of authors used the similar placement of hands on participants without any force applied. (38-40, 42)

Two trials used a ST with similar forces applied in different directions. (23, 30)

one trial did not specify the inactive manipulation applied. (27)

In trials that provided multiple techniques in the same treatment session (such as osteopathic treatment, spinal mobilization and physiotherapy) the ST was administrated 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. (35)

# Manual and controls treatments

Different manual treatments were provided:

- Physiotherapy (2 trials, 288 participants)
- Spinal manipulation (SM)/chiropractic (7 studies, 567 participants)
- Osteopathy (5 trials, 645 participants)
- Kinesiology (one trial, 58 participants)
- Articular mobilizations (5 trials, 325 participants)
- Muscular release (4 trials, 136 participants)

Five trials with multiple arms compared ST to control group (343 participants).

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 manual therapy and sham in a different zone. (37)

Characteristics of practitioner who administrated treatments were provided by 16 trials. Most of the trials involved physiotherapists (N=8), physical therapists (N=4), osteopaths (N=3) and students (N=1). Only seven studies provided information on years of practice experience of physicians involved that ranged from 6 to 17 years. (28, 31, 33-35, 38, 40, 42) The gender of practitioners was indicated in only three trials.(24, 28, 35)

#### 

Table 1: summary of main characteristics of include	ed studies
-----------------------------------------------------	------------

Study ID	N° of participants	Symptoms duration	Pain localization	Technique tested (site of application)	Type of sham procedure	Other arms	Follow-up
Antonilos- Campillo PJ 2014	40	Not reported	Cervical	Soft-tissue (cervical region)	Soft mobilization of lower limbs	None	No follow-up (outcomes collected after the interventio
Bialosky J 2014	110	> 4 months	Lumbar	Spinal Manipulation (SM) (lumbar spine)	Ineffective force applied on lumbar spine	Control group	2 weeks
Cleland JA 2005	36	>2 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the interventic
Eardley S 2013	58	> 3 years	Lumbar	Kinesiology <i>(spine)</i>	Protocol of ineffective techniques in the site of pain	Control group	7 weeks
Erdogmus S 2007	120	$\geq$ 1.4 weeks	Lumbar	Physiotherapy (spine)	Neck massage	Control group	1.5 years
Hall T 2004	24	Not reported	Lumbar	BLR technique (lower limbs)	Soft-tissue manipulation of the foot	None	24 hours
Haller H 2016	54	> 7 months	Cervical	Cranio-sacral therapy (head)	Ineffective touch of head	None	3 months
Hansen F 1993	168	≥ 18 days	Lumbar	Physiotherapy (lumbar spine and abdomen)	Intermittent traction of the spine	Intensive back muscle training	1 year
Hidalgo B 2015	32	Not reported	Lumbar	Articular mobilization (lumbar spine)	Ineffective mobilization forces applied on lumbar spine	None	2 weeks
Hoiriis K 2004	156	$\geq$ 2,3 weeks	Lumbar	SM (spine)	Ineffective force applied on spine	Medical treatment	4 weeks
Klein R 2003	61	>1 month and <5 years	Cervical	Strain- counterstain techniques (cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the interventic
Kogure A 2015	179	> 12 months and < 10 years	Lumbar	АКА-Н (sacro- iliac joint)	Ineffective force applied on sacro-iliac joint	None	6 months
Krekoukias G 2017	50	Not reported	Lumbar	Articular mobilization techniques (lumbar spine)	Hand contact with lumbar skin placement without any movement	Exercise plus TENS	5 weeks
Lascurain- Aguirrebena I 2018	40	Not reported	Cervical	Articular mobilization (Cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the interventio
Licciardone J 2003	91	$\geq$ 3 months	Lumbar	Osteopathic manual treatment (OMT) (all body)	Protocol of light touch techniques similar to OMT	Control group	6 months

					applied to all body		
Licciardone J 2013	455	$\geq$ 3 months	Lumbar	Osteopathic manual treatment (OMT) – (all body)	Protocol of light touch techniques similar to OMT applied to all body	None	8 weeks
Pires FP 2015	32	> 3 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	72 hours
Quinn F 2008	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 M 2009	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 seconds without any active contraction	None	24 hours
Senna MK 2011	93	$\geq$ 13 months	Lumbar	SM (lumbar spine)	Ineffective force applied on lumbar spine	Maintained SM	10 months
Sillevis R 2010	100	$\geq$ 23 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the intervention
Silva A 2019	28	>3 months	Cervical	Osteopathic visceral treatment (abdomen)	Hand contact on umbilical region without any movement	None	7 days
Veira-Pellez F 2014	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 M 2017	17	< 3 months	Lumbar	OMT (all spine)	Placebo mimed active treatment with an ineffective force applied.	None	7 days

# 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, (21) the majority of trials were judged with poor
   quality (N=22). Good quality was conferred on only two studies. (34, 43)
- 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 judgment were mainly related to trials involving spinal manipulations. These studies used a technique which can be easily recognized 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. (26, 30)
- Incomplete outcome data was the least common bias risk with 80% of trials judged as low risk.
   Reporting bias was evaluated unclear in 55% of trials where registration number and trial
   protocol were not reported or found.
   Other bias assumed was assumed was assumed at bias with for baseline differences of the
  - Other bias occurred was generally considered at high risk for baseline differences of the population in 30% of trials.

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

# Effects of intervention

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

tor occreation with

# **Table 2:** summary of findings of treatment effects and certainty of the evidence (GRADE) included for all comparisons.

## 1. Sham treatment (ST) compared to Manual Therapies (MT)

Patient or population: back pain Intervention: ST

Comparison: MT

0	· · · ·	ed absolute (95% CI)	Relative	Nº of	Certainty of	<b>C</b> ara and a
Outcomes	Risk with MT	Risk with ST	effect (95% CI)	participants (studies)	the evidence (GRADE)	Comments
Pain improvement assessed with: VAS score Scale from: 0 to 100		MD <b>3.86</b> higher (3.29 higher to 4.43 lower)		805 (15 RCTs)	⊕⊖⊖⊖ VERY LOW a,b	A small effect, not clinically relevant, in pain improvement wa detected in favour of MT. This analysis excluded two trials (one suspected of publication bias, on used a different scale) which increased heterogeneity levels bu did not affect overall efficacy meaningfully.
Adverse events assessed with: number of AE occurred	144 per 1.000	<b>121 per</b> <b>1.000</b> (79 to 184)	<b>RR 0.84</b> (0.55 to 1.28)	531 (6 RCTs)	⊕⊕⊖⊖ Low ª	Pooled data from 6 studies did no show any difference in AE occurrence between ST and MT.
Dropouts rate assessed with:	174 por	171 per 1.000	<b>RR 0.98</b> (0.77 to	1238	⊕⊕⊖⊖	Pooled data from 11 trials did no show difference in dropout rate
number of participants that leaved the study	174 per 1.000	(134 to 218)	1.25)	(11 RCTs)	LOW <sup>a</sup>	between ST and MT.
number of participants that	1.000		•		LOM <sup>a</sup>	between ST and MT.
number of participants that leaved the study	1.000 to Control		•	(11 RC15)	LOW a	between ST and MT.
number of participants that leaved the study 2. ST compared Patient or population: Intervention: ST Comparison: control	1.000 to Control back pain Anticipate		1.25) Relative	Nº of	Certainty of	
number of participants that leaved the study 2. ST compared Patient or population: Intervention: ST	1.000 to Control back pain Anticipate	218) ed absolute	1.25)			between ST and MT.
number of participants that leaved the study 2. ST compared Patient or population: Intervention: ST Comparison: control	1.000 to Control back pain Anticipate effects' Risk with	218) ed absolute (95% CI)	1.25) Relative effect	Nº of participants	Certainty of the evidence	

CI: Confidence interval; MD: Mean difference; RR: Risk ratio

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

# Explanations

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

b. Most of the studies were small trial.

c. Heterogeneity levels at 80%.

d. Number of participants < 400

to oper teries only

#### 

# Sham treatment versus other manual therapies

# Pain

The following outcomes on back pain are presented with a 100 mm visual analogue scale, 0 to 100; higher scores refer to worse pain. Trials using a 10mm scales were converted to 100mm 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.(25) Data from seven studies could not be extracted.

The meta-analysis at short-term showed substantial heterogeneity levels using a randomeffects 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.(28) This thought was verified with a funnel plot, which showed an asymmetric distribution with the inclusion of these two studies (*Appendix 3*). This sensitivity analysis did not influence overall effectiveness results but inconsistency levels decreased considerably at short-term. It can be deducted that a possible cause of heterogeneity was found (*Full analysis in 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<sup>2</sup>=42%, p<0.0001,very low quality of evidence downgraded two levels for very serious risk of bias and imprecision) **(Figure 3).** 

Figure 3: Forest plot of comparison ST versus MT in back pain outcome at short-term.

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) (*Appendix 4.1*).

# Success of blinding

Success of blinding was evaluated in five trials; one did not report the results. (28) 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% (23) to 83.5%. (27) In this last study patients were considered eligible if they already received SM.

One trial compared ST to an articular mobilization technique. 54.5% participants correctly guessed treatment allocation. (44)

Participants of one study that compared ST to reflexology had the lowest perception of the correct detection of allocation (46.7%). Participants in this trials were naïve to the type of treatment tested. (37)

# Dropouts

Pooled data from 11 trials at the last follow-up suggested no difference in dropouts rate between ST and MT at the end of the trials(105/612 compared to 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*).

Figure 4: Forest plot of comparison ST versus MT in number of dropouts outcome

# Adverse effects

Adverse effects were generally under-reported, six trials were included in the metaanalysis.(24-26, 34, 35, 43)

Two trials reported AE overall occurrence without specified event rates in the groups.(22, 30) AE were predominantly minor and lasted for two/three days after treatment, in the majority of trials transient worse pain, tiredness, muscle weakness and transient headache were reported.(24, 34, 35, 43)

Senna M 2011 reported the most common AE were local discomfort and tiredness but no serious complications were noted. (30)

Haller H 2016 reported two patients dropping out from the trial for recurrent headache after treatments, both Haller H and Klein R 2013 reported dizziness of one patient.

Licciardone J 2013 reported 27% of patients with AE, 2% had serious AE not related to study interventions. (22)

Overall results showed no clear difference in AE occurrence between ST and MT (32/267 compared to 38/264; RR 0.84, 95% CI 0.55 to 1.28; 531 participants, I<sup>2</sup>=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.

*Figure 5:* Forest plot of comparison ST versus MT in number of adverse events outcome at shortterm

# Sham treatment versus control

# Pain

Five studies compared ST to control, four were included in random-effect meta-analysis at short-term. (23-25, 27) Data from one trial could not be extracted. (29) Pooled data showed the presence of significant heterogeneity, therefore results are reported narratively: three trials showed no difference between ST and control on pain outcome, while Eardley S. 2013 showed an effect in favour of ST (pooled data from 4 trials: MD -6.04, 95%CI from -16.68 to 4.59, 252 participants, I<sup>2</sup>= 80%, P=0.27). The exclusion of Erdogmus S 2013 (that used a different scale) did not affect the results of effectiveness but decreased levels of heterogeneity (MD -9.72, 95%CI -19.94 to 0.51, I<sup>2</sup>= 69%, P=0.12) (*Appendix 5*).

# Dropouts

No differences were showed in the fixed-effect meta-analysis on dropout rate between ST and control in five trials (27/165 compared to 34/166; RR 0.79, 95% CI 0.51 to 1.23 ; 331 participants,  $I^2=0\%$ , P=0.30; very low quality of evidence downgraded two levels for very serious risk of bias and imprecision) *(Figure 6)*.

Figure 6: Forest plot of comparison ST versus control in number of dropouts outcome

## Adverse effects

Of the five studies comparing ST and control, only two reported AE.

One, Eardley S 2013, did not evaluate the AE occurred in control group while Erdogmus C 2007 reported that 10/40 in the control 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 sham treatment. However, the quality of evidence is very low, indicating that the true effect is probably markedly different from the estimated effect. Substantial levels of heterogeneity within the four studies analysed, showed no differences between sham treatment and control in pain reduction

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

AE 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 control group with no significant difference from ST.

The performance bias was the bias that recurred most with a possible or unclear presence of
 lack of participants blinding in 76% of the studies included in this review.

SM techniques were the treatment most evaluated (N=7). These techniques are highly 21 recognizable by patients for a popping sound emitted by the column during their 22 23 performance.(45)The fact that participants enrolled in these trials were eligible despite having 24 already received SM, threatens the validity of blinding. This thought is strengthened by the high 25 percentage of participants who recognized treatment allocation in this kind of trial (from 63.5% 26 to 83.5%). (23, 27) Additionally, five trials applied sham treatment in a different site compared 27 to pain and active treatment. This might have had important influences on sham therapy 28 29 reliability and consequently to study results.

30 Reliability of sham therapy seemed not to be related to dropouts rate, although both these 31 data were reported only in two trials. Bialosky J and Hoiriis K showed high percentages of 32 correct treatment allocation detection by participants but dropout rate between sham and MT 33 34 group did not differ. These results seem to be in conflict, nevertheless, participants could have 35 wanted to remain in the trial for several other reasons such as settings or being evaluated by an 36 expert clinician free. This possibility is reinforced by the fact that a similar dropout rate was 37 reached in the comparison sham versus control. These data suggest that dropout rate might 38 not be a dependable outcome for assessing reliability of sham therapy. The majority of trials 39 40 judged as at high or unclear risk of performance bias used a single technique evaluating its 41 effects on pain soon after its performance, or its effect after different sessions. Single 42 techniques were generally more difficult to mask, negatively affecting the validity of blinding of 43 participants. Moreover, it should be asked what result can be achieved with the application of a 44 single technique in a single therapeutic session and if the possible changes detected could be 45 46 clinically meaningful and long-lasting for the enrolled patients. 47

48 This review included generally small trials. Only 14 of 24 studies performed a sample size 49 calculation but just two of these considered MCID in this computation. The MCID is the 50 51 measure of smallest change of PROs that patients perceive as important, beneficial or harmful. 52 MCID is useful for clinicians to interpret the findings of trials and apply them in clinical practice 53 and to their decision-making. (46) An adequate sample size calculation, using MCID especially in 54 trials with PROs, is fundamental to assess the number of participants needed to detect clinically 55 relevant treatment effects. Oversized trials, which expose too many people to unnecessary 56 57 therapies, or underpowered trials, which may not achieve significant results, should be avoided. 58 (47-49) 59

60

1 2 3

4 5 6

7

8

9

10

Our results are similar to other reviews findings, notwithstanding that these reviews did not consider the difference between kinds of ST provided (hand contact or machines) and evaluated the effect of a singular type of MT (such as SM or OMT) compared to sham treatment.(50, 51)

#### Limitations

This review aimed to compare different kinds of sham therapy with different kinds of MT and control. 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. (52) Additionally, the use of machine placebo trials in the same meta-analysis could have increased diversity within trials included 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 metaanalysis 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 metaanalysis. 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.

Pair-wise 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. (53) 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 improved if both authors worked independently on bias risk assessment and then discussed any discrepancy.

#### Implications for practice and research

There is very low quality of evidence that sham compared to MT might be less effective and equally safe in the treatment of patients affected by back pain. Future studies should address

meaningful research question and improve their methodological properties to ensure patients safety and to guarantee reliability of study results

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.

Although in MT trials a true placebo is difficult to achieve, trials should also implement strategies to guarantee patients and assessors blinding, for example avoiding the inclusion of participants who already received the active treatment. Plans to avoid performance bias, such as giving similar treatment with similar localization have to be implemented.

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

Future researches should also evaluate the real effects of ST comparing it both with active treatment and to control groups. Only with this kind of design the real placebo effect in MT could be defined.

Trials should also focus on including participants with similar characteristics such as duration of symptoms (acute or chronic pain).

The majority of studies included in this review used a single technique treatment (N=11), however the clinical relevance of demonstrating the effectiveness of a singular technique is not clear. In a clinical context, most manual treatments usually involve different kinds of techniques in the same treatment session, so trials that evaluate the effectiveness of a type of treatment should include a routine of techniques in order to be more similar to clinical approaches. Studies should also consider using objective end points, not patient-reported or observerreported, with a longer period of follow-up. All these design implementations might not have a great impact on the demonstration of effectiveness of MT in BP, nevertheless, addressing meaningful researches questions, closer to the therapeutic context, could probably help to assess the real clinical effects of sham and MT.

# Conclusions

This review aimed to evaluate ST effect in MT trials. Although MT showed higher efficacy than ST, 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.

## **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.

**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.

**Data sharing:** Details of the characteristics of the included studies and data extracted are available from the corresponding author at <u>carolina.lavazza@docenti.aimoedu.it</u>.

References

Meissner K, Fässler M, Rücker G, Kleijnen J, Hróbjartsson A, Schneider A, et al.

Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. .

#### 

2. Sciences CfIOoM. International Ethical Guidelines for Health-Related Research Involving Humans, Fourth Edition. Geneva2016.

JAMA internal medicine. 2013;173(21):1941-51.

1.

3. 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).

4. Korthals-de Bos IB, Müllner M, Hoving JL, van Tulder MW, Rutten-van Mölken MP, Adèr HJ, 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(7395):1-6.

5. Kawchuk G N, Haugen R, Fritz J. A True Blind for Subjects Who Receive Spinal Manipulation Therapy. Archives of Physical Medicine and Rehabilitation. 2009;90(2):366-8.

6. Mancini F, Beaumont AL, Hu L, Haggard P, Iannetti GDD. Touch inhibits subcortical and cortical nociceptive responses. Pain. 2015;*156*(10).

7. So PS, Jiang JY, Qin Y. Touch therapies for pain relief in adults. . Cochrane Database of Systematic Reviews. 2008.

8. Cai L., He L. Placebo effects and the molecular biological components involved. General psychiatry. 2019;32(5):1-10.

Kaptchuk T. J., Miller F. G. Placebo effects in medicine. N Engl J Med. 2015;373(1):8-9.
 Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Placebo effects: biological, clinical and ethical advances. Lancet. 2010;375:686.

11. 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 C. B., Espie C. A., Epstein D. R., et al. The evidence base of sleep restriction therapy for treating insomnia disorder. . Sleep medicine reviews. 2014;18(15):415-24.
 Colloca L, Miller FG. The nocebo effect and its relevance for clinical practice.

Psychosomatic medicine. 2011;73:598-603.

14. Koes BW. How to evaluate manual therapy: value and pitfalls of randomized clinical trials. Manual therapy. 2004:183-4.

15. Chaibi A, Benth JŠ, Russell MB. Validation of placebo in a manual therapy randomized controlled trial. Scientific reports. 2015;5(1):1-8.

16. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews. 2015;4(1):1.

17. Pool JJ, Ostelo RW, Hoving JL, Bouter LM, De Vet HC. Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. . Spine. 2007;32(26):3047-51.

18. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine. 2008;33(1):90-4.

19. Farrar JT, Young Jr JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain. 2001;94(2):149-58.

3	20. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions
4	
5	Version 5.1.0 [updated March 2011].9.5.2 Identifying and measuring heterogeneity.
6	https://handbook-5-
7	<u>1.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm</u> : The
8	Cochrane Collaboration; 2011.
9	21. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of
10	Interventions Version 5.1.0 [updated March 2011] https://handbook-5-
11	1.cochrane.org/index.htm#chapter 8/8 assessing risk of bias in included studies.htm: The
12 13	
13	Cochrane Collaboration; 2011.
15	22. Licciardone JC, Minotti DE, Gatchel RJ, Kearns CM, Singh K. Osteopathic manual
16	treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. The
17	Annals of Family Medicine. 2013;11(2):122-9.
18	23. Bialosky JE, George SZ, Horn ME, Price DD, Staud R, Robinson ME. Spinal manipulative
19	therapy–specific changes in pain sensitivity in individuals with low back pain (NCT01168999)
20	The journal of pain. 2014;15(2):136-48.
21	24. Eardley S, Brien S, Little P, Prescott P, Lewith G. Professional kinesiology practice for
22	chronic low back pain: single-blind, randomised controlled pilot study Complementary
23	
24 25	Medicine Research. 2013;20(3):180-8.
25	25. Erdogmus CB, Resch KL, Sabitzer R, Müller H, Nuhr M, Schöggl A, et al. Physiotherapy-
27	based rehabilitation following disc herniation operation: results of a randomized clinical trial
28	Spine. 2007;32(19):2041-9.
29	26. Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, et al. Intensive, dynamic
30	back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back
31	pain. A randomized, observer-blind trial. Spine. 1993;18(1):98-108.
32	27. Hoiriis KT, Pfleger B, McDuffie FC, Cotsonis G, Elsangak O, Hinson R, et al. A randomized
33	clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain.
34 25	
35 36	Journal of Manipulative and Physiological Therapeutics. 2004;27(6):388-98.
30 37	28. Krekoukias G, Gelalis ID, Xenakis T, Gioftsos G, Dimitriadis Z, Sakellari V. Spinal
38	mobilization vs conventional physiotherapy in the management of chronic low back pain due to
39	spinal disk degeneration: a randomized controlled trial Journal of Manual & Manipulative
40	Therapy. 2017;25(2):66-73.
41	29. Licciardone JC, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W, et al. Osteopathic
42	manipulative treatment for chronic low back pain: A randomized controlled trial. Spine.
43	2003;28(13):1355-62.
44	30. Senna MK, Machaly SA. Does maintained spinal manipulation therapy for chronic
45	
46 47	nonspecific low back pain result in better long-term outcome? Spine. 2011;36(18).
47 48	31. Lascurain-Aguirrebeña I, Newham DJ, Casado-Zumeta X, Lertxundi A, Critchley DJ.
49	Immediate effects of cervical mobilisations on global perceived effect, movement associated
50	pain and neck kinematics in patients with non-specific neck pain. A double blind placebo
51	randomised controlled trial. Musculoskeletal Science and Practice. 2018;38:83-90.
52	32. Selkow NM, Grindstaff TL, Cross KM, Pugh K, Hertel J, Saliba S. Short-term effect of
53	muscle energy technique on pain in individuals with non-specific lumbopelvic pain: a pilot
54	study Journal of manual & manipulative therapy. 2009;19(1):14E-8E.
55	33. Antolinos-Campillo PJ, Oliva-Pascual-Vaca A, Rodríguez-Blanco C, Heredia-Rizo AM, Espí-
56	
57 58	López GV, Ricard F. Short-term changes in median nerve neural tension after a suboccipital
58 59	muscle inhibition technique in subjects with cervical whiplash: a randomised controlled trial.
60	Physiotherapy. 2014;100(3):249-55.
	28

**BMJ** Open

34. Haller H, Lauche R, Cramer H, Rampp T, Saha FJ, Ostermann T, et al. Craniosacral therapy for the treatment of chronic neck pain: a randomized sham-controlled trial. The Clinical journal of pain. 2016;32(5):441-9.

35. Klein R, Bareis A, Schneider A, Linde K. Strain–counterstrain to treat restrictions of the mobility of the cervical spine in patients with neck pain—A sham-controlled randomized trial. Complementary Therapies in Medicine. 2013;21(1):1–7.

36. Younes M, Nowakowski K, Didier-Laurent B, Gombert M, Cottin F. Effect of spinal manipulative treatment on cardiovascular autonomic control in patients with acute low back pain. Chiropractic and Manual Therapies. 2017;25(1).

37. Quinn F, Hughes CM, Baxter GD. Reflexology in the management of low back pain: A pilot randomised controlled trial. Complementary Therapies in Medicine. 2008;16(1):3-8.
38. Sillevis R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. Journal of Manual & Manipulative Therapy ,. 2010;18.

39. Cleland JA, Childs MJD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. *Manual therapy*. 2005;10(2):127-35.

40. Pires PF, Packer AC, Dibai-Filho AV, Rodrigues-Bigaton D. 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. Journal of manipulative and physiological therapeutics. 2015;38(8).

41. Silva AC, Biasotto-Gonzalez DA, Oliveira FHM, Andrade AO, Gomes CADFP, Lanza FDC, 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. . Evidence-Based Complementary and Alternative Medicine. 2018;2018:1-9.

42. Vieira-Pellenz F, Oliva-Pascual-Vaca Á, Rodriguez-Blanco C, Heredia-Rizo AM, Ricard F, Almazán-Campos G. 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. Archives of physical medicine and rehabilitation. 2014;95(9):1613-9.

43. Kogure A, Kotani K, Katada S, Takagi H, Kamikozuru M, Isaji T, 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(12):1-16.

44. Hall T, Hardt S, Schafer A, Wallin L. Mulligan bent leg raise technique-a preliminary randomized trial of immediate effects after a single intervention. Manual Therapy. 2006;11(2):130-5.

45. Dunning J, Mourad F, Barbero M, Leoni D, Cescon C, Butts R. Bilateral and multiple cavitation sounds during upper cervical thrust manipulation. BMC musculoskeletal disorders. 2013;14(1):24.

46. BC J, S E, A C-L, Furukawa TA, Patrick DL, Crawford MW et al. Minimally important difference estimates and methods: a protocol. BMJ Open. 2015;5:1-5.

47. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials. BMJ. 2009;388.

48. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. . International journal of surgery. 2012;10(1):28-55.

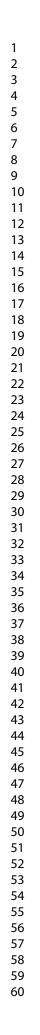
49. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PloS one. 2013;8(3).

50. Ruddock JK, Sallis H, Ness A, Perry RE. Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. . Journal of chiropractic medicine. 2016;15(3):165-83.

51. Rubinstein SM, De Zoete A, Van Middelkoop M, Assendelft WJ, De Boer MR, Van Tulder MW. 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.

52. Monroe CM. The effects of therapeutic touch on pain. Journal of Holistic Nursing. 2009;27(2):85-92.

Rubinstein SM, van Eekelen R, Oosterhuis T, de Boer MR, Ostelo RW, van Tulder MW. 53. The risk of bias and sample size of trials of spinal manipulative therapy for low back and neck pain: analysis and recommendations. Journal of manipulative and physiological therapeutics. 2014;37(8):523-41. for peer teries only



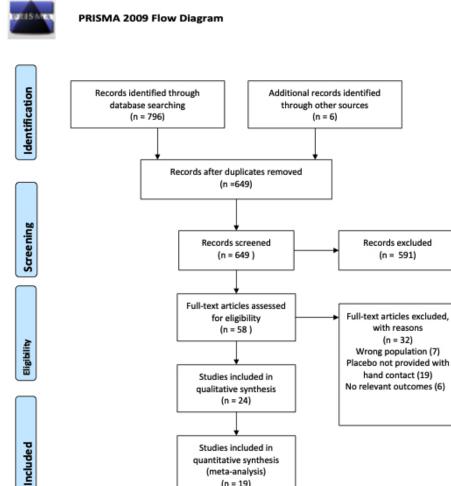


Figure 1: PRISMA flow diagram

(meta-analysis)

(n = 19)

48x56mm (298 x 300 DPI)

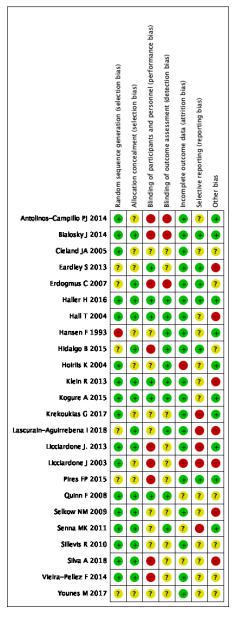


Figure 2: Risk of bias summary. Review authors' judgements about each risk of bias item for each included study.

34x93mm (298 x 299 DPI)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		sт			мт			Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean		Total	Mean		Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% CI	
4.1.1 Short-term								,		
Antolinos-Campillo PJ 2014	55.2	10.5	20	48.7	10.8	20	0.6%	6.50 [-0.10, 13.10]		970097
Bialosky j 2014		22.45		34.03		27		-2.28 [-15.20, 10.64]		
Cleland JA 2005	43.5		17			19	0.2%	17.40 [5.33, 29.47]		422242
Eardley S 2013	36	19.8	21		22.7	20	0.2%	7.00 [-6.06, 20.06]	+	224244
Erdogmus C 2007	21.6	3.13		24.13	2.5	40		Not estimable		200000
Haller H 2016	53.5	20.3	27	31.7	20.7	27	0.3%			
Hoiriis K 2004	22.1	20.2	40			34	0.4%	5.00 [-3.89, 13.89]	<u>+-</u>	
Klein R 2013	38	26	31	32	22	30	0.2%	6.00 [-6.07, 18.07]	<u> </u>	
Kogure A 2015	47.6		89	43.4	20	90	0.9%	4.20 [-1.82, 10.22]	-	
Krekouklas G 2017	58.8	9.2	25	12.2	11	25	0.074	Not estimable		
Licclardone J 2003	30.7	21.9	23	37.7		42	0.2%	-7.00 [-18.95, 4.95]		
Pires FP 2015	25.7	22.7	16	22		16	0.1%	3.70 [-11.69, 19.09]		22020
Selkow NM 2009	35.2	28	10	25	20.6	10		10.20 [-11.34, 31.74]		
Senna MK 2011	33.19			29.46	1.16	26	95.3%	3.73 [3.14, 4.32]		
Silva A 2018	47.1	17.2	14			14	0.2%	15.00 [1.18, 28.82]		
Vieira-Pellez F 2014		26.37			22.47	20	0.1%	9.09 [-6.09, 24.27]		
Younes M 2017	18.14		- 20	12.6		10	0.7%	5.54 [-1.21, 12.29]		222202
Subtotal (95% CI)	10.14	7.40	400	12.0	9.24		100.0%	3.86 [3.29, 4.43]		
Heterogeneity: $Chl^2 = 23.95$ , Test for overall effect: $Z = 13$				- 42%						
Total (95% CI)			400			405	100.0%	3.86 [3.29, 4.43]	r -	
Heterogeneity: Chi <sup>2</sup> = 23.95,	df = 14	(P = 0.6)	05): P	42%				-		_
Test for overall effect: Z = 1									-50 -25 0 25 50	
Test for subgroup difference.									Favours ST Favours MT	
Risk of bias legend										
(A) Random sequence generation	ation (sel	ection bi	as)							
(B) Allocation concealment (se			,							
(C) Blinding of participants an			formar	ce bias						
(D) Blinding of outcome asses										
(E) Incomplete outcome data			/							
(F) Selective reporting (report										
(G) Other bias	5									

Figure 3: Forest plot of comparison ST versus MT in back pain outcome at short-term.

84x50mm (300 x 298 DPI)

	ST		MT			Risk Ratio	Risk Ratio	<b>Risk of Bias</b>
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI	M-H, Fixed, 95% CI	ABCDEFO
Blakosky J 2014	1	28	1	27	1.0%	0.96 [0.06, 14.65]		9999999
Eardley S 2013	1	21	0	20	0.5%	2.86 [0.12, 66.44]		??
Erdogmus C 2007	8	40	5	40	4.6%	1.60 [0.57, 4.47]	<b></b>	? • • • • • ?
Haller H 2016	13	27	11	27	10.5%	1.18 [0.65, 2.15]	- <b>-</b>	9999999
Hansen F 1993	5	59	11	57	10.7%	0.44 [0.16, 1.18]		•??••?•
Hoirlis K 2004	13	53	13	50	12.6%	0.94 [0.49, 1.83]		9??99?
Kogure A 2015	4	69	3	90	2.8%	1.35 [0.31, 5.85]		<b></b>
Licclardone J 2003	4	23	16	46	9.9%	0.52 [0.20, 1.39]		• ? • ? • •
Licclardone J. 2013	34	225	39	230	36.8%	0.89 [0.58, 1.36]		
Senna MK 2011	20	40	7	27	6.0%	1.93 [0.95, 3.92]		
Younes M 2017	2	7	3	10	2.4%	0.95 [0.21, 4.29]		2222927
Total (95% CI)		612		626	100.0%	0.98 [0.77, 1.25]	•	
Total events	105		109					
Heterogeneity: Chi <sup>2</sup> =	9.69, df	= 10 (F	P = (0.47)	i; i² = 0	<b>%</b>		0.01 0.1 1 10 100	
Test for overall effect:	z = 0.12	(P = 0	.90)				Favours ST Favours MT	
Risk of bias legend								
(A) Random sequence	e generatio	on (sele	ction bia	s)				
(B) Allocation concealr	ment (sele	ction bi	as)					
(C) Blinding of particip	pants and	person	nel (perf	ormanc	e bias)			
(D) Blinding of outcom				()				

(G) Other bias

Figure 4: Forest plot of comparison ST versus MT in number of dropouts outcome

72x38mm (299 x 299 DPI)

1	
2	
3	
4	
5	
6	ST MT Risk Ratio Risk Ratio Risk of Bias
7	Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI A B C D E F G Eardley \$ 2013 2 21 1 20 2.6% 1.90 [0.19, 19.40] 7 9 9 9 9
8	Erdogmus C 2007 11 40 12 40 31.0% 0.92 [0.46, 1.83] 🚽 🧻 💎 🖶 🖶 😨 😨 🖓
9	Haller H 2016 8 27 6 27 15.5% 1.33 [0.53, 3.33]
10	Klein R 2013 1 31 4 30 10.5% 0.24 [0.03, 2.04] 🛛 🚽 🕒 🕒 🕀 🕀 🕀 🕐 🕐
11	Kogure A 2015 10 89 8 90 20.6% 1.26 [0.52, 3.05]
12	Total (95% CI) 267 264 100.0% 0.84 [0.55, 1.28]
13	Total events       32       38         Heterogenety: $6.79$ , df = 5 (P = 0.24); $t^2$ = 26%       0.001       0.1       1       10       1000
	Test for overall effect: Z = 0.81 (P = 0.42)         0.001         0.1         1         10         1000'           Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT
14	Risk of bias legend
15	(A) Random sequence generation (selection bias)
16	(B) Allocation concealment (selection bias) (C) Blinding of participants and personnel (performance bias)
17	(D) Blinding of outcome assessment (detection bias)
18	(E) Incomplete outcome data (attrition bias) (F) Selective reporting (reporting bias)
19	(G) Other bias
20	
	Figure 5: Forest plot of comparison ST versus MT in number of adverse events outcome at short-term
21	righte 5. To lest plot of comparison 51 versus MT in number of adverse events outcome at short-term
22	72x31mm (299 x 298 DPI)
23	72X3111111 (233 X 230 D11)
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
49	

Pag
1 2 3 4
5
6 7
8
9
10 11
12
13
14
15 16
17
18
19
20
21 22
22 23
23 24
25
26
27
28 29
29 30
31
32
33
34
35 36
36 37
38
39
40

	ST		Cont	rol		Risk Ratio	Risk Ratio	Risk of Bias
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI	M-H, Fixed, 95% CI	ABCDEF
Blakosky J 2014	1	28	0	28	1.4%	3.00 [0.13, 70.64]		
Eardley S 2013	1	21	4	25	10.6%	0.30 [0.04, 2.46]		<b>??????</b>
Erdogmus C 2007	6	40	8	40	23.2%	1.00 [0.42, 2.40]	-+-	?
Hoirlis K 2004	13	53	17	53	49.3%	0.76 [0.41, 1.41]		977997
Licclardone J 2003	4	23	5	20	15.5%	0.70 [0.22, 2.24]		••••
Total (95% CI)		165		166	100.0%	0.79 [0.51, 1.23]	•	
Total events	27		34					
Heterogeneity: Chi <sup>2</sup> =	1.84, df	= 4 (P	= 0.77);	$l^2 = 03$			has als de sad	
Heterogeneity: Chi <sup>2</sup> = Test for overall effect:				12 = 03	í		0.01 0.1 1 10 100	
				r² = 0%	í		0.01 0.1 1 10 100 Favours ST Favours Control	
				r² = 0%	6			
Test for overall effect:	z = 1.03	3 (P = 0	.30)		í			
Test for overall effect: Risk of bias legend	z = 1.03	on (sele	ction bia		1			
Test for overall effect: <u>Risk of bias legend</u> (A) Random sequence	generation ment (sele	on (sele	ction bia	s)				1
Test for overall effect: <u>Risk of bias legend</u> (A) Random sequence (B) Allocation concealr	generation ment (sele	on (sele ction b person	ction bia ias) nel (perfi	s) ormanc				1
Test for overall effect: <u>Risk of bias legend</u> (A) Random sequence (B) Allocation concealr (C) Blinding of particip	generation ment (sele pants and ne assess	on (sele ection b person ment (d	ction bia ias) nel (perfe	s) ormanc				1
Test for overall effect: <u>Risk of bias legend</u> (A) Random sequence (B) Allocation concealr (C) Blinding of particip (D) Blinding of outcom	e generation ment (sele bants and he assessme he data (at	on (sele ction b person ment (d ttrition	ction bia ias) nel (perfe	s) ormanc				1

Figure 6: Forest plot of comparison ST versus control in number of dropouts outcome

61x25mm (300 x 300 DPI)

# Appendix 1: search strategy

## Medline

1 2 3

4 5 6

7 8

9

10 11

12

13

14

15

16 17

18

19

21 22

23

24

25

26

27

29

30

31

32

33 34

37

39

43

45

46

47

48

49 50

51

52

53

54

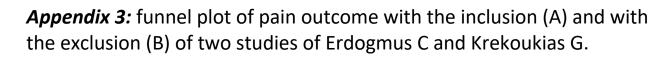
- Mesh descriptor: [Back Pain] explode all trees
- 2. dorsalgia/
- 3. backache
  - 4.(neck OR cervical) adj1 pain  $\rightarrow$  Mesh
  - 5. exp Brachial Plexus Neuropaties
- 6. exp Lumbar Plexus Neuropaties
- 7. Neck Pain/
- 8. neckache
- 9. Torticollis/
- 10. whiplash.mp 20
  - 11. cervicodynia.mp
    - 12. spondylitis/ OR spondylosis/ OR spondylolysis/ OR spondylolysthesis
  - 13.(lumbar OR dorsal OR neck OR cervical OR sciatica) adj2 (pain OR ache)
  - 14. (lumbar OR dorsal OR neck OR cervical) adj2 (discitis OR disc adj 1 herniation OR disc adj1 herniation)
  - 15. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
- 28 16. (PLACEBO\* or MOCK\* or SHAM\* or FAKE\* or VEHICLE\* or DUMM\* or ATTENTION\* CONTROL\*
  - or PSEUDO\* TREAT\* or UN?SPECIFIC\* or NON?SPECIFIC\* OR simulat\$ treatment OR inert agent) 17. Chiropractic/
  - 18. Manipulation, Chiropractic/
  - 19. chiropract\$.tw.
  - 20. (manual adj2 therap\$).mp
- 35 21. spinal manipulation.mp. or Manipulation, Spinal/ 36
  - 22. osteopath\$.tw.
- 23. Osteopathic Medicine/ 38
  - 24. Physical Therapy Modalities/ or "Physical Therapy (Specialty)"/ or physical therap\$.tw. or
- 40 physiotherap\$.tw. 41
- 25. myotherapy.mp 42
  - 26. shiatsu.mp
- 44 27.exp Therapeutic Touch/
  - 28. exp Massage/
  - 29. (neuromuscular adj therapy).mp
  - 30. 17 OR 18 OR 19 OR 20 OR OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29
    - 31. pain
  - 32. range of motion
  - 33. ROM
  - 34. 31 OR 32 OR 33
  - 35. Clinical Trial/
- 55 36. Randomized Controlled Trial/
- 56 37. controlled clinical trial/ 57
  - 38. exp RANDOMIZATION/
  - 39. PLACEBO/
- 59 40. (random\$ adj2 allocat\$).tw. 60
  - 41. single blind\$.tw.

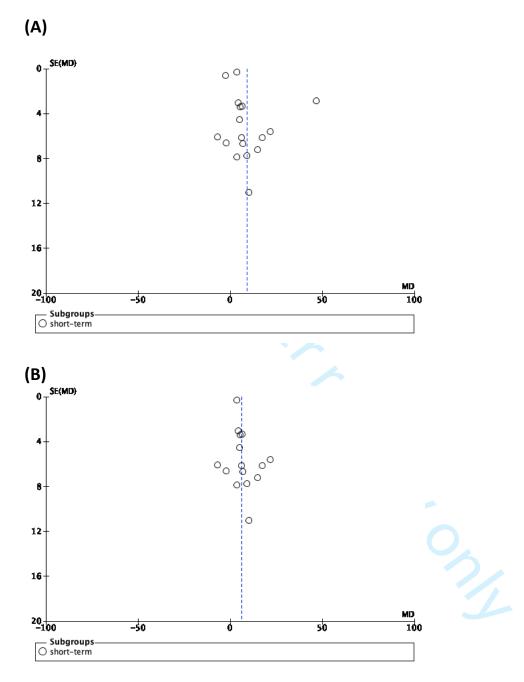
2	
3	42. double blind\$.tw.
4 5	43. placebo\$.tw.
5 6	44. 35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43
7	45. animals/
8	46. humans/
9	47. 46 NOT 45
10	48. 15 AND 16 AND 30 AND 44 AND 47
11	48. 15 AND 16 AND 50 AND 44 AND 47
12	
13 14	
15	
16	
17	
18	
19	
20 21	
21 22	
22	
24	
25	
26	
27 28	
28 29	
30	
31	
32	
33	
34 35	
35 36	
37	
38	
39	
40	
41 42	
42 43	
44	
45	
46	
47	
48 49	
49 50	
51	
52	
53	
54	
55 56	
56 57	
58	
59	
60	

# Appendix 2: data extraction form

Methods	Trial Design
	Settings
Participants	Total number of participants:
	Age:
	Gender(M/F):
O,	BMI:
	Activity:
	Duration of the symptoms:
	Location of pain (one-sided, double-sided, central, cervical, dorsal or lumbar):
	<b>Cause of pain:</b> (e.g. disc herniation, contractures, aspecific pain)
	Previous experience of the treatment provided: Y/N/ N/A
	Inclusion and exclusion Criteria:
	VAS:
	Practitioner characteristics: (years of experience, gender)
Interventions	Placebo:
	Comparator:
Outcomes	Outcomes used in the meta-analysis:
	Length of follow-up:
Notes	Difference between Placebo and active treatment:
	Placebo check for reliability:
	Adverse event:

Lost to follow-up:
 Funding source:





**Appendix 4:** forest plot of comparison pain outcome sham treatment vs manual therapies with the inclusion of two trials (Erdogmus C and Krekoukias G) at short, medium and long-term.

		ST			мт			Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean		Total	Mean		Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	
3.1.1 Short-term								,	,	
Antolinos-Campillo PJ 201		10.5				20	4.6%	6.50 [-0.10, 13.10]	-	•?•••
Białosky J 2014		22.45	-	34.03 26.1		27 19		-2.28 [-15.20, 10.64]	—	
Cleland JA 2005 Eardley S 2013	43.5 36	19.5 19.6	17 21		22.7	20	2.4%	17.40 [5.33, 29.47] 7.00 [-6.06, 20.06]		
Erdogmus C 2007	21.6	3.13		-	2.5	40	7.1%	-2.53 [-3.77, -1.29]	-	200000
Haller H 2016	53.5	20.3			20.7	27	2.6%	21.80 [10.86, 32.74]		9999999
Hoirils K 2004	22.1	20.2	-			34	3.5%	5.00 [-3.89, 13.89]	+	• ? ? • • ? •
Klein R 2013 Kogure A 2015	38 47.6	26 21.1	31 69	32 43.4	22 20	30 90	2.4% 4.9%	6.00 [-6.07, 18.07] 4.20 [-1.82, 10.22]		444442
Krekouklas G 2017	58.8	9.2		-	-	25	5.1%		-	• ? ? ? • •
Licclardone J 2003	30.7	21.9	23	-		42	2.5%	-7.00 [-18.95, 4.95]	+	• ? • ? • •
Pires FP 2015	25.7	22.7	16	22		16	1.7%		<u> </u>	<b>? ? • ? • •</b>
Selkow NM 2009 Senna MK 2011	35.2 33.19	28 1.19		25 29.46		10 26	7.2%	10.20 [-11.34, 31.74] 3.73 [3.14, 4.32]		
Silva A 2018	47.1	17.2			20.01	14	2.0%	15.00 [1.18, 28.82]		
Vieira-Pellez F 2014	29.1	26.37	20	20.01	22.47	20	1.6%	9.09 [-6.09, 24.27]	+	
Younes M 2017	18.14	7.46			6.24	10	4.5%	5.54 [-1.21, 12.29]	•	22224
Subtotal (95% CI) Heterogeneity: Tau <sup>2</sup> = 62	40· Chi -	330 75	465		00001	470		8.96 [4.45, 13.47]	•	
Test for overall effect: Z =						<i>.</i> ,,, = ;				
3.1.2 Medium-term										
Erdogmus C 2007 Haller H 2016	18.35 47.8	3.03 19.3		17.11 31.6	5.48 19	40 27	6.9X 3.0X	1.24 [-0.70, 3.18] 16.20 [5.98, 26.42]	t	?
Kogure A 2015	47.8	22.2		36		90	5.0%	10.30 [4.45, 16.15]		44444
Senna MK 2011	35.16		37	35.16		26	7.2%	0.00 [-0.64, 0.64]		••?•?
Subtotal (95% CI)			193			183	22.0%	3.93 [0.29, 7.57]	•	
Heterogeneity: Tau <sup>2</sup> = 9.3 Test for overall effect: Z =			f = 3 (P	< 0.00	01); r -	67%				
	• • • • • • • • •	0.03)								
3.1.3 Long-term										
Erdogmus C 2007	16.85	3.71		21.05		40 26	7.0%	-4.20 [-5.80, -2.60]	•	
Kogure A 2015 Licclardone J 2003	45.5 24.5	22 21.1		31.2 31.6		32	3.6× 2.4×	14.30 [5.75, 22.85] -7.10 [-19.36, 5.16]		
Senna MK 2011	36.8		-	35.53		26	7.1%	1.27 [0.34, 2.20]	-	
Subtotal (95% CI)			185			124		0.70 [-4.32, 5.71]	•	
Heterogeneity: Tau <sup>2</sup> = 18 Test for overall effect: Z =			df = 3 (	(P < 0.0	0001);	r = 93)	×			
rest for overall effect. Z =	• U.27 (P = 1	y.79)								
Total (95% CI)			843			777	100.0%	6.35 [4.02, 8.67]	•	
Heterogeneity: $Tau^2 = 19$				:4 (P < (	).00001	.);	95%	-	-50-25 0 25 50	_
Test for overall effect: Z = Test for subgroup differen				- 0.00	1 F - 4				Favours ST Favours MT	
Risk of bias legend		4.04, 0	(1	- 0.03		N. 37				
(A) Random sequence ge	neration (sel	ection b	ias)							
(B) Allocation concealmen										
(C) Blinding of participant				nce bias)						
<ul> <li>(D) Blinding of outcome as</li> <li>(E) Incomplete outcome data</li> </ul>			i dias)							
(F) Selective reporting (rep	<b>J</b>									
(G) Other bias										

# **Appendix 4.1:** Sensitivity analysis with the exclusion of Ergogmus C and Krekoukias G at short, medium and long-term

2											
3			ST			мт			Mean Difference	Mean Difference	<b>Risk of Bias</b>
4	Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
	3.1.1 Short-term			~~			~~				
5	Antolinos-Campillo PJ 2014 Bialosky J 2014		10.5 22.45	20 28	48.7 34.03	10.8 26.21	20 27	4.6%	6.50 [-0.10, 13.10] -2.28 [-15.20, 10.64]		
6	Cleland JA 2005	43.5	-	17	26.1	17.2	19	2.0%	17.40 [5.33, 29.47]		<b>•</b> ? ? ? <b>•</b> ? ?
7	Eardley S 2013	36	19.8	21	29	22.7	20	1.7%	7.00 [-6.06, 20.06]		224244
=	Erdogmus C 2007	21.6	3.13	40	24.13	2.5	40	0.0%	-2.53 [-3.77, -1.29]		? • • • • • ?
8	Haller H 2016	53.5		27	31.7	20.7	27	2.3%	21.80 [10.86, 32.74]		<b>6666666</b>
9	Hoirils K 2004	22.1	-	40	17.1	18.6	34	3.2%	5.00 [-3.89, 13.89]	+	• ? • • • ? •
10	Klein R 2013	38	26	31	32	22	30	2.0%	6.00 [-6.07, 18.07]	<u> </u>	
	Kogure A 2015 Krekouklas G 2017	47.6 58.8	21.1 9.2	89 25	43.4	20 11	90 25	5.1X 0.0X	4.20 [-1.82, 10.22] 46.60 [40.98, 52.22]	1-	
11	Licclardone   2003	30.7	-	23	37.7	26.2	42		-7.00 [-18.95, 4.95]		
12	Pires FP 2015	25.7	-	16	22	21.7	16	1.3×	3.70 [-11.69, 19.09]		220200
13	Selkow NM 2009	35.2	28	10	25	20.6	10		10.20 [-11.34, 31.74]		•••••
	Senna MK 2011	33.19	1.19	37	29.46	1.16	26	11.0%	3.73 [3.14, 4.32]	-	••?•?•
14	Silva A 2018	47.1		14		20.01	14	1.6%	15.00 [1.18, 28.82]	<u> </u>	•••???•
15	Vieira-Pellez F 2014		26.37	20		22.47	20	1.3%	9.09 [-6.09, 24.27]	+	•••?
	Younes M 2017 Subtatal (05% CI)	18.14	7.46	7 400	12.6	6.24	10 405	4.5% 45.0%	5.54 [-1.21, 12.29]	<u> </u>	??????
16	Subtotal (95% CI)	CLB - 2	3 O E -J		(n _ 0 0	E). 12 _		45.0%	6.09 [3.23, 8.96]	•	
17	Heterogeneity: Tau <sup>2</sup> = 9.37; Test for overall effect: Z = 4				r = 0.0	on r =	427				
18			,								
19	3.1.2 Medium-term										
	Erdogmus C 2007	18.35			17.11	5.48	40	9.9%	1.24 [-0.70, 3.18]	t	? • • • • • • ?
20	Haller H 2016 Kogure A 2015	47.8 46.3		27 89	31.6 36	19 17.4	27 90	2.6X 5.3X	16.20 [5.98, 26.42]		
21	Senna MK 2011	35.16			35.16	1.28	26	10.9%	10.30 [4.45, 16.15] 0.00 [-0.64, 0.64]	17	
	Subtotal (95% CI)	<b>JJ</b> .1 <b>Q</b>	1.20	193	55.14	1.20	183	28.7%	3.93 [0.29, 7.57]	•	
22	Heterogeneity: $Tau^2 = 9.14$	$Cht^2 = 2$	2.29, di	F = 3 (P	< 0.00	01); ř -	87%			·	
23	Test for overall effect: $Z = 2$	.12 (P =	0.03)	-							
24	3.1.3 Long-term										
25	Erdogmus C 2007	16.85	3.71	40	21.05	3.58	40	10.2%	-4.20 [-5.80, -2.60]		? • • • • • • ?
	Kogure A 2015	45.5	22	69	31.2	18.6	26	3.3%	14.30 [5.75, 22.85]		6666676
26	Licclardone J 2003	24.5		19	31.6	22.4	32		-7.10 [-19.36, 5.16]	-+	
27	Senna MK 2011	36.8	1.39		35.53	2.13	26	10.8%	1.27 [0.34, 2.20]		<b>••</b> • <b>?</b> • <b>?</b> ••
28	Subtotal (95% CI)			185			124		0.70 [-4.32, 5.71]	<b>•</b>	
	Heterogeneity: $Tau^2 = 18.03$			df = 3 (	(P < 0.0	0001);	i <sup>2</sup> = 93	×			
29	Test for overall effect: $Z = 0$	.27 (P = )	0.79)								
30	Total (95% CI)			778			712	100.0%	3.85 [1.98, 5.71]	•	
31	Heterogeneity: $Tau^2 = 6.19$	$Chl^2 = 1$	85.61, (	df = 22	(P < 0.	00001);	; i <sup>2</sup> = 6	6 <b>%</b>	-	-50 -25 0 25 50	_
	Test for overall effect: Z = 4					_				Favours ST Favours MT	
32	Test for subgroup difference	s: Chi <sup>2</sup> =	3.49, d	f = 2 (i	<b>?</b> = 0.17	'), I <sup>2</sup> = 4	2.7%				
33	Risk of bias legend	ntion (r-l	a ation b								
34	<ul> <li>(A) Random sequence gener</li> <li>(B) Allocation concealment (sequence)</li> </ul>			ias)							
	(C) Blinding of participants a			rformar	nce bias)						
35	(D) Blinding of outcome asse										
36	(E) Incomplete outcome data			,							
37	(F) Selective reporting (report	rting bias)									
57	(G) Other bias										

# Appendix 5: forest plot of comparison ST versus control in back pain outcome at short-term with exclusion of Erdogmus C 2007 trial

		SТ		c	Control			Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFO
Białosky j 2014	31.75	22.45	27	31.39	26.21	28	31.3%	0.36 [-12.52, 13.24]	-+-	
Eardley S 2013	36	19.8	21	57.7	19.2	21	33.2%	-21.70 [-33.50, -9.90]		??
Erdogmus C 2007	21.6	10.65	40	19.25	10.96	40	0.0%	2.35 [-2.39, 7.09]		? ?
Hoiriis K 2004	24.4	22.2	34	31.8	24	40	35.4%	-7.40 [-17.94, 3.14]		<b>9</b> ? <b>? 9 9</b> ? <b>9</b>
Total (95% CI)			82			89	100.0%	-9.72 [-21.90, 2.47]	•	
Heterogeneity: Tau <sup>2</sup> = Test for overall effect:				f = 2 (P	= 0.04)	;	9%		-100 -50 0 50 100 Favours ST Favours Control	

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias) (G) Other bias

# PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #	
TITLE				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Page 1	
ABSTRACT				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Page 5	
INTRODUCTION				
Rationale	3	Describe the rationale for the review in the context of what is already known.	Pages 9- 10	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Page 10	
2 METHODS				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Pages 6 and 10	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Pages 10-11	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Page 11	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix 1	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Page 12	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Page 12	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Page 12- 13	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Page 12	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	Pages 12-13	



# PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I <sup>2</sup> ) for each meta-analysis.	Page 13		
		Page 1 of 2	- I		
Section/topic	#	Checklist item	Reported on page #		
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Pages 12- 13		
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Page 13		
	RESULTS				
7 Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Page 15		
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Pages 16- 19		
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Page 20, figure 2		
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Pages 21- 23		
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Pages 7-8		
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Pages 20- figures 2,3,4,5,6		
<sup>3</sup> Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Appendices 4-6		
	•				
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Pages 7-8, 24		
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Page 25		
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Page 26		
		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml			

- 46 47

Page 47 of 46

BMJ Open



# PRISMA 2009 Checklist

3 4 5 6	Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Page 2
7	From: Moher D, Liberati A, Tetzla	ff J, Altm	nan DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS M	1ed 6(6): e1000097.
9	doi:10.1371/journal.pmed1000097		For more information, visit: www.prisma-statement.org.	
10				
11			Page 2 of 2	
12				
13				
14				
15				
16				
17				
18				
19 20				
20				
22				
23				
24				
25				
26				
27				
28			For more information, visit: www.prisma-statement.org. Page 2 of 2	
29 30				
30 31				
32				
33				
34				
35				
36				
37				
38				
39 40				
40 41				
41				
43				
44				
45			For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	
46				
47				

**BMJ** Open

# **BMJ Open**

#### Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pair-wise metaanalysis

Journal:	BMJ Open
Manuscript ID	bmjopen-2020-045106.R2
Article Type:	Original research
Date Submitted by the Author:	15-Mar-2021
Complete List of Authors:	Lavazza, Carolina; AIMO, Research Galli, Margherita; AIMO, Research Abenavoli, Alessandra; AIMO, Research Maggiani, Alberto; AIMO, Research
<b>Primary Subject Heading</b> :	Complementary medicine
Secondary Subject Heading:	Research methods
Keywords:	COMPLEMENTARY MEDICINE, STATISTICS & RESEARCH METHODS, Back pain < ORTHOPAEDIC & TRAUMA SURGERY





I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our <u>licence</u>.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which <u>Creative Commons</u> licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

reliez oni

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

# Title

#### Author's name

#### Address for each author

Separate each part of address with a comma (don't use a separate line for each part) Please include postcodes

#### Authors' names and positions

Use same style of name as below title Please give one job position for each author, (under the relevant address) on separate line from author's name

#### **Corresponding author**

Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pair-wise meta-analysis

Carolina Lavazza (0000-0002-9907-6819), Margherita Galli (0000-0003-4050-3131), A Abenavoli (0000-0001-8941-2308), A Maggiani (0000-0002-7980-9405) Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Carolina Lavazza associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, Margherita Galli associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, A Abenavoli associate professor, Department of Research AIMO institute, Piazzale Santuario 7, 21047 Saronno, Italy, A Maggiani professor

Correspondence to: carolina.lavazza@docenti.aimoedu.it

## Copyright/license for publication:

Carolina Lavazza The Corresponding Author of this article contained within the original manuscript which includes any diagrams & photographs within and any related or stand alone film submitted ("the Contribution") has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licencees, to permit this Contribution (if accepted) to be published in the BMJ and any other BMJ Group products and to exploit all subsidiary rights, as set out in our licence set out at:

http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyrightopen-access-and-permission-reuse.

Please tick one or more boxes as appropriate:

×

- I am the sole author of the Contribution.
- I am one author signing on behalf of all co-owners of the Contribution.
- П The Contribution has been made in the course of my employment and I am signing as authorised by my employer.
  - I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown\* acting in the course of my employment
- I am a US Federal Government employee acting in the course of my employment.
- I am not a US Federal Government employee, but some or all of my co-authors are.
- I am an employee of the UK Crown acting in the course of my employment
- I am not an employee of the UK Crown acting in the course of my employment but some/all of my coauthors are.\*

\*Such authors should consult guidance and if necessary return any completed form.

Licence: The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Conflict of interest: I have read and understood the BMJ Group policy on declaration of interests and declare the following interests: none

Ethical approval: not required.

Transparency: The manuscript's guarantor affirms that this 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.

Funding: None.

**Competing interests**: No funding were provided for this project and authors did not received support from any organisation for the submitted work. Authors have no financial relationships with any organisations that might have an interest in this project and have no other relationships or activities that could appear to have influenced this review.

in the second seco

1	
2	
3	
4	
5	Acronyma
6	Acronyms
7	
8	AE = adverse effects
9	BP = back pain
10	CI = confidence intervals
11	CRB = Cochrane risk of bias
12	
13	CT = clinical trial
14	MA = meta-analysis
15	MCID = minimal clinically important difference
16	MD = mean difference
17	MT = manual therapies
18 19	OMT = osteopathic manipulative treatment
20	
20	PROs = patient-reported outcomes
21	RCT= randomised controlled trial
23	RR = risk ratio
24	SM = spinal manipulation
25	ST = sham treatment
26	
27	
28	
29	
30	ST = sham treatment
31	
32	
33 34	
34 35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46 47	
47 48	
48	
49 50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	

# 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. Randomized controlled trials (RCTs) involving adults affected by BP (cervical and lumbar), acute or chronic, were included. Hand contact sham treatment (ST) was compared to 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 dropouts. Dichotomous outcomes were analysed using risk ratio (RR), continuous using mean difference (MD), 95% confidence intervals (CI). The minimal clinically important difference was 30 mm changes in pain score.

#### **Results:**

24 trials were included involving 2,019 participants. Very low evidence quality suggests clinically insignificant pain improvement in favour of MT compared to ST (MD 3.86, 95% CI 3.29 to 4.43) and no differences between ST and no treatment (MD -6.04, 95% CI -16.68 to 4.59).
ST reliability shows a high percentage of correct detection by participants (ranged from 46.7% to 83.5%), spinal manipulation being the most recognized technique.
Low quality of evidence suggests that AE and dropout rates were similar between ST and MT

(RR AE=0.84, 95% CI 0.55 to 1.28, RR dropouts= 0.98, 95% CI 0.77 to 1.25). A similar dropout rate was reported for no treatment (RR=0.79, 95% 0.51 to 1.23).

## **Conclusions:**

MT does not seem to have clinically relevant effect compared to ST. Similar effects were found with no treatment. The heterogeneousness 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 register: CRD42020198301

https://www.crd.york.ac.uk/prospero/display\_record.php?RecordID=198301

# Strengths and limitations of this study

#### Strengths

This systematic review and pair-wise meta-analysis:

- summarises existing evidence on the effect, reliability and application of hand contact ST in MT RCTs;
- gives suggestions for researchers on conducting methodical RCT in MT using a reliable sham procedure.

#### Limitation

- Settings and practitioner influences on ST effects were not analysed due to lack of data; •
- luences on S Luded was insuffic 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 (IMP)
 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)

29 In MT trials, placebo treatment is often provided in different modalities from trial to trial 30 although the manual techniques or treatments tested are the same. A true placebo does not 31 exist for MT and testing the effectiveness of MT requires a sham intervention. For instance, 32 33 sham treatment (ST) is commonly administrated as a light touch in the site of pain or as an 34 active treatment in a different site, (5) with no clear criterion. Such light touch might in fact 35 have a health effect and there is no evidence as to its ineffectiveness. Touch itself could have a 36 positive outcome on health (6) and active treatments could have an analgesic reflex on pain 37 38 even if administered elsewhere in the body. (7) 39

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 called nocebo. It has been estimated that up to 26% of patients in randomized control trials (RCTs) discontinue placebo due to adverse effects (AE). (9)

discontinue placeboldue to adverse effects (AL). (5)
 It is thought that these psychobiological phenomena may be related to the overall therapeutic
 context, such as treatment environment, individual patient and clinician factors (e.g. 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 a ST that mimics the active treatment and aims at blinding of participants.

1 2 3

8

9

10 11

12

13

14

15

40

41

42

43 44

45

22

23

24 25

26

27

28

29 30

31 32

33

34

35 36

37

43

44

45 46

47 48 49

50 51 52

53

54

55

56 57

58

59

60

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 to 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 sham treatment, with a better knowledge of the effect of manual contact on PROs such as pain relief and dropouts.

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 pair-wise 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 sham treatment 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 clinical trial development.

# Methods

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

The protocol registration was performed in PROSPERO (<u>http://www.crd.york.ac.uk/PROSPERO/</u>) and review registration number is **CRD42020198301.** 

#### Criteria for considering studies for this review

Only RCTs were included in this review. Quasi-randomised trials in which allocation was not strictly random (e.g. 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 (e.g. 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 10

provided by hand contact. Therefore, these studies were excluded. 11

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

ST was compared to other MT provided by any type of health care provider such as:

physiotherapist, chiropractor, osteopath, massage therapist, kinesiologist and reflexologist.

16 To assess if touch itself could have a positive health effect, ST was also compared to no treatment. Physiotherapeutic exercises were included in the analysis only if associated with 18 manual treatment.

The use of active co-interventions such as oral 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

1 2 3

4

5

6 7

8

9

12 13

14

15

17

19

20

21

22

27 28

29 30

31

32

33

34

39 40

41 42

43

44

45

46

47 48

49

50 51

52 53

54

55

56 57

58 59 60 Primary outcomes were pain intensity on a validated scale, success in the blinding of participants and AE. Secondary outcomes were number of dropouts.

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 longterm ( $\geq 6$  months). Data were extracted and analysed based on the time closest to these intervals.

#### Information sources

Search strategy (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, SPORTDiscus, PEDro, World Health Organization Clinical Trials Registration Platform, Index to Chiropractic Literature, Cochrane central register of controlled trials (CENTRAL), Clinical trials registry and metaRegister of Controlled Trials (mRCT).

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 gueries) was performed to evaluate the presence of on-going or recently completed systematic reviews. Guidelines from different organisations (e.g. 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 firstly 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 pre-defined form. Data extracted were related to settings, type of study, participants characteristics (such as localization 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. (*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 judgments was performed by *MG*. RevMan 5.3.5 was used for the graphic representation of each risk. The CRB tool results were then converted to 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% confidence intervals (CI).

Continuous outcomes, such as back pain on VAS scale, were evaluated using mean difference (MD) between ST and the MT/no treatment group with 95% CI and the standard deviation (SD). The minimal clinically important difference (MCID) between pre- 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<sup>2</sup> 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% to 100%. (23)

#### Synthesis of results

Meta-analysis of pain score, AE and dropout rates were performed using RevMan 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 localization 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 appendices.

#### Summarizing results and assessing the quality of the evidence

The quality of evidence for each outcome was evaluated with the 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://gradepro.org) was used to import data from RevMan 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 dropouts.

**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

Figure 1: PRISMA flow diagram

#### 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 United States, (24, 25, 31, 40, 41) three studies in Brazil, (42-44) one in the UK, (26) Egypt, (32) Japan (45) and Australia. (46)

No ongoing or unpublished trials were found.

#### Population

The included trials randomized a total of 2,019 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/m<sup>2</sup>.

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)

16 trials enrolled participants with low back pain (LBP), eight included participants with cervical pain (CP). (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 five (25, 26, 30) to 20 (27) 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, 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 a 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 mobilization and physiotherapy) the ST was administrated 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:

- Spinal manipulation (SM)/chiropractic (7 studies, 567 participants)
- Osteopathy (5 trials, 645 participants)
- Kinesiology (one trial, 58 participants)
- Articular mobilizations (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 manual therapy 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 six to 17 years. (30, 33, 35-37, 40, 42, 44) The gender of practitioners was indicated in only three trials.(26, 30, 37)

#### Table 1: summary of main characteristics of included studies

Study ID	N° of Symptoms participants duration		Pain localization	Technique tested (site of application)	Type of sham procedure	Other arms	Follow-up	
Antonilos- Campillo PJ 2014	pillo PJ reported		Cervical	Soft-tissue (cervical region)	Soft mobilization of lower limbs	None	No follow-up (outcomes collected after the intervention	
Bialosky J 2014	110	> 4 months	Lumbar	Spinal Manipulation (SM) <i>(lumbar spine)</i>	Ineffective force applied on lumbar spine	No treatment group	2 weeks	
Cleland JA 2005	36	>2 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the intervention	
Eardley S 2013	58	> 3 years	Lumbar	Kinesiology <i>(spine)</i>	Protocol of ineffective techniques in the site of pain	No treatment group	7 weeks	
Erdogmus S 2007	120	$\geq$ 1.4 weeks	Lumbar	Physiotherapy (Exercises of the spine and articular mobilisation techniques)	Physiotherapy Neck massage (Exercises of the spine and articular mobilisation		1.5 years	
Hall T 2004	24	Not reported	Lumbar	BLR technique (lower limbs)	Soft-tissue manipulation of the foot	None	24 hours	
Haller H 2016	54	> 7 months	Cervical	Cranio-sacral therapy (head)	Ineffective touch of head	None	3 months	
Hansen F 1993	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 B 2015	32	Not reported	Lumbar	Articular mobilization (lumbar spine)	Ineffective mobilization forces applied on lumbar spine	None	2 weeks	
Hoiriis K 2004	156	$\geq$ 2,3 weeks	Lumbar	SM (spine)	Ineffective force applied on spine	Medical treatment	4 weeks	
Klein R 2003	61	>1 month and <5 years	Cervical	Strain- counterstain techniques (cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the interventio	
Kogure A 2015	179	> 12 months and < 10 years	Lumbar	АКА-Н (sacro- iliac joint)	Ineffective force applied on sacro-iliac joint	None	6 months	
Krekoukias G 2017	50	Not reported	Lumbar	Articular mobilization techniques (lumbar spine)	Hand contact with lumbar skin placement without any movement	Exercise plus TENS	5 weeks	

Lascurain- Aguirrebena I 2018	40	Not reported	Cervical	Articular mobilization (Cervical spine)	Ineffective force applied on cervical spine	None	No follow-up (outcomes collected after the intervention
Licciardone J 2003	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 J 2013	455	$\geq$ 3 months	Lumbar	Osteopathic manual treatment (OMT) – (all body)	Protocol of light touch techniques similar to OMT applied to all body	None	8 weeks
Pires FP 2015	32	> 3 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	72 hours
Quinn F 2008	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 M 2009	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 seconds without any active contraction	None	24 hours
Senna MK 2011	93	$\geq$ 13 months	Lumbar	SM (lumbar spine)	Ineffective force applied on lumbar spine	Maintained SM	10 months
Sillevis R 2010	100	$\geq$ 23 months	Cervical	SM (thoracic spine)	Ineffective force applied on thoracic spine	None	No follow-up (outcomes collected after the intervention
Silva A 2019	28	>3 months	Cervical	Osteopathic visceral treatment (abdomen)	Hand contact on umbilical region without any movement	None	7 days
Veira-Pellez F 2014	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 M 2017	17	< 3 months	Lumbar	OMT (all spine)	Placebo mimed active treatment with an ineffective force applied.	None	7 days

#### 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.
- 17 The reasons for this judgment were mainly related to trials involving spinal manipulations.
- These studies used a technique which can be easily recognized 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 was 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.
  - Other bias occurred was generally considered as high risk for baseline differences of the population in 30% of trials.

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

#### Effects of intervention

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

tor occreation with

# **Table 2:** summary of findings of treatment effects and certainty of the evidence (GRADE) included for all comparisons.

#### 1. Sham treatment (ST) compared to Manual Therapies (MT)

Patient or population: back pain Intervention: ST

Comparison: MT

Outcomes		ed absolute (95% CI)	Relative effect	Nº of	Certainty of the evidence	Comments	
Outcomes	Risk with MT	Risk with ST	(95% CI)	participants (studies)	(GRADE)		
Pain improvement assessed with: VAS score Scale from: 0 to 100		MD <b>3.86</b> higher (3.29 lower to 4.43 higher)	5	805 (15 RCTs)	⊕⊖⊖⊖ VERY LOW <sup>a,b</sup>	A small effect, not clinically relevant, in pain improvement wa detected in favour of MT. This analysis excluded two trials (one suspected of publication bias, on used a different scale) which increased heterogeneity levels bu did not affect overall efficacy meaningfully.	
Adverse events assessed with: number of AE occurred	144 per 1.000	<b>121 per</b> <b>1.000</b> (79 to 184)	<b>RR 0.84</b> (0.55 to 1.28)	531 (6 RCTs)	⊕⊕⊖⊖ Low ª	Pooled data from 6 studies did no show any difference in AE occurrence between ST and MT.	
Dropouts rate assessed with:	174 per	<b>171 per</b> <b>1.000</b> (134 to	<b>RR 0.98</b> (0.77 to	1238 (11 RCTs)	⊕⊕⊖⊖ Low ª	Pooled data from 11 trials did no show difference in dropout rate between ST and MT.	
number of participants that leaved the study	1.000	218)	1.25)		•		
participants that leaved the study 2. ST compared	to no treatn	218)	1.25)				
participants that leaved the study	<b>to no treatn</b> back pain	218)	1.25)		2		
participants that leaved the study 2. ST compared Patient or population: Intervention: ST Comparison: No treat	to no treatm back pain ment Anticipate	218)	Relative	Nº of	Certainty of		
participants that leaved the study <b>2. ST compared</b> Patient or population: Intervention: ST	to no treatm back pain ment Anticipate	218) nent				Comments	
participants that leaved the study 2. ST compared Patient or population: Intervention: ST Comparison: No treat	to no treatm back pain ment Anticipate effects* Risk with no	218) nent ed absolute (95% CI)	Relative effect	Nº of participants	Certainty of the evidence		

CI: Confidence interval; MD: Mean difference; RR: Risk ratio

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

#### Explanations

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

b. Most of the studies were small trial.

c. Heterogeneity levels at 80%.

d. Number of participants < 400

to oper teries only

33

34 35

36

37

38 39

44 45 46

47

48

49 50

51

53

55

### Sham treatment versus other manual therapies

#### Pain

The following outcomes on back pain are presented with a 100 mm visual analogue scale, 0 to 100; higher scores refer to worse pain. Trials using a 10mm scales were converted to 100mm 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 randomeffects 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 distribution with the inclusion of these two studies (*Appendix 3*). This sensitivity analysis did not influence overall effectiveness results, but inconsistency levels decreased considerably at short-term. It can be deducted that a possible cause of heterogeneity was found (Full analysis in 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<sup>2</sup>=42%, p<0.0001, very low quality of evidence downgraded two levels for very serious risk of bias and imprecision) (Figure 3).

Figure 3: Forest plot of comparison ST versus MT in back pain outcome at short-term.

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<sup>2</sup>=91% P<0.0001, long-term I<sup>2</sup>=81% P=0.005) (*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 52 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. 54
- One trial compared ST to an articular mobilization technique. 54.5% participants correctly 56 guessed the treatment allocation. (46)
- 57 Participants of one study that compared ST to reflexology had the lowest percentage of correct 58 detection of allocation (46.7%). Participants in this trials did not know about the type of 59 treatment tested. (39) 60

#### Dropouts

Pooled data from 11 trials at the last follow-up suggested no difference in dropouts rate between ST and MT at the end of the trials (105/612 compared to 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)*.

Figure 4: Forest plot of comparison ST versus MT in number of dropouts outcome

#### Adverse effects

AE 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) AE were predominantly minor and lasted for two/three days after treatment, in the majority of trials transient worse pain, tiredness, muscle weakness and transient headache were reported. (26, 36, 37, 45)

Senna M 2011 reported the most common AE were local discomfort and tiredness but no serious complications were noted. (32)

Haller H 2016 reported two patients dropping out from the trial for recurrent headache after treatments, both Haller H and Klein R 2013 reported dizziness of one patient. (36, 37) Licciardone J 2013 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 to 38/264; RR 0.84, 95% CI 0.55 to 1.28; 531 participants, I<sup>2</sup>=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.

*Figure 5:* Forest plot of comparison ST versus MT in number of adverse events outcome at shortterm

#### Sham versus no treatment

#### Pain

Four studies compared ST to no intervention, three were included in random-effect metaanalysis 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 S. 2013 showed an effect in favour of ST (pooled data from 3 trials: MD -5.84, 95% CI from -20.46 to 8.78, 252 participants, I<sup>2</sup>= 85%, P=0.43). The exclusion of Erdogmus S 2013 (that used a 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<sup>2=</sup> 84%, P=0.33) (**Appendix 5**).

#### Dropouts

No differences were shown in the fixed-effect meta-analysis on dropout rate between ST and no intervention in four trials (14/112 compared to 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)*.

Figure 6: Forest plot of comparison ST versus no treatment in number of dropouts outcome

#### Adverse effects

Of the four studies comparing ST to no intervention, only two reported AE. One, Eardley S 2013, did not evaluate AE occurred in the no treatment group while Erdogmus C 2007 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.

AE 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 no treatment group with no significant difference from ST.

SM techniques were the treatment most evaluated (N=7). These techniques are highly recognizable 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 recognized 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 to 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 dropouts rate, although both these data were reported only in two trials. Bialosky J and Hoiriis K showed high percentages of correct treatment allocation detection by participants but dropout 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 dropout rate was reached in the comparison sham versus no treatment. These data suggest that dropout 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 to 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). (**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 JK 2016 included studies where SM was compared to 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 SM 2019 compared SM and mobilisation techniques to recommended, nonrecommended therapies and to ST. Their findings showed that 5/47 studies included attempted to blind patients to the assigned intervention by providing a 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 SM *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 metaanalysis 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 metaanalysis. 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.

Pair-wise 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 R 2014 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 multi-disciplinary 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 to no intervention. Furthermore, when blinding was guaranteed, ST showed a statistically significant effect on pain reduction in chronic LBP patients compared to 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 localization 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.

NICE 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 multi-disciplinary 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.

#### **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.

**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.

Ethical approval: due to the chosen study design, the ethical approval was not required.

**Data sharing:** Details of the characteristics of the included studies and data extracted are available from the corresponding author at <u>carolina.lavazza@docenti.aimoedu.it</u>. Extra data can be accessed via the Dryad data repository at http://datadryad.org/ with the doi: 10.5061/dryad.v9s4mw6tb

## References

1. Meissner K, Fässler M, Rücker G, Kleijnen J, Hróbjartsson A, Schneider A, et al. Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. . JAMA internal medicine. 2013;173(21):1941-51.

2. Sciences CfIOoM. International Ethical Guidelines for Health-Related Research Involving Humans, Fourth Edition. Geneva2016.

3. 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).

4. Korthals-de Bos IB, Müllner M, Hoving JL, van Tulder MW, Rutten-van Mölken MP, Adèr HJ, 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(7395):1-6.

5. Kawchuk G N, Haugen R, Fritz J. A True Blind for Subjects Who Receive Spinal Manipulation Therapy. Archives of Physical Medicine and Rehabilitation. 2009;90(2):366-8.

6. Mancini F, Beaumont AL, Hu L, Haggard P, Iannetti GDD. Touch inhibits subcortical and cortical nociceptive responses. Pain. 2015;*156*(10).

7. So PS, Jiang JY, Qin Y. Touch therapies for pain relief in adults. . Cochrane Database of Systematic Reviews. 2008.

8. Cai L., He L. Placebo effects and the molecular biological components involved. General psychiatry. 2019;32(5):1-10.

9. Kaptchuk T. J., Miller F. G. Placebo effects in medicine. N Engl J Med. 2015;373(1):8-9.

10. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Placebo effects: biological, clinical and ethical advances. . Lancet. 2010;375:686.

11. 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.

12. Miller C. B., Espie C. A., Epstein D. R., et al. The evidence base of sleep restriction therapy for treating insomnia disorder. . Sleep medicine reviews. 2014;18(15):415-24.

13. Colloca L, Miller FG. The nocebo effect and its relevance for clinical practice. Psychosomatic medicine. 2011;73:598-603.

14. Koes BW. How to evaluate manual therapy: value and pitfalls of randomized clinical trials. Manual therapy. 2004:183-4.

15. Chaibi A, Benth JŠ, Russell MB. Validation of placebo in a manual therapy randomized controlled trial. Scientific reports. 2015;5(1):1-8.

16. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. Cochrane database of systematic reviews. 2004;2.

17. Kaptchuk TJ, Stason WB, Davis RB, Legedza AR, Schnyer RN, Kerr CE, et al. Sham device v inert pill: randomised controlled trial of two placebo treatments. BMJ. 2006;332:391-7.

18. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M et al. Preferred
 reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.
 Systematic Reviews. 2015;4(1):1.
 10. Uigging IDT. Groop S. Cochrang Handback for Systematic Reviews of

19. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. <u>https://handbook-5-</u>

	1.cochrane.org/index.htm#chapter 8/8 assessing risk of bias in included studies.htm: The
	Cochrane Collaboration; 2011.
	20. Pool JJ, Ostelo RW, Hoving JL, Bouter LM, De Vet HC. Minimal clinically important change
	of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain Spine.
	2007;32(26):3047-51.
	21. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting
0	change scores for pain and functional status in low back pain: towards international consensus
1	
2	regarding minimal important change Spine. 2008;33(1):90-4.
3	22. Farrar JT, Young Jr JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes
4	in chronic pain intensity measured on an 11-point numerical pain rating scale Pain.
5	2001;94(2):149-58.
б	23. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions
3	Version 5.1.0 [updated March 2011].9.5.2 Identifying and measuring heterogeneity.
)	https://handbook-5-
)	<u>1.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm</u> : The
	Cochrane Collaboration; 2011.
	24. Licciardone JC, Minotti DE, Gatchel RJ, Kearns CM, Singh K. Osteopathic manual
	treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. The
	Annals of Family Medicine. 2013;11(2):122-9.
	25. Bialosky JE, George SZ, Horn ME, Price DD, Staud R, Robinson ME. Spinal manipulative
	therapy–specific changes in pain sensitivity in individuals with low back pain (NCT01168999)
	The journal of pain. 2014;15(2):136-48.
	26. Eardley S, Brien S, Little P, Prescott P, Lewith G. Professional kinesiology practice for
	chronic low back pain: single-blind, randomised controlled pilot study Complementary
	Medicine Research. 2013;20(3):180-8.
	27. Erdogmus CB, Resch KL, Sabitzer R, Müller H, Nuhr M, Schöggl A, et al. Physiotherapy-
	based rehabilitation following disc herniation operation: results of a randomized clinical trial
	Spine. 2007;32(19):2041-9.
	28. Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, 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(1):98-108.
	29. Hoiriis KT, Pfleger B, McDuffie FC, Cotsonis G, Elsangak O, Hinson R, et al. A randomized
	clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain.
	Journal of Manipulative and Physiological Therapeutics. 2004;27(6):388-98.
	30. Krekoukias G, Gelalis ID, Xenakis T, Gioftsos G, Dimitriadis Z, Sakellari V. Spinal
	mobilization vs conventional physiotherapy in the management of chronic low back pain due to
	spinal disk degeneration: a randomized controlled trial Journal of Manual & Manipulative
	Therapy. 2017;25(2):66-73.
	31. Licciardone JC, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W, et al. Osteopathic
	manipulative treatment for chronic low back pain: A randomized controlled trial. Spine.
	2003;28(13):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(18).
	33. Lascurain-Aguirrebeña I, Newham DJ, Casado-Zumeta X, Lertxundi A, Critchley DJ.
	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. Musculoskeletal Science and Practice. 2018;38:83-90.
	28

**BMJ** Open

34. Selkow NM, Grindstaff TL, Cross KM, Pugh K, Hertel J, Saliba S. Short-term effect of muscle energy technique on pain in individuals with non-specific lumbopelvic pain: a pilot study. . Journal of manual & manipulative therapy. 2009;19(1):14E-8E.

35. Antolinos-Campillo PJ, Oliva-Pascual-Vaca A, Rodríguez-Blanco C, Heredia-Rizo AM, Espí-López GV, Ricard F. 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(3):249-55.

36. Haller H, Lauche R, Cramer H, Rampp T, Saha FJ, Ostermann T, et al. Craniosacral therapy for the treatment of chronic neck pain: a randomized sham-controlled trial. The Clinical journal of pain. 2016;32(5):441-9.

37. Klein R, Bareis A, Schneider A, Linde K. Strain–counterstrain to treat restrictions of the mobility of the cervical spine in patients with neck pain—A sham-controlled randomized trial. Complementary Therapies in Medicine. 2013;21(1):1–7.

38. Younes M, Nowakowski K, Didier-Laurent B, Gombert M, Cottin F. Effect of spinal manipulative treatment on cardiovascular autonomic control in patients with acute low back pain. Chiropractic and Manual Therapies. 2017;25(1).

39. Quinn F, Hughes CM, Baxter GD. Reflexology in the management of low back pain: A pilot randomised controlled trial. Complementary Therapies in Medicine. 2008;16(1):3-8.

40. Sillevis R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. Journal of Manual & Manipulative Therapy ,. 2010;18.

41. Cleland JA, Childs MJD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. *Manual therapy*. 2005;10(2):127-35.

42. Pires PF, Packer AC, Dibai-Filho AV, Rodrigues-Bigaton D. 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. Journal of manipulative and physiological therapeutics. 2015;38(8).

43. Silva AC, Biasotto-Gonzalez DA, Oliveira FHM, Andrade AO, Gomes CADFP, Lanza FDC, 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. . Evidence-Based Complementary and Alternative Medicine. 2018;2018:1-9.

44. Vieira-Pellenz F, Oliva-Pascual-Vaca Á, Rodriguez-Blanco C, Heredia-Rizo AM, Ricard F, Almazán-Campos G. 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. . Archives of physical medicine and rehabilitation. 2014;95(9):1613-9.

45. Kogure A, Kotani K, Katada S, Takagi H, Kamikozuru M, Isaji T, 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(12):1-16.

46. Hall T, Hardt S, Schafer A, Wallin L. Mulligan bent leg raise technique-a preliminary randomized trial of immediate effects after a single intervention. Manual Therapy. 2006;11(2):130-5.

47. Dunning J, Mourad F, Barbero M, Leoni D, Cescon C, Butts R. Bilateral and multiple cavitation sounds during upper cervical thrust manipulation. . BMC musculoskeletal disorders. 2013;14(1):24.

48. BC J, S E, A C-L, Furukawa TA, Patrick DL, Crawford MW et al. Minimally important difference estimates and methods: a protocol. BMJ Open. 2015;5:1-5.

49. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials. BMJ. 2009;388.

50. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. . International journal of surgery. 2012;10(1):28-55.

51. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PloS one. 2013;8(3).

52. Ruddock JK, Sallis H, Ness A, Perry RE. Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. Journal of chiropractic medicine. 2016;15(3):165-83.

53. Rubinstein SM, De Zoete A, Van Middelkoop M, Assendelft WJ, De Boer MR, Van Tulder MW. 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.

54. Monroe CM. The effects of therapeutic touch on pain. Journal of Holistic Nursing. 2009;27(2):85-92.

55. Rubinstein SM, van Eekelen R, Oosterhuis T, de Boer MR, Ostelo RW, van Tulder MW. The risk of bias and sample size of trials of spinal manipulative therapy for low back and neck pain: analysis and recommendations. Journal of manipulative and physiological therapeutics. 2014;37(8):523-41.

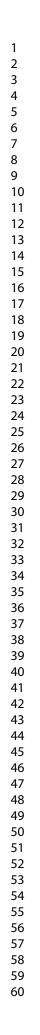
56. Defrin R, Devor M, Brill S. Tactile allodynia in patients with lumbar radicular pain (sciatica). *PAIN®*. 2014;155(12):2551-9.

57. Nijs J, Van Houdenhove B. From acute musculoskeletal pain to chronic widespread pain and fibromyalgia: application of pain neurophysiology in manual therapy practice. . Manual therapy. 2009;14(1):3-12.

58. Geri T, Viceconti A, Minacci M, Testa M, Rossettini G. 2019. Musculoskeletal Science and Practice. Manual therapy: exploiting the role of human touch. ;44.

59. Walker SC, Trotter PD, Swaney WT, Marshall A, Mcglone FP. 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 https://www.nice.org.uk/guidance/ng59/chapter/Recommendations11 December 2020 [



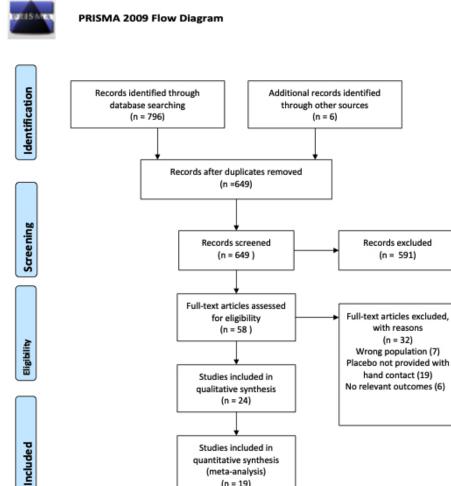


Figure 1: PRISMA flow diagram

(meta-analysis)

(n = 19)

48x56mm (298 x 300 DPI)

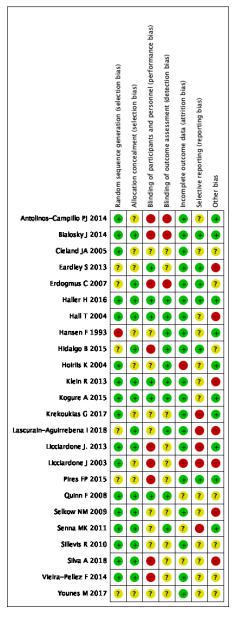


Figure 2: Risk of bias summary. Review authors' judgements about each risk of bias item for each included study.

34x93mm (298 x 299 DPI)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		sт			мт			Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean		Total	Mean		Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% CI	
4.1.1 Short-term								,		
Antolinos-Campillo PJ 2014	55.2	10.5	20	48.7	10.8	20	0.6%	6.50 [-0.10, 13.10]		970097
Bialosky j 2014		22.45		34.03		27		-2.28 [-15.20, 10.64]		
Cleland JA 2005	43.5		17			19	0.2%	17.40 [5.33, 29.47]		422242
Eardley S 2013	36	19.8	21		22.7	20	0.2%	7.00 [-6.06, 20.06]	+	224244
Erdogmus C 2007	21.6	3.13		24.13	2.5	40		Not estimable		200000
Haller H 2016	53.5	20.3	27	31.7	20.7	27	0.3%			
Hoiriis K 2004	22.1	20.2	40			34	0.4%	5.00 [-3.89, 13.89]	<u>+-</u>	
Klein R 2013	38	26	31	32	22	30	0.2%	6.00 [-6.07, 18.07]	<u> </u>	
Kogure A 2015	47.6		89	43.4	20	90	0.9%	4.20 [-1.82, 10.22]	-	
Krekouklas G 2017	58.8	9.2	25	12.2	11	25	0.074	Not estimable		
Licclardone J 2003	30.7	21.9	23	37.7		42	0.2%	-7.00 [-18.95, 4.95]		
Pires FP 2015	25.7	22.7	16	22		16	0.1%	3.70 [-11.69, 19.09]		22020
Selkow NM 2009	35.2	28	10	25	20.6	10		10.20 [-11.34, 31.74]		
Senna MK 2011	33.19			29.46	1.16	26	95.3%	3.73 [3.14, 4.32]		
Silva A 2018	47.1	17.2	14			14	0.2%	15.00 [1.18, 28.82]		
Vieira-Pellez F 2014		26.37			22.47	20	0.1%	9.09 [-6.09, 24.27]		
Younes M 2017	18.14		- 20	12.6		10	0.7%	5.54 [-1.21, 12.29]		222202
Subtotal (95% CI)	10.14	7.40	400	12.0	9.24		100.0%	3.86 [3.29, 4.43]		
Heterogeneity: $Chl^2 = 23.95$ , Test for overall effect: $Z = 13$				- 42%						
Total (95% CI)			400			405	100.0%	3.86 [3.29, 4.43]	r -	
Heterogeneity: Chi <sup>2</sup> = 23.95,	df = 14	(P = 0.6)	05): P	42%				-		_
Test for overall effect: Z = 1									-50 -25 0 25 50	
Test for subgroup difference.									Favours ST Favours MT	
Risk of bias legend										
(A) Random sequence generation	ation (sel	ection bi	as)							
(B) Allocation concealment (se			,							
(C) Blinding of participants an			formar	ce bias						
(D) Blinding of outcome asses										
(E) Incomplete outcome data			/							
(F) Selective reporting (report										
(G) Other bias	5									

Figure 3: Forest plot of comparison ST versus MT in back pain outcome at short-term.

84x50mm (300 x 298 DPI)

	ST		MT			Risk Ratio	Risk Ratio	<b>Risk of Bias</b>
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI	M-H, Fixed, 95% CI	ABCDEFO
Blakosky J 2014	1	28	1	27	1.0%	0.96 [0.06, 14.65]		9999999
Eardley S 2013	1	21	0	20	0.5%	2.86 [0.12, 66.44]		??
Erdogmus C 2007	8	40	5	40	4.6%	1.60 [0.57, 4.47]	<b></b>	? • • • • • ?
Haller H 2016	13	27	11	27	10.5%	1.18 [0.65, 2.15]	- <b>-</b>	9999999
Hansen F 1993	5	59	11	57	10.7%	0.44 [0.16, 1.18]		•??••?•
Hoirlis K 2004	13	53	13	50	12.6%	0.94 [0.49, 1.83]		9??99?
Kogure A 2015	4	69	3	90	2.8%	1.35 [0.31, 5.85]		<b></b>
Licclardone J 2003	4	23	16	46	9.9%	0.52 [0.20, 1.39]		•?•?••
Licclardone J. 2013	34	225	39	230	36.8%	0.89 [0.58, 1.36]		
Senna MK 2011	20	40	7	27	6.0%	1.93 [0.95, 3.92]		
Younes M 2017	2	7	3	10	2.4%	0.95 [0.21, 4.29]		2222927
Total (95% CI)		612		626	100.0%	0.98 [0.77, 1.25]	•	
Total events	105		109					
Heterogeneity: Chi <sup>2</sup> =	9.69, df	= 10 (F	P = (0.47)	i; i² = 0	<b>%</b>		0.01 0.1 1 10 100	
Test for overall effect:	z = 0.12	(P = 0	.90)				Favours ST Favours MT	
Risk of bias legend								
(A) Random sequence	e generatio	on (sele	ction bia	s)				
(B) Allocation concealr	ment (sele	ction bi	as)					
(C) Blinding of particip	pants and	person	nel (perf	ormanc	e bias)			
(D) Blinding of outcom				()				

(G) Other bias

Figure 4: Forest plot of comparison ST versus MT in number of dropouts outcome

72x38mm (299 x 299 DPI)

1	
2	
3	
4	
5	
6	ST MT Risk Ratio Risk Ratio Risk of Bias
7	Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI M-H, Fixed, 95% CI A B C D E F G Eardley \$ 2013 2 21 1 20 2.6% 1.90 [0.19, 19.40] 7 9 9 9 9
8	Erdogmus C 2007 11 40 12 40 31.0% 0.92 [0.46, 1.83] 🚽 🦩 🔴 😨 😨 🖓
9	Haller H 2016 8 27 6 27 15.5% 1.33 [0.53, 3.33]
10	Klein R 2013 1 31 4 30 10.5% 0.24 [0.03, 2.04] 🛛 🚽 🕒 🕒 🕀 🕀 🕀 🕐 🕐
11	Kogure A 2015 10 89 8 90 20.6% 1.26 [0.52, 3.05]
12	Total (95% CI) 267 264 100.0% 0.84 [0.55, 1.28]
13	Total events       32       38         Heterogenety: $6.79$ , df = 5 (P = 0.24); $t^2$ = 26%       0.001       0.1       1       10       1000
	Test for overall effect: Z = 0.81 (P = 0.42)         0.001         0.1         1         10         1000'           Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT         Favours ST Favours MT
14	Risk of bias legend
15	(A) Random sequence generation (selection bias)
16	(B) Allocation concealment (selection bias) (C) Blinding of participants and personnel (performance bias)
17	(D) Blinding of outcome assessment (detection bias)
18	(E) Incomplete outcome data (attrition bias) (F) Selective reporting (reporting bias)
19	(G) Other bias
20	
	Figure 5: Forest plot of comparison ST versus MT in number of adverse events outcome at short-term
21	righte 5. To lest plot of comparison 51 versus MT in number of adverse events outcome at short-term
22	72x31mm (299 x 298 DPI)
23	72X3111111 (233 X 230 D11)
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
49	

1	
2	
3	
4	
5	
6	er la susse pisto pisto pisto pisto de la companya
7	ST No treatment Risk Ratio Risk Ratio Risk of Bias Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI ABCDEFG
8	Biakosky j 2014 1 28 0 28 2.9% 3.00 [0.13, 70.64]
9	Erdogmus C 2007 B 40 B 40 45.7% 1.00 [0.42, 2.40] — 🗰 🥂 🖓 😨 💭 😨 😨 🖓
-	Lioclardone j 2003 4 23 5 20 30.6% 0.70 [0.22, 2.24] 🗕 🖷 🗣 ? 🗣 ? 🗣 ?
10	Total (95% CI) 112 113 100.0% 0.82 [0.43, 1.55]
11	Total events 14 17 Heterogenety: Ch <sup>2</sup> = 1.61, df = 3 ( $P = 0.61$ ); $P = 0\%$ 0.01 0.1 1 10 100
12	Test for overall effect:         Z = 0.62 (P = 0.54)         0.01         0.1         1         10         100'           Test for overall effect:         Z = 0.62 (P = 0.54)         Favours ST Favours Control
13	<u>Risk of bias legend</u>
14	(A) Random sequence generation (selection bias) (B) Allocation concealment (selection bias)
15	(C) Blinding of participants and personnel (performance bias)
16	(D) Blinding of outcome assessment (detection bias) (E) Incomplete outcome data (attrition bias)
17	(F) Selective reporting (reporting bias) (G) Other bias
18	(u) Other Dias
19	
20	Figure 6: Forest plot of comparison ST versus no treatment in number of dropouts outcome
21	
22	73x28mm (299 x 299 DPI)
23	
24	
24	
26	
27	
28	
29	
30	
31	

## Appendix 1: search strategy

#### Medline

1 2 3

4 5 6

7 8

9

10 11

12

13

14

15

16 17

18

19

21 22

23

24

25

26

27

29

30

31

32

33 34

37

39

43

45

46

47

48

49 50

51

52

53

54

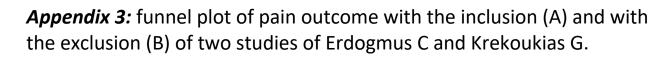
- Mesh descriptor: [Back Pain] explode all trees
- 2. dorsalgia/
- 3. backache
  - 4.(neck OR cervical) adj1 pain  $\rightarrow$  Mesh
  - 5. exp Brachial Plexus Neuropaties
- 6. exp Lumbar Plexus Neuropaties
- 7. Neck Pain/
- 8. neckache
- 9. Torticollis/
- 10. whiplash.mp 20
  - 11. cervicodynia.mp
    - 12. spondylitis/ OR spondylosis/ OR spondylolysis/ OR spondylolysthesis
  - 13.(lumbar OR dorsal OR neck OR cervical OR sciatica) adj2 (pain OR ache)
  - 14. (lumbar OR dorsal OR neck OR cervical) adj2 (discitis OR disc adj 1 herniation OR disc adj1 herniation)
  - 15. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
- 28 16. (PLACEBO\* or MOCK\* or SHAM\* or FAKE\* or VEHICLE\* or DUMM\* or ATTENTION\* CONTROL\*
  - or PSEUDO\* TREAT\* or UN?SPECIFIC\* or NON?SPECIFIC\* OR simulat\$ treatment OR inert agent) 17. Chiropractic/
  - 18. Manipulation, Chiropractic/
  - 19. chiropract\$.tw.
  - 20. (manual adj2 therap\$).mp
- 35 21. spinal manipulation.mp. or Manipulation, Spinal/ 36
  - 22. osteopath\$.tw.
- 23. Osteopathic Medicine/ 38
  - 24. Physical Therapy Modalities/ or "Physical Therapy (Specialty)"/ or physical therap\$.tw. or
- 40 physiotherap\$.tw. 41
- 25. myotherapy.mp 42
  - 26. shiatsu.mp
- 44 27.exp Therapeutic Touch/
  - 28. exp Massage/
  - 29. (neuromuscular adj therapy).mp
  - 30. 17 OR 18 OR 19 OR 20 OR OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29
    - 31. pain
  - 32. range of motion
  - 33. ROM
  - 34. 31 OR 32 OR 33
  - 35. Clinical Trial/
- 55 36. Randomized Controlled Trial/
- 56 37. controlled clinical trial/ 57
  - 38. exp RANDOMIZATION/
  - 39. PLACEBO/
- 59 40. (random\$ adj2 allocat\$).tw. 60
  - 41. single blind\$.tw.

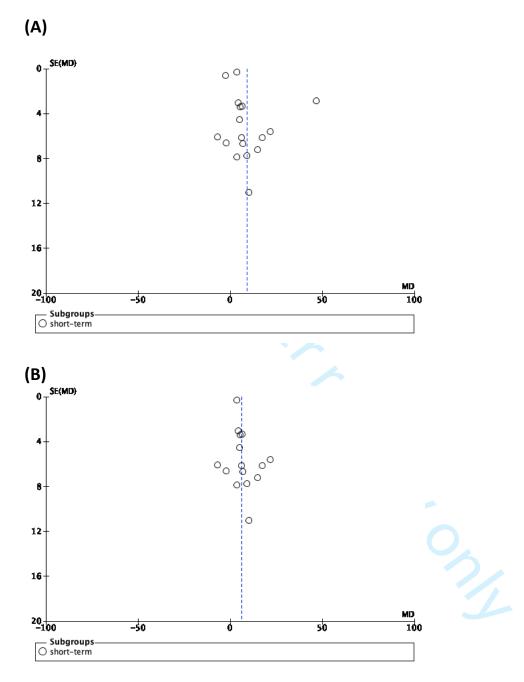
2	
3	42. double blind\$.tw.
4 5	43. placebo\$.tw.
5 6	44. 35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43
7	45. animals/
8	46. humans/
9	47. 46 NOT 45
10	48. 15 AND 16 AND 30 AND 44 AND 47
11	48. 15 AND 16 AND 50 AND 44 AND 47
12	
13 14	
15	
16	
17	
18	
19	
20 21	
21 22	
22	
24	
25	
26	
27 28	
28 29	
30	
31	
32	
33	
34 35	
35 36	
37	
38	
39	
40	
41 42	
42 43	
44	
45	
46	
47	
48 49	
49 50	
51	
52	
53	
54	
55 56	
56 57	
58	
59	
60	

## Appendix 2: data extraction form

Methods	Trial Design
	Settings
Participants	Total number of participants:
	Age:
	Gender(M/F):
O,	BMI:
	Activity:
	Duration of the symptoms:
	Location of pain (one-sided, double-sided, central, cervical, dorsal or lumbar):
	<b>Cause of pain:</b> (e.g. disc herniation, contractures, aspecific pain)
	Previous experience of the treatment provided: Y/N/ N/A
	Inclusion and exclusion Criteria:
	VAS:
	Practitioner characteristics: (years of experience, gender)
Interventions	Placebo:
	Comparator:
Outcomes	Outcomes used in the meta-analysis:
	Length of follow-up:
Notes	Difference between Placebo and active treatment:
	Placebo check for reliability:
	Adverse event:

Lost to follow-up:
 Funding source:





**Appendix 4:** forest plot of comparison pain outcome sham treatment vs manual therapies with the inclusion of two trials (Erdogmus C and Krekoukias G) at short, medium and long-term.

		ST			мт			Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean		Total	Mean		Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	
3.1.1 Short-term								,	,	
Antolinos-Campillo PJ 201		10.5				20	4.6%	6.50 [-0.10, 13.10]	-	• ? • • • ? •
Białosky J 2014		22.45	-	34.03 26.1		27 19		-2.28 [-15.20, 10.64]	—	
Cleland JA 2005 Eardley S 2013	43.5 36	19.5 19.6	17 21		22.7	20	2.4%	17.40 [5.33, 29.47] 7.00 [-6.06, 20.06]		
Erdogmus C 2007	21.6	3.13		-	2.5	40	7.1%	-2.53 [-3.77, -1.29]	-	200000
Haller H 2016	53.5	20.3			20.7	27	2.6%	21.80 [10.86, 32.74]		9999999
Hoirils K 2004	22.1	20.2	-			34	3.5%	5.00 [-3.89, 13.89]	+	• ? ? • • ? •
Klein R 2013 Kogure A 2015	38 47.6	26 21.1	31 69	32 43.4	22 20	30 90	2.4% 4.9%	6.00 [-6.07, 18.07] 4.20 [-1.82, 10.22]		444442
Krekouklas G 2017	58.8	9.2		-	-	25	5.1%		-	• ? ? ? • •
Licclardone J 2003	30.7	21.9	23	-		42	2.5%	-7.00 [-18.95, 4.95]	+	• ? • ? • •
Pires FP 2015	25.7	22.7	16	22		16	1.7%		<u> </u>	<b>? ? • ? • •</b>
Selkow NM 2009 Senna MK 2011	35.2 33.19	28 1.19		25 29.46		10 26	7.2%	10.20 [-11.34, 31.74] 3.73 [3.14, 4.32]		
Silva A 2018	47.1	17.2			20.01	14	2.0%	15.00 [1.18, 28.82]		
Vieira-Pellez F 2014	29.1	26.37	20	20.01	22.47	20	1.6%	9.09 [-6.09, 24.27]	+	
Younes M 2017	18.14	7.46			6.24	10	4.5%	5.54 [-1.21, 12.29]	•	22224
Subtotal (95% CI) Heterogeneity: Tau <sup>2</sup> = 62	40· Chi -	330 75	465		00001	470		8.96 [4.45, 13.47]	•	
Test for overall effect: Z =						<i>.</i> ,,, = ;				
3.1.2 Medium-term										
Erdogmus C 2007 Haller H 2016	18.35 47.8	3.03 19.3		17.11 31.6	5.48 19	40 27	6.9X 3.0X	1.24 [-0.70, 3.18] 16.20 [5.98, 26.42]	t	?
Kogure A 2015	47.8	22.2		36		90	5.0%	10.30 [4.45, 16.15]		44444
Senna MK 2011	35.16		37	35.16		26	7.2%	0.00 [-0.64, 0.64]		•••
Subtotal (95% CI)			193			183	22.0%	3.93 [0.29, 7.57]	•	
Heterogeneity: Tau <sup>2</sup> = 9.3 Test for overall effect: Z =			f = 3 (P	< 0.00	01); r -	67%				
	• • • • • • • •	0.03)								
3.1.3 Long-term										
Erdogmus C 2007	16.85	3.71		21.05		40 26	7.0%	-4.20 [-5.80, -2.60]	•	
Kogure A 2015 Licclardone J 2003	45.5 24.5	22 21.1		31.2 31.6		32	3.6× 2.4×	14.30 [5.75, 22.85] -7.10 [-19.36, 5.16]		
Senna MK 2011	36.8		-	35.53		26	7.1%	1.27 [0.34, 2.20]	-	
Subtotal (95% CI)			185			124		0.70 [-4.32, 5.71]	•	
Heterogeneity: Tau <sup>2</sup> = 18 Test for overall effect: Z =			df = 3 (	(P < 0.0	0001);	r = 93)	×			
rest for overall effect. Z =	• U.27 (P = 1	y.79)								
Total (95% CI)			843			777	100.0%	6.35 [4.02, 8.67]	•	
Heterogeneity: $Tau^2 = 19$				:4 (P < (	).00001	.);	95%	-	-50-25 0 25 50	_
Test for overall effect: Z = Test for subgroup differen				- 0.00	1 F - 4				Favours ST Favours MT	
Risk of bias legend		4.04, 0	(1	- 0.03		N. 37				
(A) Random sequence ge	neration (sel	ection b	ias)							
(B) Allocation concealmen										
(C) Blinding of participant				nce bias)						
<ul> <li>(D) Blinding of outcome as</li> <li>(E) Incomplete outcome data</li> </ul>			i dias)							
(F) Selective reporting (rep	<b>J</b>									
(G) Other bias										

# **Appendix 4.1:** Sensitivity analysis with the exclusion of Ergogmus C and Krekoukias G at short, medium and long-term

2											
3			ST			мт			Mean Difference	Mean Difference	<b>Risk of Bias</b>
4	Study or Subgroup 3.1.1 Short-term	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
5	Antolinos-Campillo PJ 2014	55.2	10.5	20	46.7	10.8	20	4.6%	6.50 [-0.10, 13.10]	-	• ? • • • ? •
6	Blalosky J 2014		22.45	28	-				-2.28 [-15.20, 10.64]	<u> </u>	
	Cleland JA 2005	43.5		17	26.1			2.0%	17.40 [5.33, 29.47]		<b>+</b> ??? <b>+</b> ??
7	Eardley \$ 2013	36	19.8	21	29	22.7	-	1.7%	7.00 [-6.06, 20.06]		<b>? ? <del>9</del> ? <del>9</del> 9 <b>9</b></b>
8	Erdogmus C 2007	21.6	3.13	40 27		2.5	40 27	0.0%	-2.53 [-3.77, -1.29]		? • • • • • • ?
9	Haller H 2016 Hoirlis K 2004	53.5 22.1	20.3		31.7 17.1	20.7 18.8		2.3× 3.2×	21.80 [10.86, 32.74] 5.00 [-3.89, 13.89]		
	Klein R 2013	36	26	31	32		-	2.0%	6.00 [-6.07, 18.07]	<u> </u>	
10	Kogure A 2015	47.6	21.1	69	43.4	20	90	5.1%	4.20 [-1.82, 10.22]		<b>9999979</b>
11	Krekouklas G 2017	58.8	9.2		12.2			0.0%			•••••
12	Licclardone J 2003	30.7	21.9	-	37.7	-			-7.00 [-18.95, 4.95]	+	• ? • ? • • •
	Pires FP 2015 Selkow NM 2009	25.7 35.2	22.7 28	16 10	22 25	21.7 20.6		1.3%	3.70 [-11.69, 19.09] 10.20 [-11.34, 31.74]		?? <b>?</b> ? <b>???</b>
13	Senna MK 2011	33.19	1.19		29.46	1.16			3.73 [3.14, 4.32]		
14	Silva A 2018	47.1	-	-		20.01	-	1.6%	15.00 [1.18, 28.82]		
15	Vieira-Pellez F 2014	29.1	26.37	20	20.01	22.47	20	1.3×	9.09 [-6.09, 24.27]	+	•••?•?
	Younes M 2017	18.14	7.46		12.6	6.24		4.5%	5.54 [-1.21, 12.29]		??????
16	Subtotal (95% CI)	<b></b>		400		7	405	45.0%	6.09 [3.23, 8.96]	•	
17	Heterogeneity: Tau <sup>2</sup> = 9.37; Test for overall effect: Z = 4				P = 0.0	5); r =	42%				
18	3.1.2 Medium-term										
19	Erdogmus C 2007	18.35	3.03	40	17.11	5.48	40	9.9%	1.24 [-0.70, 3.18]		? • • • • • ?
20	Haller H 2016	47.8	19.3	27	31.6	19	27	2.6%	16.20 [5.98, 26.42]		9999999
21	Kogure A 2015	46.3	22.2		36	17.4	90	5.3%	10.30 [4.45, 16.15]	-	<b>~~~</b> ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
	Senna MK 2011 Subtotal (95% CI)	35.16	1.25	37 193	35.16	1.26	26 183	10.9% 28.7%	0.00 [-0.64, 0.64]	L.	99?9?
22	Heterogeneity: Tau <sup>2</sup> = 9.14	Chf = 2	2.20 4		~ 0.00	01) P -		20.7%	3.93 [0.29, 7.57]	•	
23	Test for overall effect: $Z = 2$				~ 0.00	V1), I -	- 07/4				
24	3.1.3 Long-term										
25	Erdogmus C 2007	16.85	3.71	40	21.05	3.58	40	10.2%	-4.20 [-5.80, -2.60]	-	? • • • • • ?
26	Kogure A 2015	45.5	22	69	31.2	16.6	26	3.3%	14.30 [5.75, 22.85]		<b>~~~</b>
	Licclardone J 2003	24.5	21.1	-	31.6	22.4	32		-7.10 [-19.36, 5.16]		• ? • ? • •
27	Senna MK 2011 Subtotal (95% CI)	36.8	1.39	185	35.53	2.13	26 124	10.8X 26.3%	1.27 [0.34, 2.20] 0.70 [-4.32, 5.71]	1	9979799
28	Heterogeneity: Tau <sup>2</sup> = 18.0	2· Chi <sup>2</sup>	45 60 /			00011			0.70 [-4.32, 3.71]	Ť	
29	Test for overall effect: Z = 0			ui – Ji,		VVV1)	557	~			
30	Total (95% CI)			778			712	100.0%	3.85 [1.98, 5.71]		
31	Heterogeneity: $Tau^2 = 6.19$			df = 22	(P < 0.	00001)				-50-25 0 25 50	_
	Test for overall effect: $Z = 4$									Favours ST Favours MT	
32	Test for subgroup difference	es: Chi <sup>e</sup> =	3.49, d	f = 2 (i	<b>? = 0</b> .17	7), F = 4	2.7%				
33	Risk of bias legend	ration (cal	action b	iac)							
34	<ul> <li>(A) Random sequence gener</li> <li>(B) Allocation concealment (sequence)</li> </ul>			ia5)							
	(C) Blinding of participants a			rformar	nce bias)						
35	(D) Blinding of outcome asse										
36	(E) Incomplete outcome data										
37	(F) Selective reporting (reporting (reporting)	rting bias)									
5.	(G) Other bias										

## Appendix 5: forest plot of comparison ST versus no treatment in back pain outcome at short-term with exclusion of Erdogmus C 2007 trial

		SТ		No	treatme	nt		Mean Difference	Mean Difference	Risk of Bias
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Białosky j 2014	31.75	22.45	27	31.39	26.21	28	49.3%	0.36 [-12.52, 13.24]		
Eardley S 2013	36	19.8	21	57.7	19.2	21	50.7%	-21.70 [-33.50, -9.90]		?????
Erdogmus C 2007	21.6	10.65	40	19.25	10.96	40	0.0%	2.35 [-2.39, 7.09]		? 🗣 🗬 🖶 🗣 ?
Total (95% CI)			48			49	100.0%	-10.83 [-32.44, 10.79]	•	
Heterogeneity: Tau <sup>2</sup> Test for overall effect				df = 1 (	P = 0.03	1);	84 <b>%</b>		-100 -50 0 50 1 Favours ST Favours No tr	L <b>OO</b> eatment
Risk of bias legend										

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)
 (E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

## PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Page 1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Page 5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Pages 9- 10
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Page 10
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Pages 6 and 10
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Pages 10-11
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Page 11
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Page 12
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Page 12
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Page 12- 13
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Page 12
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	Pages 12-13



## PRISMA 2009 Checklist

Synthesis of results	sis of results 14 Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I <sup>2</sup> ) for each meta-analysis.					
		Page 1 of 2	- I			
Section/topic	ection/topic # Checklist item					
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Pages 12- 13			
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Page 13			
7 Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Page 15			
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Pages 16- 19			
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Page 20, figure 2			
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Pages 21- 23			
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Pages 7-8			
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Pages 20- figures 2,3,4,5,6			
<sup>3</sup> Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Appendices 4-6			
	•					
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Pages 7-8, 24			
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Page 25			
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Page 26			
		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml				

- 46 47

Page 47 of 46

BMJ Open



# PRISMA 2009 Checklist

3 4 5 6	Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Page 2
7	From: Moher D, Liberati A, Tetzla	ff J, Altm	nan DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS M	1ed 6(6): e1000097.
9	doi:10.1371/journal.pmed1000097		For more information, visit: www.prisma-statement.org.	
10				
11			Page 2 of 2	
12				
13				
14				
15				
16				
17				
18				
19 20				
20				
22				
23				
24				
25				
26				
27				
28			For more information, visit: www.prisma-statement.org. Page 2 of 2	
29 30				
30 31				
32				
33				
34				
35				
36				
37				
38				
39 40				
40 41				
41				
43				
44				
45			For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	
46				
47				