ELSEVIER

Physiotherapy

Physiotherapy 112 (2021) 121–134

Systematic review

The effect of spinal manipulative therapy on pain relief and function in patients with chronic low back pain: an individual participant data meta-analysis

Annemarie de Zoete [a,*], Sidney M. Rubinstein [a], Michiel R. de Boer [g], Raymond Ostelo [a], Martin Underwood [b,c], Jill A. Hayden [d], Laurien M. Buffart [e], Maurits W. van Tulder [a,f], on behalf of the International IPD-SMT group:, G. Bronfort, N.E. Foster, C.G. Maher, J. Hartvigsen, P. Balthazard, F. Cecchi, M.L. Ferreira, M.R. Gudavalli, M. Haas, B. Hidalgo, M.A. Hondras, C.Y. Hsieh, K. Learman, P.W. McCarthy, T. Petersen, E. Rasmussen-Barr, E. Skillgate, Y. Verma, . Vismara, B.F. Walker, T. Xia, N. Zaproudina

[a] *Department of Health Sciences, Faculty of Science and Amsterdam Movement Science Research Institute, Vrije Universiteit, Amsterdam, The Netherlands*
[b] *Warwick Clinical Trials Unit, Warwick Medical School, The University of Warwick, Coventry CV4 7AL, UK*
[c] *University Hospitals of Coventry and Warwickshire, Coventry CV2 2DX, UK*
[d] *Department of Community Health & Epidemiology, Dalhousie University, Halifax, Nova Scotia B3H 1V7, Canada*
[e] *Radboud UMC, Nijmegen, the Netherlands*
[f] *Department Physiotherapy & Occupational Therapy, Aarhus University Hospital, Aarhus, Denmark*
[g] *Department of General Practice and Elderly Care Medicine, UMCG, the Netherlands*

Abstract

Background A 2019 review concluded that spinal manipulative therapy (SMT) results in similar benefit compared to other interventions for chronic low back pain (LBP). Compared to traditional aggregate analyses individual participant data (IPD) meta-analyses allows for a more precise estimate of the treatment effect.
Purpose To assess the effect of SMT on pain and function for chronic LBP in a IPD meta-analysis.
Data sources Electronic databases from 2000 until April 2016, and reference lists of eligible trials and related reviews.
Study selection Randomized controlled trials (RCT) examining the effect of SMT in adults with chronic LBP compared to any comparator.
Data extraction and data synthesis We contacted authors from eligible trials. Two review authors independently conducted the study selection and risk of bias. We used GRADE to assess the quality of the evidence. A one-stage mixed model analysis was conducted. Negative point estimates of the mean difference (MD) or standardized mean difference (SMD) favors SMT.

Abbreviations: IPD, individual participant data; RCT, randomized clinical trial; LBP, low back pain; SMT, spinal manipulative therapy; PRISMA-P, Preferred Reporting Items of Systematic Reviews and Meta-Analyses Protocol; MD, mean difference; SMD, standardized mean difference; SD, standard deviation; RR, relative risk; RMDQ, Roland Morris Disability Questionnaire.

* Corresponding author at: Department Health Sciences, Faculty of Science, Vrije Universiteit, De Boelelaan 1085, Room WN U-454, 1081 HV Amsterdam, The Netherlands.

E-mail addresses: a.de.zoete@vu.nl (A. de Zoete), s.m.rubinstein@vu.nl (S.M. Rubinstein), m.r.de.boer@vu.nl (M.R. de Boer), r.ostelo@vu.nl (R. Ostelo), m.underwood@warwick.ac.uk (M. Underwood), jhayden@dal.ca (J.A. Hayden), l.buffart@amsterdamumc.nl (L.M. Buffart), maurits.van.tulder@vu.nl (M.W. van Tulder).

https://doi.org/10.1016/j.physio.2021.03.006
0031-9406/© 2021 The Author(s). Published by Elsevier Ltd on behalf of Chartered Society of Physiotherapy. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Results Of the 42 RCTs fulfilling the inclusion criteria, we obtained IPD from 21 ($n = 4223$). Most trials (s = 12, $n = 2249$) compared SMT to recommended interventions. There is moderate quality evidence that SMT *vs* recommended interventions resulted in similar outcomes on pain (MD −3.0, 95%CI: −6.9 to 0.9, 10 trials, 1922 participants) and functional status at one month (SMD: −0.2, 95% CI −0.4 to 0.0, 10 trials, 1939 participants). Effects at other follow-up measurements were similar. Results for other comparisons (SMT *vs* non-recommended interventions; SMT as adjuvant therapy; mobilization *vs* manipulation) showed similar findings. SMT *vs* sham SMT analysis was not performed, because we only had data from one study. Sensitivity analyses confirmed these findings.

Limitations Only 50% of the eligible trials were included.

Conclusions Sufficient evidence suggest that SMT provides similar outcomes to recommended interventions, for pain relief and improvement of functional status. SMT would appear to be a good option for the treatment of chronic LBP.

Systematic Review Registration Number PROSPERO CRD42015025714

© 2021 The Author(s). Published by Elsevier Ltd on behalf of Chartered Society of Physiotherapy. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Spinal manipulative therapy; Individual participant data; Low back pain

Introduction

Low back pain (LBP) is the leading cause of pain and disability worldwide, and has a major socioeconomic impact [1]. Non-pharmacological approaches are the first choice of treatment as the risk of adverse events is lower than with pharmacological approaches [2]. One non-pharmacological approach includes spinal manipulation or mobilization, collectively known as spinal manipulative therapy (SMT). SMT is used by a variety of heath care providers such as chiropractors, osteopaths, manual therapists and physiotherapists.

Many systematic reviews and meta-analyses have analysed the effects of SMT and suggest that it is an effective intervention for the reduction of pain and improvement of function [3–5]. However, recommendations for SMT in international guidelines for chronic LBP are not consistent [6–8]. Since each guideline development group is using the same evidence, this is likely to be a consequence of differences in how groups approach appraisal and interpretation of the evidence.

One disadvantage of traditional meta-analyses, is that aggregate data are extracted at the study-level and the investigator is dependent upon how the data is analysed and presented. Individual participant data (IPD) meta-analysis circumvents the issues of poor reporting and not correcting for baseline covariates, because the individual data are available, resulting in more precise and potentially, a more valid estimate of the effect.

Our recent systematic review for SMT for chronic LBP [5] reflects some of the potential limitations of traditional aggregate meta-analysis. For example, the authors of included studies used different definitions of LBP, included a few subacute LBP patients, used different frequencies of treatments, and different analytic techniques ranging from a t-test to sophisticated regression models. In an IPD meta-analysis, some of these problems can be resolved.

The specific objective of this IPD meta-analysis was to assess the effectiveness of SMT compared to any other conservative therapy for primary outcomes (i.e. pain and back-related function) and secondary outcomes (i.e. quality of life, recovery, return-to-work, medication use and treatment satisfaction) at one, three, six and twelve months in adults with chronic LBP.

Methods

This study was conducted according to the Preferred Reporting Items of Systematic Reviews and Meta-Analyses for IPD (PRISMA-IPD) guidelines [9] (Appendix eTable 1). The protocol was registered with PROSPERO (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=25714) and approved by the Scientific Review Board of the Vrije Universiteit Amsterdam and by the Ethical Committee of the VU University Medical Centre Amsterdam (Projectnr. 2015.544).

A detailed description of our study design and procedures was published previously [10]. The methodology presented here gives a brief overview.

Data sources and searches

Search methods for identification of new studies

We included RCTs published from the year 2000. We limited this inclusion, because it is difficult to trace authors of older trials, and there is a high probability that these data would not be accessible. More importantly, more recent studies of SMT for low-back pain are of better methodological quality. Therefore, it is unlikely that this delineator will have introduced undesirable bias [11]. Studies in the 2011 Cochrane Review which examined the effect of SMT for chronic LBP were included [12]. In addition, we updated the search in December 2016 following the same procedure used in the Cochrane review (Appendix eTable 10) [5,10]. This was supplemented with reference checking of systematic reviews and meta-analyses, and personal communication. A recent update of the search (May 2018) resulted in the identification of five new trials [5], all of which were small in size and considered to have a high risk of bias. A update search from May 2018 until October 2020 identified five studies [13–17],

three of which are small in size. The two large-sized studies of which one examined SMT vs recommended therapies and the other SMT as adjuvant therapy, reported similar results to ours.

Study selection

Type of studies and participants

Inclusion criteria. Only randomized clinical trials (RCTs) were included. Studies were included if they recruited adults (≥18 years of age) with chronic (≥12 weeks duration) LBP. LBP is defined as LBP not attributed to a specific pathology (e.g. infection fracture, tumour or radicular syndrome). Participants with diffuse leg pain due to a low-back condition were included as were participants from primary or secondary care. In those studies where a mixed population was included (e.g. subacute and chronic), where possible, we included only those participants with >12 weeks of LBP.

Exclusion criteria. We excluded studies that: 1) used an inadequate randomization procedure (e.g. alternate allocation, allocation based on birth date); 2) included participants with LBP and other conditions such as pregnancy or post-operative participants; 3) tested the immediate effect of a single treatment only; and 4) compared the effects of a multimodal therapy including SMT to another therapy or any other study design whereby the contribution of SMT could not be isolated.

Types of interventions

Experimental intervention. Studies of spinal manipulation (i.e. high-velocity low-amplitude techniques) as well as mobilization (i.e. low-velocity low-amplitude techniques) were included.

Comparisons. We analysed the following comparisons: 1) SMT *vs* recommended interventions including non-drug treatment (e.g. exercise), and drug treatment (e.g. NSAIDs); 2) SMT *vs* non-recommended interventions (e.g. diathermy), 3) sham 'placebo' SMT; 4) SMT + intervention *vs* intervention alone; 5) high-velocity low-amplitude SMT *vs* low-velocity low-amplitude SMT (i.e. manipulation *vs* mobilization).

We based the definition of 'recommended' and 'non-recommended' interventions on recent international guidelines for LBP from the USA [8], the UK [6], the Netherlands [7] and COST B13 European guidelines [18]. We categorized an intervention into 'recommended' or 'non-recommended' when this was consistently stated in at least two of these guidelines.

Types of outcome measures

Primary outcomes were self-reported pain and back-specific functional status.

Secondary outcomes included self-reported health-related quality of life, return-to-work, global improvement (i.e. perceived recovery), treatment satisfaction and analgesic use.

Data extraction and quality of assessment

Risk of bias in individual studies

The 13 risk of bias criteria (scored as 'low risk', 'high risk' or 'unclear risk') recommended by the Cochrane Back and Neck group were used (Appendix eTable 2) [19]. The risk of bias was conducted by two independent reviewers (SMR, AdeZ). To adjudicate disagreement, a third reviewer (RO) was contacted.

Data of all participants was sought from the authors of the studies fulfilling the inclusion criteria. We extracted study characteristics, patient characteristics, types of outcomes, duration of follow-up, description of experimental and control interventions.

Preparing data for analyses

We first compared the original data with the published data to check for completeness and where necessary and possible, attempted to resolve any discrepancies. All variables were harmonized in a data harmonization platform [10].

All outcomes were pooled following a decision rule (Appendix eTable 5). All pain scores were converted to a 0–100 points pain scale. To allow pooling of different functional status measures, we recoded the individual scores into Z-scores for each separate time point using pooled standard deviations as denominator ($Z\ score = \frac{xi-\bar{x}}{SD}$). Analysing these Z-scores resulted in standardized mean differences (SMD's). To ease interpretation of SMD's, we converted these to a mean difference (MD) for the 24 point Roland Morris Disability Questionnaire (RMDQ), by multiplying the SMD with the population standard deviation (SD) of the studies measuring RMDQ ($SD_{pooled} = \sqrt{\sum_{i=0}^{n} \frac{(n_i-1)*S^2}{(n_i-1)}}$ n_i = sample size for each trial; S = standard deviation for each trial).

For quality of life, physical and mental component scales of SF12 and SF36 were combined.

Other secondary outcomes were all dichotomized (Appendix eTable 5). However, data were often insufficient (less than 3 trials) to perform any analyses for these outcomes. Adverse events were not included in our protocol but we did examine these data.

Data synthesis and analysis

All analyses were based on the intention-to-treat (ITT) principle. Our primary analyses consisted of one-stage IPD meta-analyses for the five main comparisons at one, three, six and twelve months follow-up (see protocol) [10]. These chosen intervals are standard follow-up moments for treatment in LBP. We did not examine the effect of SMT directly

post-intervention as there was a large variation in duration and frequency of treatments among the studies. Furthermore, many studies contained no follow-up data immediately following the end of treatment. Longitudinal analyses for all time points simultaneously were not performed as the models were deemed too computationally demanding.

Analyses were conducted using a random-effects analysis of covariance model adjusting for baseline outcome using REML (restricted maximum likelihood), where a separated intercept and separate residual variance for each study is specified. Models extended with a separate baseline adjustment term per trial did not demonstrate convergence in most analyses and we omitted them from all analyses [20].

The pooled treatment effect of SMT was estimated using an MD or SMD (for continuous outcomes) or as an odds ratio (for dichotomous outcomes) including the 95% CI. Negative MD for pain and SMD for function favours SMT, while positive MD for quality of life favours SMT.

We did not assess the effects of imputing missing data on outcomes. We addressed the missing outcome data (see results: characteristics of studies).

Subgroup and sensitivity analyses

Subgroup analyses were pre-specified in our protocol [10] and conducted for the following variables: 1) type of clinician (i.e. chiropractor *vs* other); 2) 'multi-modal' SMT (i.e. SMT delivered alone as opposed to in conjunction with other modalities have limiting or no effect); 3) country where the study was conducted (USA *vs* other); 4) only chronic LBP participants (some trials included participants with subacute LBP), and 5) only trials with exercise therapy as a comparator. We conducted sensitivity analyses for studies: 1) with low risk of bias on random sequence generation and allocation concealment, 2) with overall low risk of bias (defined as fulfilling six or more of the criteria items); 3) with a follow-up period of eight weeks (data from eight weeks follow-up analysed with the three months instead of one month and 4) where we were able to reproduce published results.

Furthermore, sensitivity analyses were performed by calculating functional status scores ourselves instead of using the received overall score. Also, we examined the different functional status measures (e.g. RMDQ) and pain scales (e.g. average pain, pain intensity), separately.

Lastly, to examine whether the RCTs included in this IPD meta-analysis were a representative sample of all known RCTs published since 2000, we conducted a two-stage sensitivity analysis wherein we examined the effect sizes of RCTs included in this IPD meta-analysis *vs* those which were eligible for inclusion, but for which no IPD was available (using published aggregate data) [5].

Assessment of clinical relevance was defined as a small, medium or large effect and based on the recommendations of the Cochrane Back and Neck group [21,22]. The overall quality of the evidence for each outcome was evaluated using GRADE [23] adapted for IPD (see Appendix eTable 6).

Results

Identification of trials

In total, 43 RCTs met our inclusion criteria, of which 21 (50%) provided data [24–44] (Fig. 1) representing 4223 participants. In three trials, the results differed from the published results for the primary outcomes by more than five percent (i.e. 5 points on a 0–100 point Visual analogue scale, and 1.2 points for the 0–24 RMDQ), which we determined to be a relevant difference. Of these, one trial provided only data from participants who gave consent to share their data [40]. For another study, we received data for more participants and longer follow-up than published [38] while for the third study, baseline data were very similar but our results of the analyses deviated somewhat from the published results due to different patient numbers and use of different statistical techniques; This was a small trial ($n = 41$) and therefore, these deviations were not likely to influence the results presented here [41].

Characteristics of studies

Of these 21 RCT's, 12 evaluated the effect of SMT *vs* recommended interventions of which eight were compared to exercise therapy [26,27,29,30,32–34,38,40–43], one evaluated the effects of SMT *vs* sham SMT [44], five evaluated the effect of SMT *vs* non-recommended interventions [24,31,33,37,39], five evaluated the effect of SMT as an adjuvant therapy [25,33,35,36,43] and three evaluated the effect of manipulation *vs* mobilization [28,32,39] (Appendix eTable 3).

Sample sizes ranged from 21 to 1334 (median = 192; IQR = 45–271). However, some trials included multiple arms, and some included non-chronic LBP patients; Therefore, the sample size for a given comparison should be considered potentially smaller. The included trials varied with respect to the recruitment method, type of SMT technique, number and duration of treatments and type of practitioner (Appendix eTable 3).

Of the 4223 participants, 2249 were randomized to the SMT group and 1974 to the comparison group. Table 1 presented the patient characteristics at baseline for SMT *vs* recommended interventions. Data for the other comparisons are tabulated in eTable 7 (Appendix).

Missing data for primary outcomes ranged from 11% at one month to 21% at 12 months. The UK BEAM trial provided the largest dataset ($n = 1334$) and as a result, contributed most to the missing outcome data (50% of the total amount). The UK BEAM authors did not find a difference across randomized groups between responders and non-responders and drop-out appeared to be unrelated to the treatment [43].

Risk of bias

Approximately three quarters of the studies ($n = 15$) reported an adequate random sequence generation

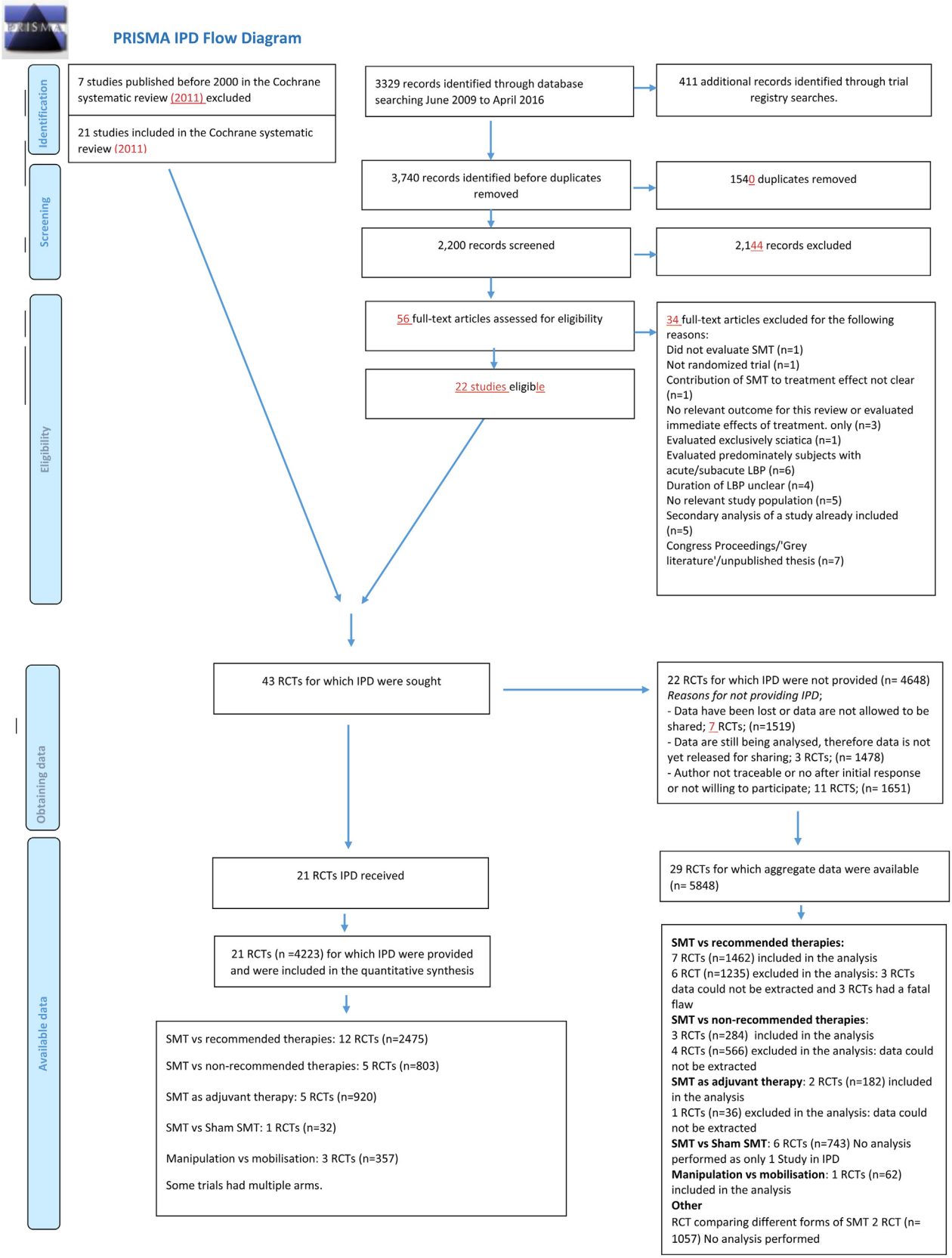

Fig. 1. Flow diagram of study inclusion.

Table 1
Patient characteristics at baseline for groups receiving SMT *vs* groups receiving recommended interventions (*s* = 12; *n* = 2475).

Demographic data	SMT	Recommended interventions
Age, mean (SD) years (s = 11, *n* = 2409)	47 (14)	47 (14)
Sex, *n* (%) female (s = 11, *n* = 2412)	667 (57)	684 (55)
Body mass index, mean (SD) (s = 8, *n* = 1434)	27 (5)	27 (5)
Ethnicity, *n* (%) white (s = 5, *n* = 861)	409 (91)	388 (88)
Lifestyle factors		
Physical activity, (%) (s = 6, *n* = 824)		
Low (1 or less than once a week)	115 (32)	166 (36)
Medium (2–3× a week)	146 (40)	166 (36)
High (more than 3× a week)	100 (28)	131 (28)
Smoking, *n* (%) non-smokers (s = 6, *n* = 1173)	451 (80)	453 (75)
Alcohol use, *n* (%)	*	*
Socio-demographics		
Marital status, *n* (%) married; living with a partner (s = 6, *n* = 1173)	397 (69)	404 (68)
Level of education, *n* (%) low/middle (s = 7, *n* = 1672)	600 (68)	534 (68)
Income, *n* (%)	*	*
Employment status, *n* (%) at work (s = 9, *n* = 2126)	818 (78)	770 (72)
Nature and severity of LBP		
Duration of LBP, *n* (%) less than 12 months (s = 7, *n* = 1252)	121 (20)	149 (23)
Leg pain, *n* (%) (s = 5, *n* = 1038)	320 (59)	281 (57)
Previous LBP treatment received, *n* (%) (s = 5, *n* = 930)	258 (28)	218 (23)
Previous physiotherapy for low back pain received, *n* (%) (s = 5, *n* = 771)	64 (8)	72 (9)
Previous SMT for low back pain received, *n* (%) (s = 6, *n* = 988)	209 (21)	111 (11)
Used medication for low back, *n* (%) (s = 6, *n* = 1018)	200 (20)	269 (26)
Non-specific, *n* (%)	*	*
Comorbidities	*	*
Type of treatment	*	*
Psychosocial factors	SMT	Control
Depression, *n* (%) (s = 5, *n* = 1297)	43 (6)	75 (13)
Treatment preference/expectations	*	*
Primary outcomes		
Pain		
Combined pain score at baseline, mean (SD), (s = 12, *n* = 2441)	49.5 (22.3)	49.8 (21.6)
Combined pain score at one month, mean (SD), (s = 10, *n* = 1948)	34.2 (23.0)	35.8 (23.9)
Combined pain score at three months, mean (SD), (s = 9, *n* = 1673)	27.92 (23.0)	32.1 (24.3)
Combined pain score at six months, mean (SD), (s = 8, *n* = 1321)	27.35 (23.1)	32.3 (23.9)
Combined pain score at twelve months, mean (SD), (*s* = 10, *n* = 1816)	31.80 (26.8)	33.3 (25.4)
Functional status		
RMDQ sum score at baseline, mean (SD), (s = 9, *n* = 2174)	9.0 (5.0)	10.1 (5.4)
RMDQ sum score at one month, mean (SD), (s = 8, *n* = 1760)	5.6 (5.0)	6.7 (5.4)
RMDQ sum score at three months, mean (SD), (s = 8, *n* = 1648)	4.8 (5.1)	5.5 (5.3)
RMDQ sum score at six months, mean (SD), (s = 8, *n* = 1348)	5.0 (5.4)	6.3 (6.0)
RMDQ sum score at twelve months, mean (SD), (s = 7, *n* = 1575)	5.4 (5.7)	6.2 (5.92)
Secondary outcomes		
SF36 Physical Component Scale of SF36 at baseline, mean (SD), (s = 5, *n* = 1362)	40.7 (7.2)	41.1 (7.6)
SF36 Physical Component Scale of SF36 at one month, mean (SD), (s = 3, *n* = 865)	44.1 (7.9)	45.7 (8.1)
SF36 Physical Component Scale of SF36 at three months weeks, mean (SD), (s = 4, *n* = 1154)	46.7 (8.2)	46.9 (8.5)
SF36 Physical Component Scale of SF36 at six months, mean (SD), (s = 5, *n* = 839)	47.3 (7.8)	47.9 (7.7)
SF36 Physical Component Scale of SF36 at twelve months, mean (SD), (s = 5, *n* = 1249)	46.4 (8.6)	46.8 (8.8)
SF36 Mental Component Scale of SF36 at baseline, mean (SD), (s = 5, *n* = 1362)	43.8 (9.1)	45.1 (9.6)
SF36 Mental Component Scale of SF36 at one month, mean (SD), (s = 3, *n* = 865)	45.5 (8.8)	46.9 (8.8)
SF36 Mental Component Scale of SF36 at three months, mean (SD), (s = 4, *n* = 1154)	46.9 (9.0)	47.3 (9.5)
SF36 Mental Component Scale of SF36 at six months, mean (SD), (s = 5, *n* = 839)	46.9 (9.1)	48.0 (8.9)
SF36 Mental Component Scale of SF36 at twelve months, mean (SD), (*s* = 5, *n* = 1249)	45.7 (8.9)	46.2 (10.0)
Medication use at baseline, *n* (% medication use) (s = 3, *n* = 668)	145 (22)	216 (32)
Medication use at one month, *n* (% medication use) (s = 3, *n* = 646)	84 (13)	146 (23)
Medication use at three months, *n* (% medication use) (s = 3, *n* = 626)	78 (13)	132 (21)
Medication use at six months, *n* (% medication use) s = 3, *n* = 593)	67 (11)	143 (24)
Medication use at twelve months weeks, *n* (% medication use) (s = 3, *n* = 582)	82 (14)	141 (24)

SD = standard deviation; s = number of studies; *n* = number of participants; * combining categories was not meaningful or no data available.

and allocation concealment (Appendix eTable 4) [24–27,29–32,34,37–40,42,43]. Fifteen trials provided an adequate overview of withdrawals or drop-outs and were able to keep these to a minimum for the subsequent follow-up measurements [24–29,31,33,34,36–40,44].

Effect of SMT on primary and secondary outcomes: one stage meta-analysis

Negative point estimates of the mean difference (MD) or standardized mean difference (SMD) favours SMT.

1) ***SMT vs recommended Interventions***

Pain and function improved by the end of treatment and this improvement was sustained up to twelve months after randomization for all groups (Appendix eFigs. 3 and 4).

Primary outcomes

Pain. There is moderate quality evidence that SMT has similar benefit to recommended interventions at all time points (largest difference at three months; Table 2).

Functional status. There is moderate quality evidence that SMT has similar benefit to recommended interventions at all time points (largest difference at one month; Table 2).

A subgroup analysis for SMT *vs* exercise showed similar results (see Appendix eTable 8).

Secondary outcomes

There is moderate quality evidence that SMT results in a medium reduction in medication use compared to recommended interventions at two of the four time points (largest difference at six months. For all other secondary analyses, there is low to high quality evidence that SMT has a similar benefit to recommended interventions (Table 3).

2) ***SMT vs non-recommended interventions***

Primary outcomes

Pain. There is moderate quality evidence that SMT has similar benefit compared to non-recommended interventions at one and six months (largest difference at six months). There are insufficient data for the three and twelve months analyses (Table 2).

Functional status: There is moderate quality evidence that SMT has similar benefit compared to non-recommended interventions at one, three, and six months (largest difference at six months). There are insufficient data for the twelve months analysis (Table 2).

Secondary outcomes

Quality of life

There is low quality evidence that SMT has a similar benefit to non-recommended interventions at one and six months (largest difference at six months). There are insufficient data for the three and twelve months analyses (Table 3).

3) ***SMT vs Sham SMT***

The analysis for this comparison was not performed, because we only had data from one study [44].

4) ***SMT + intervention vs intervention alone***

Primary outcomes

Pain. There is moderate quality evidence that SMT + intervention has a similar benefit compared to intervention alone at one, three and twelve months and low quality evidence that SMT has a similar benefit to intervention alone at six months (largest difference at one month) (Table 2).

Functional status. There is moderate quality evidence that SMT + intervention has similar benefit compared to intervention alone at one, three and twelve months and low quality evidence that SMT + intervention has similar benefit compared to the intervention alone at six months (largest difference at three months) (Table 2).

Secondary outcomes

Quality of life

There is moderate quality evidence that SMT + intervention has similar benefit compared to the intervention alone at one, three and twelve months and low quality evidence that SMT + intervention has similar benefit to the intervention alone at 6 months (largest difference at twelve months) (Table 3).

5) ***Manipulation vs mobilization***

Pain. There is moderate quality evidence that manipulation has a similar benefit compared to mobilization at one month (Table 2).

Functional status. There is moderate quality evidence that manipulation has a similar benefit compared to mobilization at one month (Table 2).

There are no data for the other time points and secondary outcomes.

Subgroup and sensitivity analyses

The results from all one-stage sensitivity analyses suggest similar results for pain and functional status at all time points (Appendix eTable 8).

We found no differences in pain and functional status between RCTs included and eleven eligible RCTs not included in the IPD repository (Table 4 and Appendix eTable 9). The results of the two-stage analysis were comparable with the one-stage analysis. Sensitivity analysis, including studies published since 2016, did not change our results.

Discussion

Our results suggest there is moderate quality evidence that SMT has similar effects as recommended treatments for pain reduction and improved functional status at short-, intermediate- and long-term follow-up. Additionally, there is moderate evidence that SMT has similar effects for pain relief and improvement in function when compared to non-recommended therapies and when examined as an adjuvant

Table 2
Main treatment effects and GRADE summary of findings for all comparisons for the primary outcomes. Regression coefficients (β) and 95% confidence intervals (CI) of the intervention effects of random-effect models adjusted for baseline using REML (one stage analysis) are presented.

Time measurement	Difference in effect (β) (95% CI)	# studies	N	Quality of the evidence (and reason for downgrading)	Comments
Comparison 1: SMT *vs* recommended therapies					
Outcome: pain[a]					
1 month	MD −3.0, 95% CI −6.9 to 0.9	10	1922	Moderate (inconsistency)	
3 months	MD −6.6, 95% CI −13.0 to −0.2	9	1647	Moderate (inconsistency)	
6 months	MD −5.6, 95% CI −9.6 to −1.5	8	1321	Moderate (inconsistency)	
12 months	MD −2.5, 95% CI −7.1 to 2.1	10	1791	Moderate (inconsistency)	
Outcome: functional status					SMD converted to a MD on the 24 point RMDQ scale
1 month	SMD −0.2, 95% CI −0.4 to 0.0	10	1939	Moderate (inconsistency)	−0.8
3 months	SMD −0.1, 95% CI −0.4 to 0.1	11	1892	Moderate (inconsistency)	−0.6
6 months	SMD −0.2, 95% CI −0.3 to 0.0	9	1490	Moderate (inconsistency)	−0.8
12 months	SMD −0.1, 95% CI −0.3 to 0.1	10	1826	Moderate (inconsistency)	−0.5
Comparison 2: SMT *vs* non-recommended therapies					
Outcome: pain[a]					
1 month	MD −6.6 95% CI −10. 8 to −2.3	5	755	Moderate (inconsistency)	
3 months	Not enough data				
6 months	MD −8.3, 95% CI −20.5 to 3.8	3	419	Moderate (imprecision)	
12 months	Not enough data				
Outcome: functional status					SMD converted to a MD on the 24 point RMDQ scale
1 month	SMD −0.3, 95% CI −0.6 to 0.0	5	835	Moderate (inconsistency)	−0.6
3 months	SMD −0.0, 95% CI −0.8 to 0.7	2	375	Moderate (inconsistency)	−0.9[c]
6 months	SMD −0.3, 95% CI −0.6 to −0.0	3	414	Moderate (imprecision)	−0.9
12 months	Not enough data				
Comparison 3: SMT *vs* sham SMT					
No results only data of one study					
Comparison 4: SMT + intervention *vs* intervention alone					
Outcome: pain[a]					
1 month	MD −7.44, 95% CI −12.7 to −2.1	5	762	Moderate (inconsistency)	
3 months	MD −5.2, 95% CI −11.0 to 0.7	2	619	Moderate (inconsistency)	
6 months	MD −1.4, 95% CI −6.7 to 3.8	2	222	Low (inconsistency, imprecision)	
12 months	MD −2.2, 95% CI −5.9 to 1.4	2	603	Moderate (inconsistency)	
Outcome: functional status[b]					
1 month	MD −0.6, 95% CI −2.3 to 1.1 [a]	4	746	Moderate (inconsistency)	
3 months	MD −1.3, 95% CI −2.6 to −0.1[a]	3	681	Moderate (inconsistency)	
6 months	MD −1.0, 95% CI −2.1 to 0.1[a]	2	218	Low (imprecision, inconsistency)	
12 months	MD −0.9, 95% CI −1.6 to −0.2[a]	2	626	Moderate (inconsistency)	
Comparison 5: Manipulation *vs* mobilization					
Outcome: pain					
1 month	MD −1.5, 95% CI −6.8 to 3.9	3	321	Moderate (limitations)	
Not enough data for other time points					
Outcome: Functional status					SMD converted to a MD on the 24 point RMDQ scale
1 month	SMD 0.0, 95% CI −0.0 to 0.1	3	356	Moderate (limitations)	−0.6
Not enough data for other time points					

Negative difference in effect indicates higher estimated decrease in pain or improvements in function for SMT group compared to the control.
MD = mean difference of combined pain score on a 0–100 scale.
SMD = standardized mean difference of combined functional status score.
[a] Pain measured on a 0–100 point scale.
[b] All studies in the SMT + intervention *vs* intervention alone measured Roland Morris Disability questionnaire, therefore we use a mean difference.
[c] Based on one small study.

Table 3
Main treatment effects and GRADE summary of findings for all secondary outcomes.

Comparison 1: SMT *vs* recommended therapies

Time measurement	Difference in effect (95% CI)	# studies	N	Quality of the evidence (and reason for downgrading)
Outcome: Quality of life: Physical Component Scale of SF36 and SF12 combined				
1 month	MD −0.6, 95% CI −1.4 to 0.1	4	844	High
3 months	MD −0.2, 95% CI −1.0 to 0.7	3	967	High
6 months	MD −0.3, 95% CI −1.5 to 0.91	4	688	High
12 months	MD 0.1, 95% CI −0.8 to 1.0	4	1055	High
Outcome: Mental Component Scale of SF36 and SF12 combined				
1 month	MD 0.4, 95% CI −0.4 to 1.2	4	844	High
3 months	MD 0.8, 95% CI −0.0 to 1.6	3	967	High
6 months	MD −0.1, 95% CI −1.4 to 1.2	4	688	High
12 months	MD 0.5, 95% CI −0.9to 2.0	4	1055	High
Outcome: Recovery[a]: Yes *vs* No				
1 month	OR 1.3, 95% CI 0.9 to 1.9	2	499	Moderate (inconsistency)
3 months	OR 1.2, 95% CI 0.8 to 1.8	3	538	Moderate (inconsistency)
6 months	OR 1.1, 95% CI 0.8 to 1.6	3	651	Moderate (inconsistency)
12 months	OR 0.8, 95% CI 0.5 to 1.2	2	445	Moderate (inconsistency)
Outcome: Medication Use[a]: Yes vs No				
1 month	OR 0.7, 95% CI 0.5 to 1.0	3	646	Moderate (inconsistency)
3 months	OR 0.7, 95% CI 0.4 to 1.2	3	626	Moderate (inconsistency)
6 months	OR 0.5, 95% CI 0.3 to 0.9	3	593	Moderate (inconsistency)
12 months	OR 0.7, 95% CI 0.3 to 1.3	3	582	Moderate (inconsistency)
Outcome: Return to work[a]: Yes vs No				
1 month	Not enough data			
3 months	OR 1.0, 95% CI 0.5 to 1.9	3	190	Low (inconsistency, imprecision)
6 months	OR 0.6, 95% CI 0.3 to 1.3	3	189	Low (inconsistency, imprecision)
12 months	OR 1.3, 95% CI 0.6 to 2.7	3	180	Low (inconsistency, imprecision)
Outcome: Satisfaction[a]: Yes vs No				
1 month	OR 0. 8, 95% CI 0.4 to 1.6)	2	319	Low (inconsistency, imprecision)
3 months	OR 6.6, 95% CI 1.5 to 29.9	2	429	Low (inconsistency, imprecision)
6 months	Not enough data			
12 months	Not enough data			

Comparison 2: SMT *vs* non-recommended therapies

Outcome: Quality of life: Physical Component Scale of SF36 and SF12 combined				
1 month	MD 0.4, 95% CI −0.8 to 1.6	3	345	Low (inconsistency, imprecision)
3 months	Not enough data			
6 months	MD −0.1, 95% CI −2.5 to 2.2	2	202	Low (inconsistency, imprecision)
12 months	Not enough data			
Outcome: Quality of life: Mental Component Scale of SF36 and SF12 combined				
1 month	MD 1.0, 95% CI −3.1; 5.2	3	345	Low (inconsistency, imprecision)
3 months	Not enough data			
6 months	MD 1.7, 95% CI −0.6; 4.0	2	202	Low (inconsistency, imprecision)
12 months	Not enough data			
Other outcomes not enough data				

Comparison 3: SMT *vs* sham SMT

No results only data of one study

Comparison 4: SMT + intervention *vs* intervention alone

Outcome: Quality of life: Physical Component Scale of SF36 and SF12 combined				
1 month	MD 0.1, 95% CI −1.1 to 1.5	3	708	Moderate (inconsistency)
3 months	MD 0.3, 95% CI −0.7 to 1.3	2	619	Moderate (inconsistency)

Table 3 (*Continued*)

Comparison 1: SMT *vs* recommended therapies				
Time measurement	Difference in effect (95% CI)	# studies	N	Quality of the evidence (and reason for downgrading)
6 months	MD −1.4, 95% CI −2.9 to 0.1	2	221	Low (inconsistency, imprecision)
12 months	MD −2.2, 95% CI −5.9 to 1.4	2	603	Moderate (inconsistency)
Outcome: Quality of life: Mental Component Scale of SF36 and SF12 combined				
1 month	MD −0.2, 95% CI −1.6 to 1.2	3	708	Moderate (inconsistency)
3 months	MD 1.9, 95% CI −0.7 to 4.5	2	619	Moderate (inconsistency)
6 months	MD 1.8, 95% CI −0.1 to 3.7	2	221	Low (inconsistency, imprecision)
12 months	MD 1.0, 95% CI −0.3 to 2.2	2	605	Moderate (inconsistency)
Other outcomes not enough data				
Comparison 5: Manipulation *vs* mobilization				
All outcomes not enough data				

Positive difference in effect indicates higher increased quality of health for SMT group compared to the control.
MD = mean difference.
OR = odds ratio.
[a] Recovery was classified as 'recovered' if the participant scored more than 50% improvement or were (much) better or had no symptoms. Medication use was classified for those using taking any using medication for LBP, while not taking any medication was classified as no medication use. Return to work was classified as participants had returned to work or if there were no sick days recorded. Satisfaction was classified as 'satisfied with care' if participants were (completely) satisfied or had scores >75%.

therapy. We have no results for the SMT *vs* sham comparison, because we could only include one study. Finally, there is moderate quality evidence that manipulation has similar effects as mobilisation.

Our results are consistent with the recently published aggregate data review [5] and with other recently published systematic reviews [4,45,46].

It is somewhat difficult to interpret these findings, particularly when SMT demonstrates similar effects to recommended and non-recommended therapies or when examined as an adjuvant therapy. This appears confusing and requires explanation. Firstly, most studies we identified examined the effect of SMT *vs* recommended therapies. In general, these studies were larger, had more data on follow-up time-points and were of better methodological quality (i.e. low risk of bias) than the studies in the other comparisons. Meaning, these findings were more robust and therefore, we have more confidence in their effect estimate. Even though for all these comparisons, there is generally moderate quality evidence according to GRADE. While there are general guidelines for applying GRADE, there is no consensus. For example, we used a general rule-of-thumb when evaluating 'imprecision' in accordance with what might be considered an 'optimal information size. Applying a more stringent optimal information size would result in lower quality evidence for SMT *vs* non-recommended therapies or SMT as an adjuvant therapy, but not when applying this criterion to SMT *vs* recommended therapies (because the latter analyses included more than 1000 subjects). Secondly, categorizing interventions into recommended or non-recommended interventions was not always straightforward (e.g. myofascial therapy), and therefore, open for interpretation. While a sensitivity analysis could have helped to resolve this issue, the data were not sufficiently robust to make this possible. Lastly, the categorization of an intervention as 'non-recommended' does not imply that these interventions do not have an effect or are dangerous or ill-advised. While trials whereby patients are 'blinded' (i.e. sham) would help to resolve this issue; in our estimation, no single study was adequately able to do so. An important difference of our IPD analysis compared to traditional aggregate meta-analyses is that we could adjust for the covariates, baseline pain and functional status, and were not dependent upon how these data were reported in the original publications. This has increased precision of our estimates compared to aggregate data meta analyses, but did not lead to a different conclusion for the main effects.

It will be difficult to justify the required financial and participant resources for further trials comparing SMT *vs* current recommended therapies, as this is unlikely to change our overall conclusions. Others have previously made the same observation with regard to trials of exercise treatment for low back pain. A 2019 IPD meta-analysis of exercise therapy for low back pain has also produced precise estimates for effectiveness [47]. Therefore, future studies should focus on cost-effectiveness, optimal dosage, delivery route to minimize side-effects, specificity of the location treated and maximize the non-specific effects of care, instead of reproducing the same type of trials.

Table 4
Representativeness of the pooled effects of studies providing data for the IPD study and those not providing data. Two stage analysis; SMT *vs* recommended therapies.

Representativeness	Number of studies	Difference in effect (CI 95%)	Test of heterogeneity		Prediction Interval
Outcome			I^2	*P*-value	
Pain		Mean difference			
Combined pain score one month					
All eligible studies	16	−2.4 (−4.9; 0.1)	66%	<0.001	−2.4 (−11.6; 6.9)
Studies providing data	10	−2.5 (−5.9; 0.9)	75%	<0.001	−2.7 (−13.8;8.8)
Studies not providing data	6	−2.2 (−6.1; 1.7)	55%	0.02	−2.218 (−12.9; 8.5)
Combined pain score three months					
All eligible studies	14	−3.1 (−7.0; 0.8)	80%	<0.001	−3.1 (−18.6; 12.4)
Studies providing data	9	−5.9 (−11.4; −0.5)	86%	<0.001	−5.9 (−25.0; 13.0)
Studies not providing data	5	1.1 (−2.7; 4.9)	29%	0.196	1.1 (−7.4; 9.6)
Combined pain score six months					
All eligible studies	13	−4.1 (−6.7; −1.6)	66%	<0.001	−4.1 (−13.5; 5.2)
Studies providing data	8	−4.8 (−8.9; −0.6)	72%	0.001	−4.8 (−17.8;8.3)
Studies not providing data	5	−3.5 (−7.0; −0.1)	61%	0.009	−3.5 (−13.8; −6.4)
Combined pain score twelve months					
All eligible studies	12	−1.8 (−4.8; 1.3)	71%	<0.001	−1.8 (−12.5; 9.0)
Studies providing data	10	−2.1 (−6.5; 2.2)	79%	<0.001	−2.1 (−16.8; 12.5)
Studies not providing data	2	−0.9 (−3.5; 1.8)	0%	0.6	−0.9 (−6.7; 5.0)
Functional status		**Difference in Z-score**			
Combined Functional Status one month					
All eligible studies	13	−0.2 (−0.3; −0.0)	49%	0.014	−0.2 (−0.6; 0.3)
Studies providing data	9	−0.3 (−0.5; −0.0)	65%	0.003	−0.3 (−1.0; 0.4)
Studies not providing data	4	−0.1 (−0.2; 0.0)	0%	0.739	−0.1 (−0.2; 0.1)
Combined Functional Status three months					
All eligible studies	15	−0.1 (−0.2; 0.1)	75%	<0.001	−0.1 (−0.8; 0.7)
Studies providing data	11	−0.2 (−0.4; 0.0)	79%	<0.001	−0.2 (−0.9; 0.6)
Studies not providing data	4	0.2 (−0.1; 0.4)	45%	0.089	0.2 (−0.5; 0.8)
Combined Functional Status six months					
All eligible studies	13	−0.1 (−0.2; 0.0)	56%	0.002	−0.1 (−0.6; 0.4)
Studies providing data	9	−0.2 (−0.4; −0.0)	69%	0.001	−0.2 (−0.9; 0.5)
Studies not providing data	4	0.0 (−0.1; 0.1)	0%	0.778	0.0 (−0.1; 0.2)
Combined Functional Status twelve months					
All eligible studies	13	−0.2 (−0.3; 0.1)	72%	<0.001	−0.2 (−0.7; 0.4)
Studies providing data	10	−0.2 (−0.4; 0.1)	79%	<0.001	−0.2 (−0.9; 0.6)
Studies not providing data	3	−0.1 (−0.3; 0.1)	43%	0.132	−0.1 (−0.7; 0.5)

CI = confidence interval; $I^2 = I^2$ statistic, which is the percentage of total variance that can be explained by heterogeneity, and 25% is considered low, 50% moderate, and 75% high heterogeneity.

Strengths and limitations

The most important strength is the sample size and the diversity of studies, meaning these results are likely to be broadly generalizable to clinical practice. On the other hand, this diversity in studies led to a difference in methodological quality and types of outcomes and covariates across trials and introduced statistical heterogeneity, which can introduce difficulties in interpreting the data. We investigated this diversity with sensitivity analyses and two-stage analyses for primary outcomes at all follow-up measurements, but could not explain the statistical heterogeneity. Our understanding of the effects of SMT would improve if we had an understanding of the aetiology of LBP and how SMT works.

The most important limitation is potential selection bias. We included only 50% of the eligible trials, which is comparable to other IPD studies [47,48]. In a two-stage analysis we examined the effect sizes of those that were eligible but did not provide data. Results suggest only small differences between included studies and those for which we were unable to source the original data, indicating that the RCTs included are likely to be a representative sample of all published studies. Also, the range of studies based upon publication date and methodological quality of the studies we included is comparable with the non-included studies in the recently published review [5]. Therefore, this facilitates an effective comparison of interventions across trials.

Also, our review differs slightly from our protocol with regard to the classification of the comparator. In our protocol, we classified therapies into effective and non-effective, whereas in this review we classified them into recommended and non-recommended therapies. It was thought this would best help translation of findings to clinical practice. This did not affect the reported result, but was more a wording difference. Finally, longitudinal analyses would have provided us with more information on the individual pattern of changes over time. In the future, these can be run, when programs are able to process these large amounts of data.

Implications for clinicians

SMT is similarly effective as recommended and non-recommended interventions and when added an adjuvant therapy, in reducing pain and improving function in patients with chronic LBP. For patients with chronic LBP, SMT is a treatment option. SMT can be delivered as a standalone therapy, although it is typically offered within the constructs of a broader treatment package, together with exercise therapy or combined with usual care, as is recommended in recent national guidelines for low back pain [6–8]. This is important because SMT is by nature a passive treatment. Therefore, to prevent inappropriate behaviour and to empower patients to take control of their condition it is vital that practitioners impart evidence-based messages about passive interventions such as SMT. The choice of treatment should be the result of a shared decision-making process, taking patient preferences and clinicians experience and skills into account. No more research is needed to support these recommendations. Further similar research is unlikely to change these conclusions.

Adverse events were often not recorded and when recorded, were measured differently across trials. Consequently, we were not able to pool these data. These data did not provide more information than the adverse events described in our systematic review of aggregate data [5].

Conclusion/clinical implication

Sufficient evidence suggest that SMT provides similar outcomes to recommended therapies for pain relief and improvement of functional status. SMT would appear to be a good option for the treatment of chronic LBP.

Ethical approval: The study protocol was approved by the Review Board of the coordinating institution (EMGO Institute VU University Amsterdam). The protocol has also been approved by the Ethical Committee (Ref. No. 2015.554) of the VU University.

Funding: This systematic review was funded by a grant from European Chiropractic Union Research Fund Contract No A.14.03.

Conflict of interest: ADZ works in clinical practice as a chiropractor and treats patients with chronic low-back pain.

Key messages

- Randomised controlled trials of varying methodological quality and size have examined the benefits and harms of spinal manipulative therapy (SMT) for the treatment of chronic low back pain. These trials have been summarised in systematic reviews with varying results. SMT is not currently recommended as a first line treatment for chronic low back pain and its effects are uncertain.
- SMT is a good option for the treatment of chronic low back pain.
- Future trials of SMT for low back pain should include an economic evaluation and a better description of the qualitative and quantitative components of SMT (e.g. context of the visit, patient beliefs, and preferences and factors that are likely to influence treatment). Future initiatives should also focus on standardizing the manner in which inclusion and exclusion criteria, outcomes and moderators are defined, measured and reported. This will facilitate an effective comparison of interventions across trials.

SMR reports grants from European Chiropractor's Union, grants from Netherlands Chiropractic Association, grants from European Centre for Chiropractic Research Excellence, grants from Belgian Chiropractic Association, during the conduct of the study; and SMR works in clinical practice as a chiropractor and treats patients with chronic low-back pain.

MU was Chair of the NICE accreditation advisory committee until March 2017 for which he received a fee. He is chief investigator or co-investigator on multiple previous and current research grants from the UK National Institute for Health Research, Arthritis Research UK and is a co-investigator on grants funded by the Australian NHMRC. He is a NIHR Senior Investigator. He has received travel expenses for speaking at conferences from the professional organisations hosting the conferences. He is a director and shareholder of Clinvivo Ltd. that provides electronic data collection for health services research. He is part of an academic partnership with Serco Ltd. related to return to work initiatives. He is a co-investigator on a study receiving support in kind from Orthospace Ltd. He is an editor of the NIHR journal series, and a member of the NIHR Journal editors group, for which he receives a fee. He has published multiple papers on LBP some of which are referenced in this paper. He was lead author on one study included in the IPD meta-analysis. The authors declare they have no competing interests.

Acknowledgements

We thank Martin Roosenberg for the advice in writing the article and correcting the English. We thank the Euro-

pean Chiropractic Union for receiving the 2nd price ECCRE Research Award at the ECU congress 2018 in Budapest, Hungary.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.physio.2021.03.006.

References

[1] GBD Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1789–858.

[2] Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, *et al.* Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 2018;391(10137):2368–83.

[3] Coulter ID, Crawford C, Hurwitz EL, Vernon H, Khorsan R, Suttorp Booth M, *et al.* Manipulation and mobilization for treating chronic low back pain: a systematic review and meta-analysis. Spine J 2018;18(5):866–79.

[4] Ruddock JK, Sallis H, Ness A, Perry RE. Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. J Chiropr Med 2016;15(3):165–83.

[5] Rubinstein SM, de Zoete A, van Middelkoop M, Assendelft WJJ, 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:l689.

[6] NICE guideline [NG59]. Low back pain and sciatica in over 16s: assessment and management; 2016 https://www.nice.org.uk/guidance/NG59/chapter/Recommendations#non-invasive-treatments-for-low-back-pain-and-sciatica.

[7] Bons SCS, Borg MAJP, Van den Donk M, Koes BW, Kuijpers T, Ostelo RWJG, et al. NHG guideline for aspecific low-back pain. https://www.nhg.org/standaarden/samenvatting/aspecifieke-lagerugpijn#idp23613872.2017.

[8] Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of P. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017;166(7):514–30.

[9] Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, *et al.* Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. JAMA 2015;313(16):1657–65.

[10] de Zoete A, de Boer MR, van Tulder MW, Rubinstein SM, Underwood M, Hayden JA, *et al.* Rational and design of an individual participant data meta-analysis of spinal manipulative therapy for chronic low back pain-a protocol. Syst Rev 2017;6(1):21.

[11] 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. J Manipulative Physiol Ther 2014;37(8):523–41.

[12] Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. Cochrane Database Syst Rev 2011;(2):CD008112.

[13] Grande-Alonso M, Suso-Marti L, Cuenca-Martinez F, Pardo-Montero J, Gil-Martinez A, La Touche R. Physiotherapy based on a biobehavioral approach with or without orthopedic manual physical therapy in the treatment of nonspecific chronic low back pain: a randomized controlled trial. Pain Med 2019;20(12):2571–87.

[14] de Oliveira Meirelles F, de Oliveira Muniz Cunha JC, da Silva EB. Osteopathic manipulation treatment versus therapeutic exercises in patients with chronic nonspecific low back pain: a randomized, controlled and double-blind study. J Back Musculoskelet Rehabil 2020;33(3):367–77.

[15] Ford JJ, Slater SL, Richards MC, Surkitt LD, Chan AYP, Taylor NF, *et al.* Individualised manual therapy plus guideline-based advice vs advice alone for people with clinical features of lumbar zygapophyseal joint pain: a randomised controlled trial. Physiotherapy 2019;105(1):53–64.

[16] Goertz CM, Long CR, Vining RD, Pohlman KA, Walter J, Coulter I. Effect of usual medical care plus chiropractic care vs usual medical care alone on pain and disability among US service members with low back pain: a comparative effectiveness clinical trial. JAMA Netw Open 2018;1(1):e180105.

[17] Schulz C, Evans R, Maiers M, Schulz K, Leininger B, Bronfort G. Spinal manipulative therapy and exercise for older adults with chronic low back pain: a randomized clinical trial. Chiropr Man Therap 2019;27:21.

[18] Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, *et al.* Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006;15(Suppl 2):S192–300.

[19] Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, *et al.* 2015 updated method guideline for systematic reviews in the cochrane back and neck group. Spine (Phila Pa 1976) 2015;40(21):1660–73.

[20] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010;340:c221.

[21] Furlan AD, Pennick V, Bombardier C, van Tulder M, Group EBCBR. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976) 2009;34(18):1929–41.

[22] Cohen J. Statistical power analysis for the behavioral sciences. 2nd edn Hillsdale: Lawrence Earlbaum Associates; 1988.

[23] Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration Version 510; 2011 [updated March 2011]. Available from wwwhandbookcochraneorg.

[24] Balthazard P, de Goumoens P, Rivier G, Demeulenaere P, Ballabeni P, Deriaz O. Manual therapy followed by specific active exercises versus a placebo followed by specific active exercises on the improvement of functional disability in patients with chronic non specific low back pain: a randomized controlled trial. BMC Musculoskelet Disord 2012;13:162.

[25] Bronfort G, Hondras MA, Schulz CA, Evans RL, Long CR, Grimm R. Spinal manipulation and home exercise with advice for subacute and chronic back-related leg pain: a trial with adaptive allocation. Ann Intern Med 2014;161(6):381–91.

[26] Bronfort G, Maiers MJ, Evans RL, Schulz CA, Bracha Y, Svendsen KH, *et al.* Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. Spine J 2011;11(7):585–98.

[27] Cecchi F, Molino-Lova R, Chiti M, Pasquini G, Paperini A, Conti AA, *et al.* Spinal manipulation compared with back school and with individually delivered physiotherapy for the treatment of chronic low back pain: a randomized trial with one-year follow-up. Clin Rehabil 2010;24(1):26–36.

[28] Cook C, Learman K, Showalter C, Kabbaz V, O'Halloran B. Early use of thrust manipulation versus non-thrust manipulation: a randomized clinical trial. Man Ther 2013;18(3):191–8.

[29] Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, *et al.* Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. Pain 2007;131(1–2):31–7.

[30] Gudavalli MR, Cambron JA, McGregor M, Jedlicka J, Keenum M, Ghanayem AJ, *et al.* A randomized clinical trial and subgroup analysis

to compare flexion-distraction with active exercise for chronic low back pain. Eur Spine J 2006;15(7):1070–82.
[31] Haas M, Vavrek D, Peterson D, Polissar N, Neradilek MB. Dose-response and efficacy of spinal manipulation for care of chronic low back pain: a randomized controlled trial. Spine J 2014;14(7):1106–16.
[32] Hondras MA, Long CR, Cao Y, Rowell RM, Meeker WC. A randomized controlled trial comparing 2 types of spinal manipulation and minimal conservative medical care for adults 55 years and older with subacute or chronic low back pain. J Manipulative Physiol Ther 2009;32(5):330–43.
[33] Hsieh CY, Adams AH, Tobis J, Hong CZ, Danielson C, Platt K, *et al.* Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. Spine (Phila Pa 1976) 2002;27(11):1142–8.
[34] Skillgate E, Vingard E, Alfredsson L. Naprapathic manual therapy or evidence-based care for back and neck pain: a randomized, controlled trial. Clin J Pain 2007;23(5):431–9.
[35] Verma Y, Goyal M, Narkeesj D. Pain, range of motion and back strength in chronic low back pain before and after lumbar mobilisation. J Physiother Res 2013;(3):48–57.
[36] Vismara L, Cimolin V, Menegoni F, Zaina F, Galli M, Negrini S, *et al.* Osteopathic manipulative treatment in obese patients with chronic low back pain: a pilot study. Man Ther 2012;17(5):451–5.
[37] Walker BF, Hebert JJ, Stomski NJ, Losco B, French SD. Short-term usual chiropractic care for spinal pain: a randomized controlled trial. Spine (Phila Pa 1976) 2013;38(24):2071–8.
[38] Wilkey A, Gregory M, Byfield D, McCarthy PW. A comparison between chiropractic management and pain clinic management for chronic low-back pain in a national health service outpatient clinic. J Altern Complement Med 2008;14(5):465–73.
[39] Xia T, Long CR, Gudavalli MR, Wilder DG, Vining RD, Rowell RM, *et al.* Similar effects of thrust and nonthrust spinal manipulation found in adults with subacute and chronic low back pain: a controlled trial with adaptive allocation. Spine (Phila Pa 1976) 2016;41(12):E702–9.
[40] Zaproudina N, Hietikko T, Hanninen OO, Airaksinen O. Effectiveness of traditional bone setting in treating chronic low back pain: a randomised pilot trial. Complement Ther Med 2009;17(1):23–8.
[41] Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. Man Ther 2003;8(4):233–41.
[42] Petersen T, Larsen K, Nordsteen J, Olsen S, Fournier G, Jacobsen S. The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: a randomized controlled trial. Spine (Phila Pa 1976) 2011;36(24):1999–2010.
[43] UK Beam Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ 2004;329(7479):1377.
[44] Hidalgo B, Pitance L, Hall T, Detrembleur C, Nielens H. Short-term effects of Mulligan mobilization with movement on pain, disability, and kinematic spinal movements in patients with nonspecific low back pain: a randomized placebo-controlled trial. J Manipulative Physiol Ther 2015;38(6):365–74.
[45] Chou R, Deyo R, Friedly J, Skelly A, Hashimoto R, Weimer M, *et al.*, Rockville (MD) Noninvasive treatments for low back pain. AHRQ comparative effectiveness reviews; 2016.
[46] Franke H, Franke JD, Fryer G. Osteopathic manipulative treatment for nonspecific low back pain: a systematic review and meta-analysis. BMC Musculoskelet Disord 2014;15:286.
[47] Hayden JA, Wilson MN, Stewart S, Cartwright JL, Smith AO, Riley RD, *et al.* Exercise treatment effect modifiers in persistent low back pain: an individual participant data meta-analysis of 3514 participants from 27 randomised controlled trials. Br J Sports Med 2019;0:1–16.
[48] Buffart Lm, Kalter J, Sweegers Mg, Courneya Ks, Newton Ru, Aaronson Nk, *et al.* Effects and moderators of exercise on quality of life and physical function in patients with cancer: an individual patient data meta-analysis of 34 RCTs. Cancer Treat Rev 2017;52:91–104.

Available online at www.sciencedirect.com
ScienceDirect