REVIEW ARTICLE

Spinal manipulative therapy in older adults with chronic low back pain: an individual participant data meta-analysis

Alan Jenks[1] · Annemarie de Zoete[1] · Maurits van Tulder[2,3] · Sidney M. Rubinstein[1] · International IPD-SMT group

Abstract

Purpose Many systematic reviews have reported on the effectiveness of spinal manipulative therapy (SMT) for low back pain (LBP) in adults. Much less is known about the older population regarding the effects of SMT.

Objective To assess the effects of SMT on pain and function in older adults with chronic LBP in an individual participant data (IPD) meta-analysis.

Setting Electronic databases from 2000 until June 2020, and reference lists of eligible trials and related reviews.

Design and subjects Randomized controlled trials (RCTs) which examined the effects of SMT in adults with chronic LBP compared to interventions recommended in international LBP guidelines.

Methods Authors of trials eligible for our IPD meta-analysis were contacted to share data. Two review authors conducted a risk of bias assessment. Primary results were examined in a one-stage mixed model, and a two-stage analysis was conducted in order to confirm findings.

Main outcomes and measures Pain and functional status examined at 4, 13, 26, and 52 weeks.

Results 10 studies were retrieved, including 786 individuals, of which 261 were between 65 and 91 years of age. There is moderate-quality evidence that SMT results in similar outcomes at 4 weeks (pain: mean difference [MD] -2.56, 95% confidence interval [CI] -5.78 to 0.66; functional status: standardized mean difference [SMD] -0.18, 95% CI -0.41 to 0.05). Second-stage and sensitivity analysis confirmed these findings.

Conclusion SMT provides similar outcomes to recommended interventions for pain and functional status in the older adult with chronic LBP. SMT should be considered a treatment for this patient population.

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-Analysis Protocol

✉ Alan Jenks
a.d.jenks@vu.nl

1 Faculty of Science, Department of Health Sciences, Vrije Universiteit, Gebouw MF, Flexruimte, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

2 Faculty of Movement and Behavioral Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

3 Department of Physiotherapy and Occupational Therapy, Aarhus University Hospital, Aarhus, Denmark

Purpose

Low back pain (LBP) is one of the leading causes of disability and a lower quality of life in older adults [1]. It is one of the top three reasons why this population group visits a general practitioner [2]. The prevalence and burden of LBP increase with age [3], yet little is known about the effectiveness of interventions for LBP in older adults [1].

Older adults with LBP are underrepresented in randomized controlled trials (RCTs) [4], and LBP remains ubiquitous among older adults in retirement [5]. Studies have demonstrated that LBP in older adults often lasts longer than 3 months [6] and is usually undertreated or mismanaged [7].

Older adults tend to have more than one illness, and the odds of having LBP are higher in older adults with multiple comorbidities [8]. Therefore, it is important to identify treatment options which are safe and effective for this population [9]. Guidelines advocate non-pharmacological treatments for

LBP, such as complementary health approaches [10]. Finding safe and effective treatments for the older adult with LBP should be a priority. One such non-pharmacological approach is spinal manipulative therapy (SMT), which is a technique used worldwide by a variety of healthcare providers, such as chiropractors, osteopaths, and physiotherapists.

Many systematic reviews and meta-analyses have analyzed the effects of SMT [11]. Their results suggest that it is an effective intervention for both the reduction in pain and the improvement in function, two of the core domains in LBP trials [12]. Systematic reviews examining the effectiveness of various non-pharmacological treatments in older adults with LBP [13, 14] identified only three studies that assessed the effect of SMT. Two of the three trials were included in this analysis, and the third trial was excluded due to average age below 55. Therefore, it is difficult to draw valid conclusions for this patient population due to the lack of trials.

Given this, one approach to examine the effectiveness of SMT in older adults with LBP is to perform individual participant data (IPD) meta-analysis. This type of analysis has distinct advantages over traditional aggregate meta-analysis. In an IPD meta-analysis, we can select certain individuals since we have the individual data for each participant. This is more efficient than setting up new trials, particularly if the data are sufficient in order to allow for a meaningful analysis. Additional advantages of IPD include allowing the investigator to analyze the data independently of how the data were reported in the original publication. This is in contrast to the traditional aggregate approach in which meta-analyses extract data at the study level, meaning that the author(s) of the review must rely on how the data were analyzed and presented originally. Additionally, IPD makes it possible to correct for baseline covariates which may influence the results, enabling a more precise, and thereby potentially more valid, calculation of the effect estimates [15].

In short, LBP is a common cause of disability in the older adult [16], and our current knowledge of LBP in this patient group is limited [17]. The objective of this IPD meta-analysis is to assess the effectiveness of SMT versus interventions recommended by the guidelines at 1-, 3-, 6- and 12-month follow-up in older 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 [18] (Appendix 1).

IPD database

A detailed description of the IPD database design and the procedures followed was published previously [19]. The protocol [19] for the original study upon which this analysis is based was registered with PROSPERO (https://www.crd.york.ac.uk/prospero/25714). The database includes the raw data from 21 RCTs, which were published between 2000 and April 2018 [11], examining the effects of SMT on chronic LBP. This study used the IPD database defined above and represents a secondary analysis as defined a priori in PROSPERO. We updated the search from April 2018 to June 2020 identifying one trial that met our inclusion criteria [20].

Study selection

Trials examining the effects of SMT versus recommended therapies in the older age-group with chronic low back pain were included.

Inclusion criteria: Patients with chronic LBP with or without leg pain, defined as LBP of > 12 weeks of duration and not attributable to a recognizable, known specific pathology (e.g., infection, fracture, tumor, radicular syndrome, or herniation) were included. Additionally, trials from primary or secondary care settings were included. When a mixed population was involved (e.g., subacute and chronic), only those participants with > 12 weeks of LBP were included. For this IPD meta-analysis, we selected only those trials that had included participants aged 55 and older.

Exclusion criteria: Studies were excluded if they: (1) used an inadequate randomization procedure (e.g., alternate allocation, allocation based on birth date); (2) included patients with LBP and other conditions, such as pregnancy or postoperative patients; (3) tested the immediate effect of a single treatment only; (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; and (5) included patients where there was a contraindication to SMT.

Types of interventions

Experimental intervention: Spinal manipulation (i.e., high-velocity low-amplitude [HVLA] techniques) and mobilization (i.e., low-velocity low-amplitude [LVLA] techniques) were defined as SMT.

Comparison: We addressed the effects on pain and functional status of SMT versus interventions (e.g., exercise therapy, usual care) that are consistently recommended in international guidelines [21–24], while SMT is not [25]. The determination for recommended therapy was based on Rubinstein et al. [11]. We categorized an intervention into

'recommended' when this was consistently stated in at least two of the guidelines.

Types of outcome measures

Primary outcomes were pain and back-specific functional status, as recommended in the core set of outcome measurements in LBP [12].

Data extraction and quality assessment

Risk of bias assessment

The 13 risks of bias criteria recommended by the Cochrane Back and Neck group were used [26] (Appendix 2). These 13 criteria are used to identify selection bias, performance bias, attrition bias, detection bias, and selective outcome reporting bias.

Data extracted were study characteristics, patient characteristics, types of outcomes, duration of follow-up, and descriptions of experimental and control interventions.

Preparing data for analyses

The original data were compared with the published data to check for completeness. All variables were then harmonized in a data harmonization platform developed for a previous IPD analysis [27].

All outcomes were pooled following a decision rule (Appendix 3). All pain scores were converted to a pain scale (range 0–100 where a higher score indicates more pain) following a decision rule. 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 the nominator $\left(Z \text{ score} = x_i - \overline{x}/\text{SD}\right)$. Analyzing these Z-scores resulted in standardized mean differences (SMDs). To ease interpretation of SMDs, 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 $\left(\text{SD}_{\text{pooled}} = \sqrt{\sum_{i=0}^{n} \frac{(n_i-1)*S^2}{(n_i-1)}}; n_i = \text{sample}\right.$ size for each trial; S = standard deviation for each trial).

Data analysis and synthesis

All analyses were based on the intention-to-treat principle. Our primary analyses consisted of one-stage IPD meta-analysis at 4, 13, 26, and 52 weeks of follow-up. We chose these specific intervals as they are standard follow-up moments in the treatment of LBP [11]. We did not examine the effects of SMT post-intervention as there were large inter-study

variations in the number and frequency of treatments and, consequently, the duration of therapies and follow-up data for the period immediately following the end of treatment. Lastly, longitudinal analyses were not performed as the models are too complex and do not converge.

Analyses were conducted using a random-effects model that was adjusted for baseline using the restricted maximum likelihood (REML) method, where a separate intercept and a separate residual variance for each study are specified. However, in most analyses, these models did not demonstrate convergence. Instead, we present the results adjusted for baseline and with a random intercept and common residual variance [28].

The pooled treatment effects of SMT were estimated with a mean difference or Z-score for continuous outcomes, including the 95% confidence interval (CI).

Sensitivity and subgroup analyses

In order to examine whether the RCTs included in this IPD meta-analysis were a representative sample of all RCTs assessing the effects of SMT in older adult patients, we conducted a two-stage sensitivity analysis wherein we examined the effect sizes of RCTs both included in this IPD meta-analysis and those which were eligible for inclusion, but for which no IPD were available. For the latter, we used published aggregate data of those eligible trials that had an average age above 55.

A post hoc sensitivity analysis was performed as the one-stage and two-stage estimates at 4 weeks for the outcome pain were not similar.

Lastly, we performed a moderator analysis for age. Age was dichotomized into 55–64 and 65 years and older. This moderator was analyzed using a one-stage random-effect IPD meta-analysis or get rid of 'a' before one-stage. The baseline outcome, treatment, age, and interaction between treatment and age were included as fixed effects. Study-specific intercepts were also included as fixed effects. Random treatment and interaction effects were added to the model. We performed these analyses for each time point and age separately to facilitate convergence of models. Centering the patient-level covariates about their study-specific means enabled us to separate the within- and across-study interactions [28]. The within-study interaction explained the patient-level variation in treatment response, while the across-study interaction represented the age effect on study level. We present the within-study interactions. A negative interaction coefficient indicates a more positive or less negative estimate of the intervention effects of SMT vs comparison for the group 65 years and older compared to 55–64 years old.

We refrained from presenting stratified results for subgroups of moderator variables, because these included a combination of within- and across-study information due

to differences in proportions of persons within the separate subgroups between studies.

Synthesis of evidence

The overall quality of the evidence for each outcome was evaluated using the GRADE approach [29] (Appendix 4), and assessment of clinical relevance was defined as small, medium, or large effect [26, 30]. Results from meta-epidemiological studies suggest that selection bias (i.e., randomization) and performance bias (i.e., blinding) are perhaps the more important forms of bias which influence treatment effects [31]; therefore, we focused on these two aspects when considering 'limitations' as part of the GRADE process.

Results

In total, of the 21 trials in the IPD database, ten RCTs met the inclusion criteria, all of which provided data for the primary analysis [32–41] (Table 1) (Fig. 1). One trial [42] that did not provide data and had an average age above 55 was used in the second-stage analysis. In total, 786 participants aged ≥ 55 years were examined (403 were randomized to the SMT group and 383 were randomized to the comparison group) (Table 2). Two studies [20, 42] fit the inclusion criteria but did not provide individual data. (Table 3) Their aggregate published results were used in the two-stage analyses (Table 4) as they had an average age above 55. We identified 261 participants (from a total 786) older than 65 years of age originating from seven studies [32–37, 40], representing a third of all cases. Of the 261 patients, three quarters of them came from three studies [33, 34, 36] and were evenly distributed between treatment arms.

Description of studies

Of the ten RCTs, nine compared SMT to exercise therapy [32–39, 41] and one evaluated the effects of SMT compared to standard medical care [40] (consisting of drug and non-drug therapies). The included trials varied with respect to recruitment method, type of SMT technique, number and duration of treatments, and type of practitioner (Table 1).

Sample sizes ranged from 5 to 220 (median = 78.6; interquartile range [IQR] = 16–132). It should be noted that some trials had multiple arms, and some included non-chronic LBP patients; therefore, the sample size for a given comparison should be considered to be smaller.

The patient characteristics at baseline for SMT versus recommended interventions are presented in Table 2. The average age of all participants was 63 years (standard deviation [SD] 6.7), and slightly more than half (58.4%) were women.

Risk of bias

Approximately 80% of the studies (*n* = 8/10) reported an adequate random sequence generation and allocation concealment [32, 33, 35, 36, 38, 40, 41]. Two trials provided an adequate overview of withdrawals or dropouts and were able to keep these to a minimum for the subsequent follow-up measurements [34, 40].

Missing data for primary outcomes ranged from 12% at 4 weeks to 21% at 52 weeks.

Effects of SMT vs recommended interventions

Pain and function improved by the end of treatment, and this improvement was sustained up to 12 months after randomization for all groups (Table 3).

One-stage analysis

Pain

There is moderate quality evidence that SMT has similar benefits to recommended interventions at all time points for pain (Table 3). The mean difference (MD) for SMT compared to recommended interventions is − 2.56 (95% CI − 5.78 to 0.66; scale 0–100) after 1 month, and these effects appear similar over the subsequent 12 months (Table 4). Further analysis on the group of patients 65 and older showed similar effects − 2.46 (95% CI − 7.41 to 2.48; scale 0–100) after 1 month and appear similar over the subsequent 12 months (Appendix 5).

Functional status

There is moderate quality evidence that SMT has similar benefits to recommended interventions at all time points for functional status (Table 3). The comparison of SMT and recommended interventions for functional status outcome demonstrated a SMD of − 0.18 (95% CI − 0.41. to 0.05; scale 0–100) (− 0.85 on RMDQ 24-point scale) after 1 month and remained similar over the subsequent 12 months (SMD − 0.15; 95% CI − 0.38 to 0.08; scale 0–100) (− 0.76 on RMDQ 24-point scale) (Table 3). Further analysis on the group of patients 65 and older showed similar effects − 0.32 (95% CI − 0.57 to − 0.08; scale 0–100) (− 0.79 on RMDQ 24-point scale) after 1 month and appear similar over the subsequent 12 months − 0.40 (95% CI − 0.77 to − 0.02) (− 0.73 on RMDQ 24-point scale) (Appendix 5).

Table 1 Descriptives of studies evaluating the effects of SMT on outcomes included in the database (n = 10) in alphabetical order of first author of patients 55 +

Author (year) Acronnym	Country	N	Interventions	Duration of LBP according to inclusion criteria	Type of manipulator	Type of manipulation	Max no. of treatments allowed and duration of treatment
Bronfort (2011)	USA	57	1. Supervised exercise (n=101) 2. Spinal manipulative therapy (n=100) 3. Home exercise and advice (n=100);	>6 weeks	Chiropractor (n=9)	Manipulation	Participants were discharged from care if the treating clinician felt that maximum clinical benefit was obtained. 12 wks of care
Cecchi (2010)	Italy	144	1. Back school (n=70); 2. Individualized physiotherapy (n=70); 3. Spinal manipulative therapy (n=70)	>6 mo	Physician (n=2)	Manipulation and mobilization	4–6 sessions per week for 4–6 wks
Ferreira (2007)	Australia	132	1. General exercise (n=80) 2. Motor control exercise 3. Spinal manipulative therapy	>3 mo	Physical therapist (n=?)	Mobilization or manipulation; Maitland	12 over 8 wks
Gudavelli (2006)	USA	37	1. Flexion distraction mobilization (n=123) 2. Exercise therapy (n=112)	>3 mo	Chiropractor (n=?)	Mobilization (flexion distraction)	16 over 4 wks
Hondras (2009)	USA	220	1. High-velocity low-amplitude spinal manipulative therapy (n=96) 2. Low-velocity variable amplitude spinal mobilization (n=95) 3. Medical care (n=49)	>4 wks	Chiropractor (n=4)	Manipulation or mobilization (flexion distraction) (depending upon grp. assignment)	12 over 6 wks
Hsieh (2002)	USA	16	1. Back school (n=48) 2. Myofascial therapy (n=51) 3. Joint manipulation (n=49) 4. Combination of treatments 2 & 3 (n=52)	>3 wks to <6 mo	Chiropractor (n=?)	Manipulation	9 over 3 wks
Petersen (2011)	Denmark	14	1. McKenzie therapy (n=175) 2. Spinal manipulative therapy (n=175)	>6 wks	Chiropractor (n=3)	Manipulation or mobilization	max 15 over 12 wks
Rasmussen-Bar (2003)	Sweden	5	1. Stabilizing training group (n=24) 2. Manual therapy group (n=23)	>6 wks	Manual therapist (n=?)	MOB	6 over 6 wks

Table 1 (continued)

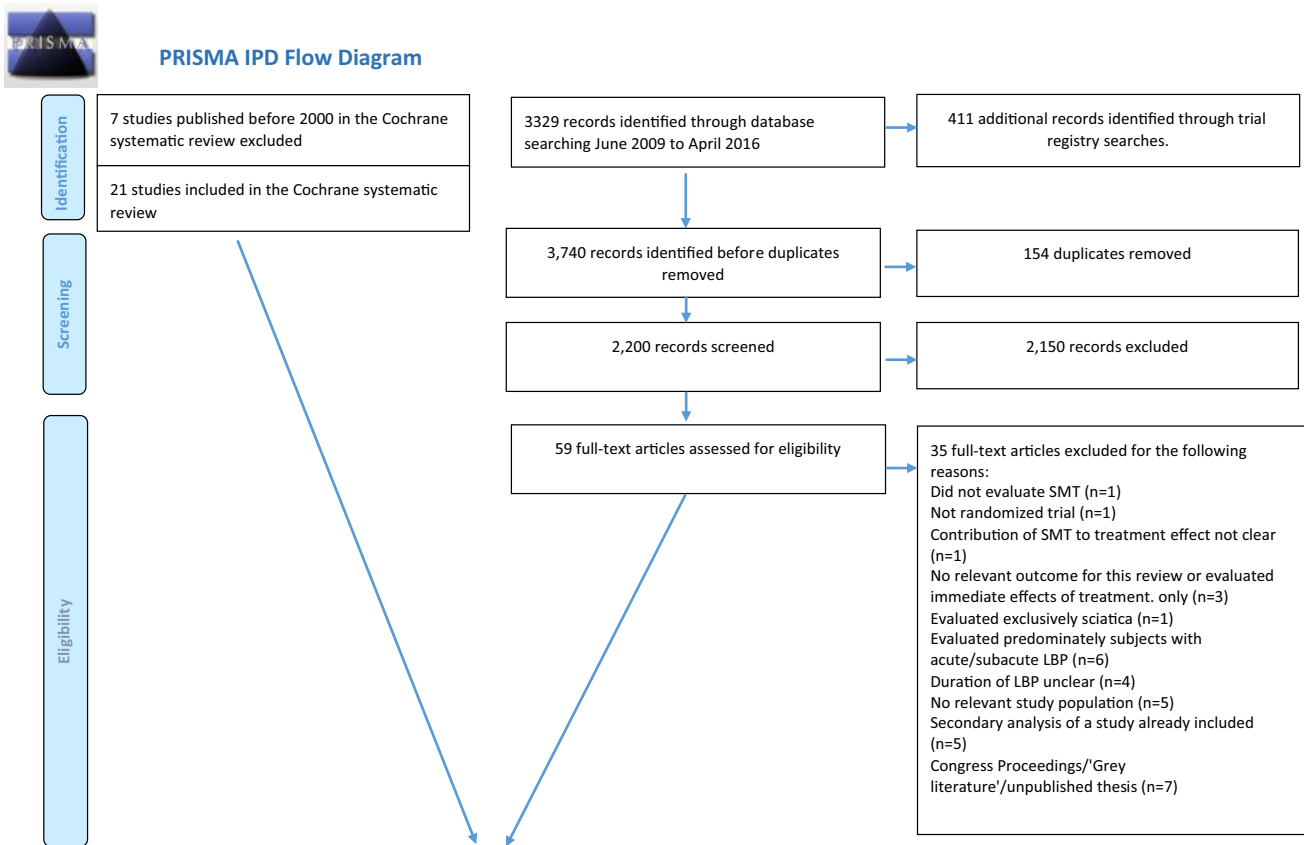| Author (year) Acronnym	Country	N	Interventions	Duration of LBP according to inclusion criteria	Type of manipulator	Type of manipulation	Max no. of treatments allowed and duration of treatment
Skillgate (2007)	Sweden	43	1. Naprapathy ($n=206$) 2. Standard care or 'evidence-based' care (provided by physician) ($n=203$)	>2 wks	Naprapath ($n=8$)	Manipulation or mobilization	6 over 6 wks
UK Beam (2004)	UK	118	1. Best care in general practice ($n=338$) 2. Best care plus exercise alone ($n=310$) 3. Best care plus private manipulation alone ($n=180$) 4. Best care plus NHS manipulation alone ($n=173$) 5. Best care plus private manipulation plus exercise ($n=172$) 6. Best care plus NHS manipulation plus exercise ($n=161$)	(Essentially) > 3 wks	Chiropractor, osteopath, or physiotherapist ($n=84$)	Manipulation or mobilization	8 over 12 wks

wks weeks; *mo* months; ? unclear/unknown

PRISMA IPD Flow Diagram

```
┌─────────────┐
│Identification│   ┌──────────────────────────────┐   ┌──────────────────────────────┐   ┌──────────────────────────────┐
│             │   │7 studies published before 2000 │   │3329 records identified through │   │411 additional records identified│
│             │   │in the Cochrane systematic      │   │database searching June 2009 to │   │through trial registry searches. │
│             │   │review excluded                 │   │April 2016                      │   │                                │
│             │   └──────────────────────────────┘   └──────────────────────────────┘   └──────────────────────────────┘
│             │   ┌──────────────────────────────┐
│             │   │21 studies included in the      │
│             │   │Cochrane systematic review      │
│             │   └──────────────────────────────┘
```


Fig. 1 Flow diagram of study inclusion

Flow diagram contents:

- 7 studies published before 2000 in the Cochrane systematic review excluded
- 21 studies included in the Cochrane systematic review
- 3329 records identified through database searching June 2009 to April 2016
- 411 additional records identified through trial registry searches.
- 3,740 records identified before duplicates removed
- 154 duplicates removed
- 2,200 records screened
- 2,150 records excluded
- 59 full-text articles assessed for eligibility
- 35 full-text articles excluded for the following reasons:
  Did not evaluate SMT (n=1)
  Not randomized trial (n=1)
  Contribution of SMT to treatment effect not clear (n=1)
  No relevant outcome for this review or evaluated immediate effects of treatment. only (n=3)
  Evaluated exclusively sciatica (n=1)
  Evaluated predominately subjects with acute/subacute LBP (n=6)
  Duration of LBP unclear (n=4)
  No relevant study population (n=5)
  Secondary analysis of a study already included (n=5)
  Congress Proceedings/'Grey literature'/unpublished thesis (n=7)

Sensitivity analyses and subgroup analysis

We identified two trials to be included in the two-stage analysis, one from the original systematic review [42] and the other from our updated search [20]. We included the aggregate results of these studies in the second-stage analysis after going through a risk of bias assessment. The two-stage analysis showed a MD similar to the one-stage analysis except for pain at 4 weeks (Table 4). The difference at 4 weeks was a result of two studies that included 5 patients, yet had a large effect on recommended therapies. The second-stage analysis confirmed the results of the one-stage analysis at all time points, showing robustness of the effect in both analyses (Appendix 5). A subgroup analysis using age as a moderator showed similar results to a previous IPD [43], that age does moderate any effect of the treatment (Appendix 5) (Tables 5 and 6).

Discussion

These results suggest that SMT has similar effects to recommended interventions, mainly exercise therapy, at the short, intermediate, and long term. This is the first IPD meta-analysis to examine the effects of SMT in older adults with LBP, although admittedly, the majority of subjects (two-thirds) were between 55 and 65 years of age; therefore, these results should perhaps be interpreted with caution. However, if there were big differences in effects, this might have become intuitively obvious from this subgroup analysis [43, 44]. Using age as a moderator also did not change the effects at all time points. The importance of these findings cannot be sufficiently underscored. Given the growing aging population and the burden of LBP, there is a need to provide safe, conservative treatments. These data provide support for the use of SMT in this population.

These findings have important implications. The recent Lancet series [45] suggests that SMT should be considered a second treatment option, following the more commonly recommended treatments for chronic LBP (e.g., exercise). Our results suggest that SMT produces similar effects to other commonly recommended interventions for older patients with LBP. This is particularly pertinent because prior to this analysis, these effects were unclear. However, a note of caution is perhaps necessary because we did not examine adverse reactions in detail. These data were not registered in any systematic way in the individual studies and were
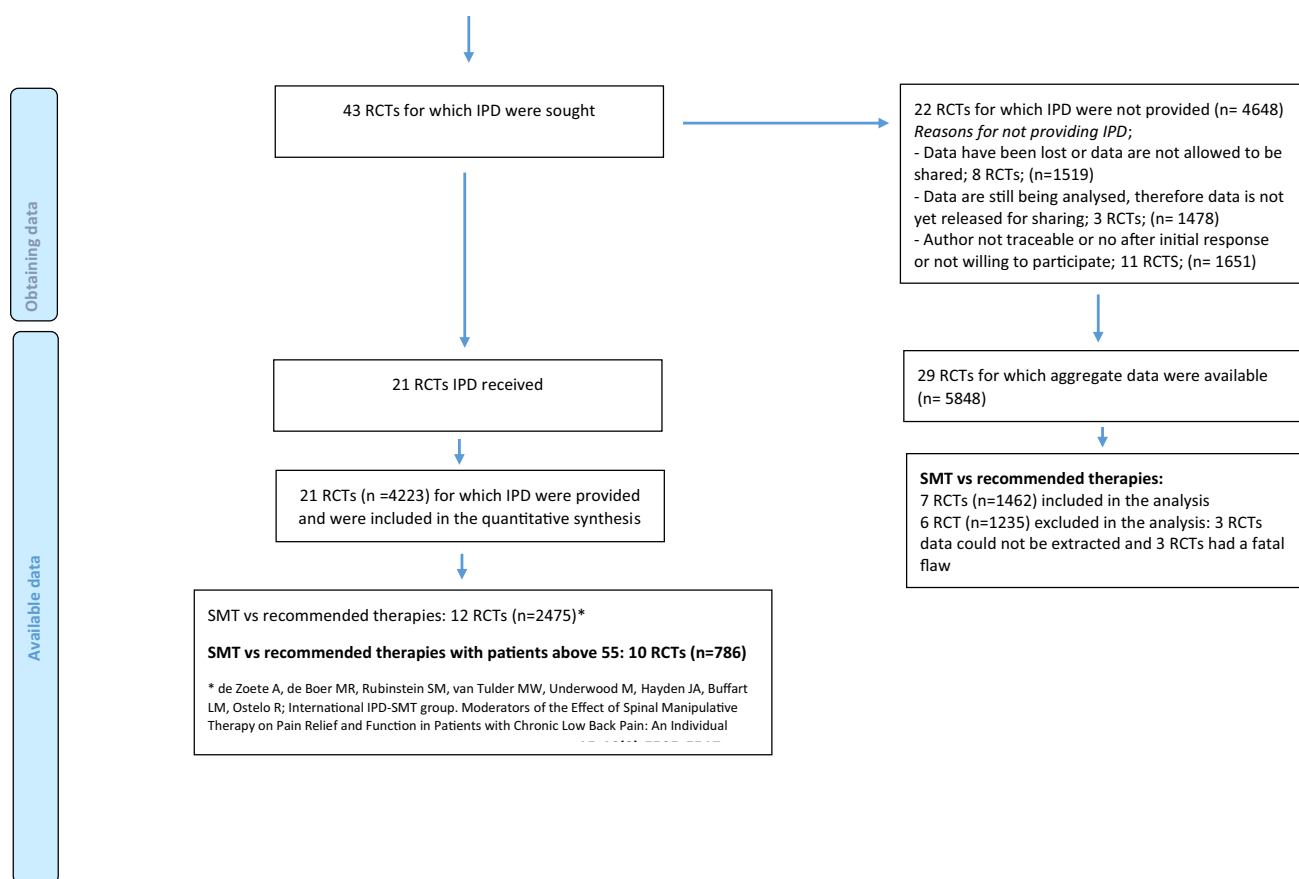
Fig. 1 (continued)

Table 2 Descriptives of studies evaluating the effects of SMT on outcomes used in the two-stage analysis ($n=2$) in alphabetical order of first author of patients 55 +

Author	Country	N	Avg. age (SD)	Interventions	Duration of LBP according to inclusion criteria	Type of manipulator	Type of manipulation	Max no. of treatments allowed and duration of treatment
Dougherty (2014)	USA	136	77.02 (6.8)	Spinal manipulation or flexion distraction. Sham treatment was detuned ultrasound	> 3 wks	Chiropractor	Manipulation or sham	2 × a week for 4 weeks, duration of tx ~ 15 min
Schultz (2020)	USA	241	73.6 (5.5)	Home exercise program, supervised exercise or spinal manipulative therapy + home exercise program	> 6 wks	Chiropractor + exercise therapists	Manipulation or exercise program	12 weeks of care

Table 3 Patient characteristics at baseline for groups receiving SMT vs groups receiving recommended interventions

	SMT vs recommended interventions ($m = 10$; $n = 786$)	
Demographic data	SMT ($n = 403$)	Recommended interventions ($n = 383$)
Age, mean (SD) years ($m = 10$, $n = 786$)	62.2 (6.2)	63.8 (7.0)
Sex, *n* (%) female ($m = 10$, $n = 786$)	220 (55)	239 (62)
Body Mass Index, mean (SD) ($m = 7$, $n = 596$)	28.0 (5.2)	27.9 (5.0)
Ethnicity, *n* (%) white ($m = 4$, $n = 330$)	93.2	89.2
Lifestyle factors		
Physical activity (%) ($m = 5$, $n = 135$)	55 (13.6)	80 (20.9)
Low (1 or less than once a week)	17 (4.2)	28 (7.3)
Medium (2–3×a week)	24 (3.5)	25 (6.5)
High (more than 3× a week)	14 (3.5)	27 (7.0)
Smoker, *n* (%) non-smokers ($m = 6$, $n = 494$)	37 (9.2)	62 (16.2)
Socio-demographics		
Marital status, *n* (%) married; living with a partner ($m = 5$, $n = 372$)	290 (72)	226 (59)
Level of education, *n* (%) low/middle ($m = 5$, $n = 372$)	300 (74)	187 (49)
Employment status, *n* (%) at work ($m = 8$, $n = 696$)	370 (91)	362 (85)
Nature and severity of LBP		
Duration of LBP, *n* (%) less than 12 months ($m = 6$, $n = 206$)	272 (68)	206 (54)
Leg pain, *n* (%) ($m = 5$, $n = 1038$)	219 (54)	109 (29)
Previous LBP treatment received, *n* (%) ($m = 4$, $n = 109$)	252 (63)	160 (42)
Previous physiotherapy for low back pain received, *n* (%) ($m = 1$, $n = 11$)	4 (1)	11 (3)
Previous SMT for low back pain received, *n* (%) ($m = 1$, $n = 43$)	177 (44)	43 (11)
Used medication for low back, *n* (%) ($m = 1$, $n = 19$)	24(6)	19(5)
Psychosocial factors	SMT	Control
Depression, *n* (%) ($m = 1$, $n = 57$)	11.3(11)	9.58(10)
Primary outcomes		
Pain		
Combined pain score at baseline, mean (SD), ($m = 10$, $n = 375$)	400(99)	375(98)
Combined pain score at 4 weeks, mean (SD), ($m = 9$, $n = 333$)	370(92)	333(87)
Combined pain score at 13 weeks, mean (SD), ($m = 7$, $n = 222$)	159(40)	222(58)
Combined pain score at 26 weeks, mean (SD), ($m = 7$, $n = 259$)	146(36)	259(68)
Combined pain score at 52 weeks, mean (SD), ($m = 8$, $n = 531$)	192(47)	295(77)
Functional status		
RMDQ sum score at baseline, mean (SD), ($m = 2$, $n = 195$)	64(16)	131(34)
RMDQ sum score at 4 weeks, mean (SD), ($m = 2$, $n = 195$)	64(16)	131(34)
RMDQ sum score at 13 weeks, mean (SD), ($m = 2$, n $= 195$)	64(16)	131(34)
RMDQ sum score at 26 weeks, mean (SD), ($m = 2$, $n = 195$)	64(16)	131(34)
RMDQ sum score at 52 weeks, mean (SD), ($m = 2$, $n = 195$)	64(16)	131(34)
ODI sum score at baseline, mean (SD),($m = 1$,n$= 5$)	*	*

SD standard deviation; *m* number of studies; *n* number of participants; *RMDQ* Roland Morris Disability Questionnaire; *ODI* Oswestry Disability Index

*Combining categories was not meaningful or no data available

not directly available; therefore, uncommon and potentially serious adverse reactions cannot be ruled out.

Importantly, our results appear consistent with recent systematic reviews using aggregate data on the effects of SMT for adults with LBP [11] as well as older adults [13, 14]. An important difference of our IPD analysis compared to traditional aggregate meta-analyses is that we could adjust for the baseline pain and functional status and were not dependent

Table 4 Main treatment effects and GRADE summary of findings for SMT vs recommended interventions for the primary outcomes

Comparison 1: SMT versus recommended therapies

Time measurement	Difference in effect (β) (95% CI)	# studies	N	Quality of the evidence (and reason for downgrading)	Comments
Outcome: Pain					
1 month	MD − 2.56, 95% CI − 5.78 to 0.66	9	693	Moderate (inconsistency)	
3 months	MD − 7.54, 95% CI − 15.38 to 0.30	7	372	Moderate (inconsistency)	
6 months	MD − 6.47, 95% CI − 13.22 to − 0.30	7	405	Moderate (inconsistency)	
12 months	MD − 1.74, 95% CI − 9.08 to 5.55	8	477	Moderate (inconsistency)	
Outcome: Functional status					SMD converted to a MD on the 24-point RMDQ scale
1 month	SMD − 0.18, 95% CI − 0.41 to 0.05	9	707	Moderate (inconsistency)	− 0.85
3 months	SMD − 0.23, 95% CI − 0.41 to 0.05	9	583	Moderate (inconsistency)	− 1.08
6 months	SMD − 0.25, 95% CI − 0.43 to 0.07	8	586	Moderate (inconsistency)	− 1.27
12 months	SMD − 0.15, 95% CI − 0.38 to 0.08	8	487	Moderate (inconsistency)	− 0.76

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.

MD median difference; *SMD* standard median difference

upon how these data were reported in the original publications. This adjustment increased the precision of our estimates compared to aggregate data meta-analyses, but did not lead to a different conclusion for the main effects.

Adverse events were often not reported by trial authors, and when reported there was no uniformity in how this was done, particularly for older patients; therefore, these data do not provide more information than the adverse events described in our systematic review of aggregate data [11]. The adverse events which were reported are likely to be more serious events for which reporting was required, or were unrelated to SMT [20]. Nevertheless, there may be a theoretically increased risk with SMT which would need to be examined in future studies and compared to recommendations like exercise therapy (e.g., osteoporosis). In short, the risk of (major) adverse events is likely to be very low and may reflect adaptation by the therapist for this patient population.

Strengths and weaknesses

These results should be interpreted in light of a few strengths and limitations. The most important strength is that we included 786 patients from ten trials in our analysis. Furthermore, these patients came from 10 of the 11 trials that could have provided data, which minimized selection bias. Additionally, all trials provided data for pain and functional status for all the time points analyzed; lastly, the one-stage

estimates were confirmed by the two-stage analysis, suggesting that our effects estimates were robust.

Study limitations

There are, however, some important limitations. Inclusion bias cannot be ruled out. We may have missed some important studies published after 2018. In order to determine whether this might be the case, we performed a cursory search of the literature in PubMed (up to June 2020). We identified 18 potential articles. Upon further analysis, 17 were excluded for various reasons, including younger age, lack of randomization, a protocol, or other type of comparison (e.g., SMT as adjuvant therapy). In short, only one study fulfilled the inclusion criteria which could have been included in an update. We analyzed that trial [20] in the two-stage analysis, and those results were consistent with the one-stage analysis.

Additionally, selection bias cannot be ruled out; 11 of the 21 studies identified in the search did not provide IPD data. Of those, three included patients older than 55 [46–48], but relatively few subjects would have been included because the average age was under 55 (SDs ranged from 12 to 15). One [42] trial included subjects with an average age over 55 and was examined in our second-stage analysis. Again, those results were consistent with our one-stage analysis. This suggests that our analysis is representative of all subjects that could have been included and, therefore, robust.

Table 5 Representativeness of the pooled effects of studies providing data for the IPD study and those not providing data

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 4 weeks					
All eligible studies + *s	11	2.06 (− 6.84; 10.97)	93.6	0.000	2.06 (− 32.83; 36.95)
Studies providing data	9	2.44 (− 8.54; 13.42)	94.8	0.000	2.44 (− 38.12; 42.99)
Studies not providing data	2	0.19 (− 13.77; 14.16)	78.8	0.030	0.19 (-;-)
Combined pain score 13 weeks					
All eligible studies + *s	11	− 4.65 (− 9.18; − 0.12)	82.4	0.000	− 4.65 (− 20,88; 11.58)
Studies providing data	9	− 5.92 (− 11.38; − 0.46)	86.3	0.000	− 5.92 (− 24.97; 12.95)
Studies not providing data	2	− 1.61 (− 6.46; 3.23)	0.0	0.370	− 1.61 (− 33.03; 29.81)
Combined pain score 26 weeks					
All eligible studies + *s	9	− 3.79 (− 7.14; − 0.44)	63.6	0.003	− 3.79 (− 14.14; 6.56)
Studies providing data	7	− 4.21 (− 8.47; − 0.05)	74.6	0.001	− 4.21 (− 17.97; 9.55)
Studies not providing data	2	− 1.81 (− 6.66; 3.04)	0.0	0.786	− 1.81 (− 33.22; 29.60)
Combined pain score 52 weeks					
All eligible studies	9	− 1.18 (− 5.27; 2.91)	78.1	0.000	− 1.18 (− 14.84; 12.48)
Studies providing data	8	− 1.11 (− 5.73; 3.51)	80.8	0.000	− 1.11 (− 16.48; 14.25)
Studies not providing data	1	− 1.30 (− 7.22; 4.62)	100		− 1.30 (-;-)
Functional status		Difference in Z-score			
Combined functional status 4 weeks					
Studies providing data	8	− 0.20 (− 0.43; 0.03)	72.6	0.001	− 0.20 (− 0.92; 0.52)
All eligible studies + *s	10	− 0.09 (− 0.31; 0.12)	83.1	0.000	− 0.09 (− 0.85; 0.66)
Studies not providing data	2	0.17 (0.07; 0.28)	0.0	0.858	0.07 (− 0.50; 0.85)
Combined functional status 13 weeks					
All eligible studies	10	− 0.06 (− 0.24; 0.12)	75.1	0.00	− 0.06 (− 0.77; 0.65)
Studies providing data	9	− 0.20 (− 0.43; 0.03)	78.9	0.00	− 0.20 (− 0.89; 0.57)
Studies not providing data	1	0.18 (− 0.07; 0.42)	45.4	0.089	0.18 (− 0.47; 0.82)
Combined functional status 26 weeks					
All eligible studies	9	− 0.11 (− 0.24; 0.03)	56.1	0.002	− 0.11 (− 0.56; 0.34)
Studies providing data	8	− 0.23 (− 0.44; − 0.01)	69.4	0.001	− 0.23 (− 0.91; 0.45)
Studies not providing data	1	0.01 (− 0.11; 0.13)	0.0	0.778	0.01 (− 0.14; 0.16)
Combined functional status 52 weeks					
All eligible studies	9	− 0.15 (− 0.31; 0.05)	72.2	0.00	− 0.15 (− 0.72; 0.42)
Studies providing data	8	− 0.17 (− 0.39; 0.05)	78.9	0.00	− 0.17 (− 0.92; 0.57)
Studies not providing data	1	− 0.10 (− 0.31; 0.11)	43.4	0.132	− 0.10 (− 0.68; 0.48)

Two-stage analysis; SMT versus recommended therapies

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.

*s = 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.

Implications for clinicians

SMT appears to be similarly effective to recommended therapies for reducing pain and improving function in older patients with chronic LBP, meaning SMT may be delivered as a stand-alone therapy. Future research should focus on identifying which older adults are best suited for SMT, taking lifestyle factors, comorbidities, and level of physical activity into account.

Table 6 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 4 weeks					
All eligible studies	10	2.84 (−6.75; 12.44)	93.5	0.000	2.84 (−33.85; 39.53)
All eligible studies + *s	11	2.06 (−6..84; 10.97)	93.6	0.000	2.06(−32.83; 36.95)
Studies providing data	9	2.44 (−8.54; 13.42)	94.8	0.000	2.44 (−38.12;42.99)
Studies w/o *r*h	7	−2.78 (−6.30; 0.74)	27.2	0.22	−2.78(−10.61; 5.05)
Studies w/o *r*h & w/ *s	8	−3.39 (−6.46; −0.32)	23.5	0.24	−3.39(−9.87; 3.08)
Studies not data providing	2				
Combined pain score 13 weeks					
All eligible studies	10	−4.70 (−9.71; 0.31)	84.0	0.000	−4.70 (−22,34; 12.95)
Studies providing data	9	−5.92 (−11.38; −0.46)	86.3	0.000	−5.92 (−24.97; 12.95)
All eligible studies w/ *s	11	−4.65 (−9.18; −0.12)	82.4	0.000	−4.65 (−20.88; 11.58)
Studies not providing data	2				
Combined pain score 26 weeks					
All eligible studies	8	−4.20 (−7.90; −0.05)	66.1	0.003	−4.14 (−15.57; 7.17)
Studies providing data	7	−4.21 (−8.47; −0.05)	74.6	0.001	−4.21 (−17.97;9.55)
All eligible studies + *s	9	−3.79 (−7.14; −0.44)	63.6	0.003	−3.79 (−14.14; 6.56)
Studies not providing data	2				
Combined pain score 52 weeks					
All eligible studies	9	−1.18 (−5.27; 2.91)	78.1	0.000	−1.18 (−14.84; 12.48)
Studies providing data	8	−1.11 (−5.73; 3.51)	80.8	0.000	−1.11 (−16.48;14.25)
Studies not providing data	1				
Functional status		Difference in Z-score			
Combined functional status 4 weeks					
All eligible studies	9	−0.12 (−0.35; 0.12)	84.5	0.000	−0.20 (−0.93; 0.69)
Studies providing data	8	−0.20 (−0.43; 0.03)	72.6	0.001	−0.20 (−0.92; 0.52)
All eligible studies + *s	10	−0.09 (−0.31; 0.12)	83.1	0.000	−0.09 (−0.85; 0.66)
Studies not providing data	2				
Combined functional status 13 weeks					
All eligible studies	10	−0.06 (−0.24; 0.12)	75.1	0.00	−0.06 (−0.77; 0.65)
Studies providing data	9	−0.20 (−0.43; 0.03)	78.9	0.00	−0.20 (−0.89;0.57)
Studies not providing data	1	0.18 (−0.07; 0.42)	45.4	0.089	0.18 (−0.47; 0.82)
Combined functional status 26 weeks					
All eligible studies	9	−0.11 (−0.24; 0.03)	56.1	0.002	−0.11 (−0.56; 0.34)
Studies providing data	8	−0.23 (−0.44; −0.01)	69.4	0.001	−0.23 (−0.91; 0.45)
Studies not providing data	1	0.01 (−0.11; 0.13)	0.0	0.778	0.01 (−0.14; 0.16)
Combined functional status 52 weeks					
All eligible studies	9	−0.15 (−0.31; 0.05)	72.2	0.00	−0.15 (−0.72; 0.42)
Studies providing data	8	−0.17 (−0.39; 0.05)	78.9	0.00	−0.17 (−0.92; 0.57)
Studies not providing data	1	−0.10 (−0.31; 0.11)	43.4	0.132	−0.10 (−0.68; 0.48)

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

*r = Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low back pain. Manual Therapy. 2003; 8(4):233–41

*h = 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

*s = 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

Conclusion

SMT is equally as effective as recommended interventions for the treatment of chronic low back pain in the older adult. Over three quarters of the data came from adults aged 55–64, yet sensitivity analysis in the second stage and using age as a moderator showed results were similar across all age-groups. Therefore, SMT should be considered a treatment option in this patient population.

Appendix 1 PRISMA-IPD (Preferred reporting items for systematic review and meta-analysis individual patient data) checklist: recommended items to address in a systematic review protocol

PRISMA-IPD checklist of items to include when reporting a systematic review and meta-analysis of individual participant data (IPD)

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data	Article completed
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	Article completed page 2
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	Article completed page 3
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups	Article completed page 5

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable	PROSPERO CRD42015025714 (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=25714) Protocol: https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-017-0413-y
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated	Article completed page 4
Identifying studies-information sources	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings; use of study registers and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation	Article completed page 5
Identifying studies-search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated	Appendix 5
Study selection processes	9	State the process for determining which studies were eligible for inclusion	Article completed page 4
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study)	Article completed page 5–6
		If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators	
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardizing or translating variables within the IPD datasets to ensure common scales or measurements across studies	Completed

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done	Completed
Risk of bias assessment in individual studies	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any datasynthesis	Completed
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome	Completed
Synthesis methods	14	Describe the meta-analysis methods used to synthesize IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): Use of a one-stage or two-stage approach. How effect estimates were generated separately within each study and combined across studies (where applicable). Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. Use of fixed or random effects models and any other model assumptions, such as proportional hazards. How (summary) survival curves were generated (where applicable). Methods for quantifying statistical heterogeneity (such as I^2 and τ^2). How studies providing IPD and not providing IPD were analyzed together (where applicable). How missing data within the IPD were dealt with (where applicable)	Completed
Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analyzed as potential effect modifiers, and whether these were pre-specified	Completed
Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables	Completed
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified	Completed

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram	Appendix x
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD	Completed and Table 1
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none	There were none
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions	Article completed
Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot	Table x,x
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based	Completed and Table x and x
		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials	
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice	

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables	Completed Table x
Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarize the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available	Completed and Table x
Discussion			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome	Article completed page 8
Strengths and limitations	25	Discuss anyimportant strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available	Article completed page 8
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence	Article completed page 8–9
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research	Article completed page 9
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support	Article completed page 10

A1–A3 denote new items that are additional to standard PRISMA items. A4 has been created as a result of re-arranging content of the standard PRISMA statement to suit the way that systematic review IPD meta-analyses are reported.

Appendix 2: Criteria for a judgment of 'low risk of bias' for the sources of bias

1	A random (unpredictable) assignment sequence. Examples of adequate methods are coin toss (for studies with 2 groups), rolling a dice (for studies with 2 or more groups), drawing of balls of different colors, drawing of ballots with the study group labels from a dark bag, computer-generated random sequence, pre-ordered sealed envelopes, sequentially-ordered vials, telephone call to a central office, and pre-ordered list of treatment assignments. Examples of inadequate methods are: alternation, birth date, social insurance/security number, date in which they are invited to participate in the study, and hospital registration number
2	Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient
3	Index and control groups are indistinguishable for the patients or if the success of blinding was tested among the patients and it was successful
4	Index and control groups are indistinguishable for the care providers or if the success of blinding was tested among the care providers and it was successful

5	Adequacy of blinding should be assessed for each primary outcome separately. This item should be scored 'low risk' if the success of blinding was tested among the outcome assessors and it was successful or:
	-For patient-reported outcomes in which the patient is the outcome assessor (e.g., pain, disability): the blinding procedure is adequate for outcome assessors if participant blinding is scored 'low risk';-for outcome criteria assessed during scheduled visit and that supposes a contact between participants and outcome assessors (e.g., clinical examination): the blinding procedure is adequate if patients are blinded, and the treatment or adverse effects of the treatment cannot be noticed during clinical examination;-for outcome criteria that do not suppose a contact with participants (e.g., radiography, magnetic resonance imaging): the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed when assessing the main outcome;-for outcome criteria that are clinical or therapeutic events that will be determined by the interaction between patients and care providers (e.g., co-interventions, hospitalization length, treatment failure), in which the care provider is the outcome assessor: the blinding procedure is adequate for outcome assessors if item ''4'' (caregivers) is scored 'low risk';-for outcome criteria that are assessed from data of the medical forms: the blinding procedure is adequate if the treatment or adverse effects of the treatment cannot be noticed on the extracted data
6	The number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given. If the percentage of withdrawals and drop-outs does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias a 'low risk' is scored. (N.B. these percentages are arbitrary, not supported by literature)
7	All randomized patients are reported/analyzed in the group they were allocated to by randomization for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and co-interventions
8	All the results from all pre-specified outcomes have been adequately reported in the published report of the trial. This information is either obtained by comparing the protocol and the report, or in the absence of the protocol, assessing that the published report includes enough information to make this judgment. In the absence of a protocol, the authors judged this item to have been met when the expected outcomes (i.e. pain and functional status) were reported
9	Groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with neurological symptoms, and value of main outcome measure(s)
10	If there were no co-interventions or they were similar between the index and control groups
11	The reviewer determines if the compliance with the interventions is acceptable, based on the reported intensity, duration, number and frequency of sessions for both the index intervention and control intervention(s). For example, physiotherapy treatment is usually administered for several sessions; therefore it is necessary to assess how many sessions each patient attended. For single-session interventions (e.g., surgery), this item is irrelevant
12	Timing of outcome assessment should be identical for all intervention groups and for all primary outcome measures
13	Other types of biases. For example:-When the outcome measures were not valid. There should be evidence from a previous or present scientific study that the primary outcome can be considered valid in the context of the present.-Industry-sponsored trials. The conflict of interest (COI) statement should explicitly state that the researchers have had full possession of the trial process from planning to reporting without funding bodies with potential COI having any possibility to interfere in the process. If, for example, the statistical analyses have been done by a funding body with a potential COI, usually 'unsure' is scored

Appendix 3: Decision rules–primary and secondary outcomes

Analysis plan of outcome pain

All of the RCTs in the repository had asked participant to rate or mark on a numerical rating scale or a visual analogue scale that described either their average pain at the present time or over a defined weeks or months. This item was presented either as a single standalone instrument or as an item that was part of a collective pain measurement:

Analysis of average pain and pain intensity separately

For the analyses of average pain, one of the following instruments from each trial, where available, was chosen (in descending order):

(1) Individual visual analogue scale (VAS) on average pain today
(2) Average pain over the past one week
(3) The individual item of the Von Korff pain intensity score that is equivalent to the VAS if it is available

For the analyses of pain intensity, one of the following instruments from each trial, where available, was chosen (in descending order):

(1) Individual visual analogue scale (VAS) on pain intensity today (0–10 or 0–100)
(2) Pain intensity over the past one week (0–10 or 0–100)
(3) The individual item of the Von Korff pain intensity score that is equivalent to the VAS if it is available
(4) Roland Morris pain score (0–6). (Divided by 6 and multiplied by 10).

All measures will be scaled to 0–100 scale.

Combining average pain and pain intensity

For the analyses of pain, one of the following instruments from each trial, where available, was chosen (in descending order):

(1) Individual visual analogue scale (VAS) on average pain today
(2) average pain over the past one week
(3) the average pain item of the Von Korff pain intensity score that is equivalent to the VAS if it is available
(4) individual visual analogue scale (VAS) on pain intensity today (0–10 or 0–100)
(5) Pain intensity over the past one week (0–10 or 0–100)
(6) The summary score of the Von Korff pain intensity score
(7) Roland Morris pain score (0–6). (Divided by 6 and multiplied by 10).

All measures were scaled to 0–100 scale.

Analysis plan of outcome functional status

All of the RCTs but one in the repository had asked participant to rate or mark the functional status. Different functional status questionnaires were used.

For the analyses of functional status, the sum scores of the following instruments where analyzed separately:

- RMDQ: Only sum score of all studies will be analyzed
- ODI: Sum score of all studies will analyzed

Combined measure of functional status

For the analyses of combined measure of functional status, one of the following instruments from each trial, when available, was chosen (in descending order):

o RMDQ
p ODI
q Von Korff disability scale
r Other functional status questionnaires

One stage analysis

- The individual scores will be recoded into z-scores by subtracting the individual score from the mean score at baseline, and dividing the result by the mean standard deviation at baseline. Subsequently, the pooled z-scores will be used for further analyses.

Two stage analysis

- Standard mean difference will be calculated for each study.

Appendix 4: The GRADE approach to evidence synthesis

GRADE was used to evaluate the quality of the evidence for each primary outcome.

The quality of evidence is categorized as follows:

- *High* (⊙⊙⊙⊙): further research is very unlikely to change the confidence in the estimate of effect.
- *Moderate* (⊙⊙⊙○): further research is likely to have an important impact in the confidence in the estimate of effect.
- *Low* (⊙⊙○○): further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- *Very low* (⊙○○○): any estimate of effect is very uncertain.

The evidence was graded upon the following five domains (i.e. limitations/risk of bias, inconsistency, indirectness, imprecision, publication bias) in the following manner:

Limitations/risk of bias

Limitations in the study design refers to the way in which the various forms of bias may influence the estimates of the treatment effect.

We examined all studies for the following forms of bias:

- Selection bias (random sequence generation, allocation concealment, group similarities at baseline);
- Performance bias (blinding of participants and/or healthcare providers);
- Attrition bias (drop outs and intention-to-treat analysis);
- Detection bias (blinding of the outcome assessors and timing of outcome assessment);
- Reporting bias (selective reporting).

There is evidence that selection bias, specifically concealment of the allocation, and performance bias are most closely associated with treatment effect (Juni 2001; Savovic 2017).Therefore, we considered downgrading the quality of the evidence as follows:

- By one level when the majority of subjects (> 50%) came from studies with selection bias (specifically, the allocation concealment was not conducted properly) and performance bias was present;
- By two levels when the majority of subjects (> 50%) came from studies with selection bias (specifically the allocation concealment was not conducted properly) and performance bias and bias was present in one or more other category.

Inconsistency

Inconsistency refers to an unexplained heterogeneity of results. Widely differing estimates of the treatment effect (i.e. heterogeneity or variability in results) across studies suggest true differences in the underlying treatment effect. Inconsistency may arise from differences in the populations (e.g. patients treated for low-back pain in primary care may demonstrate a different treatment response than those treated in secondary or tertiary care; or those with non-specific low-back pain may demonstrate different effects as opposed to those with radiating pain), differences in the interventions (e.g. high-velocity SMT versus low-velocity SMT), or differences in the timing of the outcome measurements. The results of the second stage analysis will be used as this cannot be elicited from the one-stage analysis.

We considered downgrading the quality of the evidence as follows:

- *By one level*: when the heterogeneity or variability in results was large (e.g. I^2 statistic value > 50%, representing potentially substantial heterogeneity). I^2 statistic value will be collected from the two-stage analysis.
- *By two levels*: when the heterogeneity or variability in results was large AND there was inconsistency arising

from differences in the populations, interventions, or outcomes.

Indirectness

Indirectness refers to the generalizability of the findings. Indirectness may be a problem and diminish our confidence if the population, type of intervention, comparator, or outcome in the included randomized trials differs broadly from the research question being addressed in this review. In systematic review, study with mixed population studies (acute/subacute/chronic), studies which included a majority of subjects with radiating pain, or the majority of subjects were referred from a secondary or tertiary professional (or setting)) are included in the analysis. In IPD, generalizability in many instances were solved because many variables describing these issues were present in the raw data, for example by excluding patients from the analyses e.g. only including chronic patients or performing a moderator analysis leg pain vs no leg pain.

In cases where these data were not available, we considered downgrading the quality of the evidence as follows:

- *By one level*: when there is indirectness in only one area.
- *By two levels*: when there is indirectness in two or more areas.

Imprecision

Imprecision refers to limitations in the interpretation of the results when studies include relatively few participants and few events, leading to wide confidence intervals (CIs) surrounding the estimate of the effect, and thus resulting in uncertainty about the treatment effect.

For dichotomous outcomes, we considered imprecision for either of the following two reasons:

(a) There is only one study; when there is more than one study, the total number of events is less than 300 (a threshold rule-of-thumb value) (Mueller 2007).
(b) The 95% CI around the pooled effect includes both 1) no effect and 2) appreciable benefit or appreciable harm. The threshold for'appreciable benefit' or appreciable harm' is a relative risk reduction (RRR) or relative risk increase (RRI) greater than 25%.

For continuous outcomes, we considered imprecision for either of the following two reasons.

(a) There is only one study; when there is more than one study, the total population size is less than 400 (a threshold rule-of-thumb value) (Mueller 2007).

(b) The 95% CI includes no effect and the upper or lower confidence limit crosses an effect size of 0.5 or mean difference of 20 mm in either direction.

We considered downgrading the quality of the evidence as follows:

- *By one level*: when there is imprecision due to (a) or (b) for a continuous or dichotomous outcome.
- *By two levels*: when there is imprecision due to (a) and (b) for a continuous or dichotomous outcome.

Publication bias

Publication bias refers to bias introduced as a result of the selective publication of studies, typically leading to an underestimation of the effect from studies demonstrating a 'negative' effect which are under-reported.

We considered downgrading the quality of evidence as follows:

- *By one level*: when the funnel plot suggests publication bias.

Appendix 5 Search strategy

1. #1 MeSH descriptor Back explode all trees
2. #2 MeSH descriptor Buttocks, this term only
3. #3 MeSH descriptor Leg, this term only
4. #4 MeSH descriptor Back Pain explode tree 1
5. #5 MeSH descriptor Back Injuries explode all trees
6. #6 MeSH descriptor Low Back Pain, this term only
7. #7 MeSH descriptor Sciatica, this term only
8. #8 (low next back next pain)
9. #9 (lbp)
10. #10 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9)
11. #11 MeSH descriptor Musculoskeletal Manipulations explode all trees
12. #12 MeSH descriptor Chiropractic explode all trees
13. #13 manip*
14. #14 MeSH descriptor Osteopathic Medicine explode all trees
15. #15 osteopath*
16. #16 chiropract*
17. #17 (#11 OR #12 OR #13 OR #14 OR #15 OR #16)
18. #18 (#17 AND #10)
19. #19 (#18)

MEDLINE search strategy

1. Clinical Trial.pt.
2. randomized.ab,ti.
3. placebo.ab,ti.
4. dt.fs.
5. randomly.ab,ti.
6. trial.ab,ti.
7. groups.ab,ti.
8. or/1–7
9. Animals/
10. Humans/
11. 9 not (9 and 10)
12. 8 not 11
13. dorsalgia.ti,ab.
14. exp Back Pain/
15. backache.ti,ab.
16. (lumbar adj pain).ti,ab.
17. coccyx.ti,ab.
18. coccydynia.ti,ab.
19. sciatica.ti,ab.
20. sciatica/
21. spondylosis.ti,ab.
22. lumbago.ti,ab.
23. exp low back pain/
24. or/13–23
25. exp Manipulation, Chiropractic/
26. exp Manipulation, Orthopedic/
27. exp Manipulation, Osteopathic/
28. exp Manipulation, Spinal/
29. exp Musculoskeletal Manipulations/
30. exp Chiropractic/
31. manipulation.mp.
32. manipulate.mp.
33. exp Orthopedics/
34. exp Osteopathic Medicine/
35. or/25–34
36. 12 and 24 and 35
37. limit 36 to yr = "2007–2008"

EMBASE search strategy

1. Clinical Article/
2. exp Clinical Study/
3. Clinical Trial/
4. Controlled Study/
5. Randomized Controlled Trial/
6. Major Clinical Study/
7. Double Blind Procedure/
8. Multicenter Study/
9. Single Blind Procedure/
10. Phase 3 Clinical Trial/
11. Phase 4 Clinical Trial/
12. crossover procedure/

13 placebo/
14 or/1–13
15 allocat$.mp
16 assign$.mp
17 blind$.mp
18 (clinic$ adj25 (study or trial)).mp
19 compar$.mp
20 control$.mp
21 cross?over.mp
22 factorial$.mp
23 follow?up.mp
24 placebo$.mp
25 prospectiv$.mp
26 random$.mp
27 ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp
28 trial.mp
29 (versus or vs).mp
30 or/15–29
31 14 and 30
32 human/
33 Nonhuman/
34 exp ANIMAL/
35 Animal Experiment/
36 33 or 34 or 35
37 32 not 36
38 31 not 36
39 37 and 38
40 38 or 39
41 dorsalgia.mp
42 back pain.mp
43 exp BACKACHE/
44 (lumbar adj pain).mp
45 coccyx.mp
46 coccydynia.mp
47 sciatica.mp
48 exp ISCHIALGIA/
49 spondylosis.mp
50 lumbago.mp
51 exp Low back pain/
52 or/41–51
53 exp CHIROPRACTIC/
54 exp Orthopedic Manipulation/
55 exp Manipulative Medicine/
56 exp Osteopathic Medicine/
57 manipulation.mp
58 manipulate.mp
59 exp Orthopedics/
60 osteopathy.mp
61 or/53–60
62 40 and 52 and 61

CINAHL search strategy

Yields 44 for 2007–2008.

1. Randomized Controlled Trials.mp.
2. clinical trial.pt.
3. exp Clinical Trials/
4. (clin$ adj25 trial$).tw.
5. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).tw.
6. exp PLACEBOS/
7. placebo$.tw.
8. random$.tw.
9. exp Study Design/
10. (latin adj square).tw.
11. exp Comparative Studies/
12. exp Evaluation Research/
13. Follow-Up Studies.mp.
14. exp Prospective Studies/
15. (control$ or prospectiv$ or volunteer$).tw.
16. Animals/
17. or/1–15
18. 17 not 16
19. dorsalgia.ti,ab.
20. exp Back Pain/
21. backache.ti,ab.
22. (lumbar adj pain).ti,ab.
23. coccyx.ti,ab.
24. coccydynia.ti,ab.
25. sciatica.ti,ab.
26. exp SCIATICA/
27. spondylosis.ti,ab.
28. lumbago.ti,ab.
29. exp low back pain/
30. or/19–29
31. exp CHIROPRACTIC/
32. exp MANIPULATION, CHIROPRACTIC/
33. exp MANIPULATION, ORTHOPEDIC/
34. exp MANIPULATION, OSTEOPATHIC/
35. manipulation.mp.
36. manipulate.mp.
37. exp Manual Therapy/
38. exp ORTHOPEDICS/
39. exp OSTEOPATHY/
40. or/31–39
41. 18 and 30 and 40

International IPD-SMT group Bronfort G, Cecchi F, Ferreira ML, Gudavalli MR, Hidalgo B, Hondras MA, Hsieh CJ, Petersen T, Rasmussen-Barr E, Skillgate E, UK Beam Trial Team.

Availability of data and materials Not applicable.

Declarations

Conflict of interest SMR, ADJ, and AdeZ work in clinical practice as chiropractors and treat older adults with chronic low back pain. SMR reports grants from European Chiropractor's Union, grants from Netherlands Chiropractic Association, grants from Centre for Chiropractic Research Excellence, and grants from Belgian Chiropractic Association, during the conduct of the study.

Consent for publication Not applicable.

Ethical approval The study protocol was approved by the Scientific Review Board of the coordinating institution Vrije Universiteit Amsterdam and by the Ethical Committee of the VU University Medical Center Amsterdam. (Projectnr. 2015.544). Risk of bias assessment: Evaluation of the risk of bias was conducted by two independent reviewers (SMR, AdeZ). To adjudicate disagreement, a third reviewer (ADJ) was contacted.

References

1. Wong AY, Karppinen J, Samartzis D (2017) Low back pain in older adults: risk factors, management options and future directions. Scoliosis Spinal Disord 12:14. https://doi.org/10.1186/s13013-017-0121-3

2. St Sauver JL, Warner DO, Yawn BP, Jacobson DJ, McGree ME, Pankratz JJ, Melton LJ 3rd, Roger VL, Ebbert JO, Rocca WA (2013) Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. Mayo Clin Proc 88:56–67. https://doi.org/10.1016/j.mayocp.2012.08.020

3. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-allah F, Abdela J, Abdelalim A, Abdollahpour I, Abdulkader RS, Abebe Z, Abera SF, Abil OZ, Abraha HN, Abu-Raddad LJ, Abu-Rmeileh NME, Accrombessi MMK, Acharya D, Acharya P et al (2018) 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 392:1789–1858. https://doi.org/10.1016/s0140-6736(18)32279-7

4. Paeck T, Ferreira ML, Sun C, Lin CW, Tiedemann A, Maher CG (2014) Are older adults missing from low back pain clinical trials? A systematic review and meta-analysis. Arthritis Care Res (Hoboken) 66:1220–1226. https://doi.org/10.1002/acr.22261

5. Palacios-Cena D, Alonso-Blanco C, Hernandez-Barrera V, Carrasco-Garrido P, Jimenez-Garcia R, Fernandez-de-las-Penas C (2015) Prevalence of neck and low back pain in community-dwelling adults in Spain: an updated population-based national study (2009/10-2011/12). Eur Spine J 24:482–492. https://doi.org/10.1007/s00586-014-3567-5

6. Hartvigsen J, Frederiksen H, Christensen K (2006) Back and neck pain in seniors-prevalence and impact. Eur Spine J 15:802–806. https://doi.org/10.1007/s00586-005-0983-6

7. American Geriatrics Society Panel on Pharmacological Management of Persistent Pain in Older Persons? why is this repeated and where are the authors? (2009) Pharmacological management of persistent pain in older persons. J Am Geriatr Soc 57:1331–1346. https://doi.org/10.1111/j.1532-5415.2009.02376.x

8. Stewart Williams J, Ng N, Peltzer K, Yawson A, Biritwum R, Maximova T, Wu F, Arokiasamy P, Kowal P, Chatterji S (2015) Risk factors and disability associated with low back pain in older adults in low- and middle-income countries. Results from the WHO study on Global AGEing and Adult Health (SAGE). PLoS ONE 10:e0127880. https://doi.org/10.1371/journal.pone.0127880

9. Pohontsch NJ, Heser K, Loffler A, Haenisch B, Parker D, Luck T, Riedel-Heller SG, Maier W, Jessen F, Scherer M (2017) General practitioners' views on (long-term) prescription and use of problematic and potentially inappropriate medication for oldest-old patients—a qualitative interview study with GPs (CIM-TRIAD study). BMC Fam Pract 18:22. https://doi.org/10.1186/s12875-017-0595-3

10. Sibbritt DWAJ (2010) Back pain amongst 8910 young Australian women: a longitudinal analysis of the use of conventional providers, complementary and alternative medicine (CAM) practitioners and self-prescribed CAM. Clin rheumatol. https://doi.org/10.1007/s10067-009-1299-4

11. Rubinstein SM, de Zoete A, van Middelkoop M, Assendelft WJJ, de Boer MR, van Tulder MW (2019) 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 364:l689. https://doi.org/10.1136/bmj.l689

12. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, Costa LO, Foster NE, Grotle M, Koes BW, Kovacs FM, Lin CW, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Ostelo RW (2015) Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J 24:1127–1142. https://doi.org/10.1007/s00586-015-3892-3

13. Nascimento P, Costa LOP, Araujo AC, Poitras S, Bilodeau M (2019) Effectiveness of interventions for non-specific low back pain in older adults. A systematic review and meta-analysis. Physiotherapy 105:147–162. https://doi.org/10.1016/j.physio.2018.11.004

14. Amaral LKB, Souza MB, Campos MGM, Mendonca VA, Bastone A, Pereira LSM, Mascarenhas RO, Oliveira VC (2020) Efficacy of conservative therapy in older people with nonspecific low back pain: a systematic review with meta-analysis and GRADE recommendations. Arch Gerontol Geriatr 90:104177. https://doi.org/10.1016/j.archger.2020.104177

15. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, Rovers M (2015) Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. PLoS Med 12:e1001855. https://doi.org/10.1371/journal.pmed.1001855

16. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, Yusuf S (2015) The burden of disease in older people and implications for health policy and practice. Lancet 385:549–562. https://doi.org/10.1016/s0140-6736(14)61347-7

17. Jenks AD, Hoekstra T, Axen I, de Luca K, Field J, Newell D, Hartvigsen J, French SD, Koes B, van Tulder MW, Rubinstein SM (2020) Back complaints in the elders-chiropractic (BACE-C): protocol of an international cohort study of older adults with low back pain seeking chiropractic care. Chiropr Man Ther 28:17. https://doi.org/10.1186/s12998-020-00302-z

18. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF, Group P-ID (2015) Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. JAMA 313:1657–1665. https://doi.org/10.1001/jama.2015.3656

19. de Zoete A, de Boer MR, van Tulder MW, Rubinstein SM, Underwood M, Hayden JA, Kalter J, Ostelo R (2017) Rationale and design of an individual participant data meta-analysis of spinal manipulative therapy for chronic low back pain-a protocol. Syst Rev 6:21. https://doi.org/10.1186/s13643-017-0413-y

20. Schulz C, Evans R, Maiers M, Schulz K, Leininger B, Bronfort G (2019) Spinal manipulative therapy and exercise for older adults with chronic low back pain: a randomized clinical trial. Chiropr Man Ther 27:21. https://doi.org/10.1186/s12998-019-0243-1

21. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians (2017) Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American college of physicians. Ann Intern Med 166:514–530. https://doi.org/10.7326/M16-2367

22. Bernstein IA, Malik Q, Carville S, Ward S (2017) Low back pain and sciatica: summary of NICE guidance. BMJ 356:i6748. https://doi.org/10.1136/bmj.i6748

23. Bons SCS, Borg MAJP, Van den Donk M, Koes BW, Kuijpers T, Ostelo RWJG, Schaafstra A, Spinnewijn WEM, Verburg-Oorthuizen AFE, Verweij HA (2017) NHG-Standaard Aspecifieke lagerugpijn. Huisarts Wet 60(2):54–84

24. Cost B (2006) European guidelines for the management of low back pain. Eur Spine J 15:s125–s127. https://doi.org/10.1007/s00586-006-1066-z

25. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, Ferreira PH, Fritz JM, Koes BW, Peul W, Turner JA, Maher CG, Buchbinder R, Hartvigsen J, Cherkin D, Foster NE, Maher CG, Underwood M, van Tulder M, Anema JR, Chou R, Cohen SP, Menezes Costa L, Croft P, Ferreira M, Ferreira PH, Fritz JM, Genevay S, Gross DP, Hancock MJ, Hoy D, Karppinen J, Koes BW, Kongsted A, Louw Q, Öberg B, Peul WC, Pransky G, Schoene M, Sieper J, Smeets RJ, Turner JA, Woolf A (2018) Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 391:2368–2383. https://doi.org/10.1016/s0140-6736(18)30489-6

26. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, Bronfort G, van Tulder MW, Editorial Board of the Cochrane Back NG (2015) 2015 Updated method guideline for systematic reviews in the Cochrane back and neck group. Spine 40:1660–1673. https://doi.org/10.1097/BRS.0000000000001061

27. Kalter J, Sweegers MG, Verdonck-de Leeuw IM, Brug J, Buffart LM (2019) Development and use of a flexible data harmonization platform to facilitate the harmonization of individual patient data for meta-analyses. BMC Res Notes 12:164. https://doi.org/10.1186/s13104-019-4210-7

28. Riley RD, Lambert PC, Abo-Zaid G (2010) Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 340:c221. https://doi.org/10.1136/bmj.c221

29. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (2019) Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane

30. J. C (1988) Statistical power analysis for the behavioral sciences. https://doi.org/10.4324/9780203771587

31. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, Sterne JAC (2018) Association between risk-of-bias assessments and results of randomized trials in cochrane reviews: the ROBES meta-epidemiologic study. Am J Epidemiol 187:1113–1122. https://doi.org/10.1093/aje/kwx344

32. Bronfort G, Maiers MJ, Evans RL, Schulz CA, Bracha Y, Svendsen KH, Grimm RH Jr, Owens EF Jr, Garvey TA, Transfeldt EE (2011) Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. Spine J Off J N Am Spine Soc 11:585–598. https://doi.org/10.1016/j.spinee.2011.01.036

33. Cecchi F, Molino-Lova R, Chiti M, Pasquini G, Paperini A, Conti AA, Macchi C (2010) 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 24:26–36. https://doi.org/10.1177/0269215509342328

34. Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, Maher CG, Refshauge KM (2007) Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. Pain 131:31–37. https://doi.org/10.1016/j.pain.2006.12.008

35. Gudavalli MR, Cambron JA, McGregor M, Jedlicka J, Keenum M, Ghanayem AJ, Patwardhan AG (2006) A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. Eur Spine J 15:1070–1082. https://doi.org/10.1007/s00586-005-0021-8

36. Hondras MA, Long CR, Cao Y, Rowell RM, Meeker WC (2009) 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 Manip Physiol Ther 32:330–343. https://doi.org/10.1016/j.jmpt.2009.04.012

37. Hsieh CY, Adams AH, Tobis J, Hong CZ, Danielson C, Platt K, Hoehler F, Reinsch S, Rubel A (2002) Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. Spine 27:1142–1148

38. Petersen T, Larsen K, Nordsteen J, Olsen S, Fournier G, Jacobsen S (2011) 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 36:1999–2010. https://doi.org/10.1097/BRS.0b013e318201ee8e

39. Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I (2003) Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. Man Ther 8:233–241. https://doi.org/10.1016/s1356-689x(03)00053-5

40. Skillgate E, Vingard E, Alfredsson L (2007) Naprapathic manual therapy or evidence-based care for back and neck pain: a randomized, controlled trial. Clin J Pain 23:431–439. https://doi.org/10.1097/AJP.0b013e31805593d8

41. Underwood M (2004) United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of

physical treatments for back pain in primary care. BMJ 329:1377. https://doi.org/10.1136/bmj.38282.669225.AE

42. Dougherty PE, Karuza J, Dunn AS, Savino D, Katz P (2014) Spinal Manipulative therapy for chronic lower back pain in older veterans: a prospective, randomized, placebo-controlled trial. Geriatr Orthop Surg Rehabil 5:154–164. https://doi.org/10.1177/2151458514544956

43. de Zoete A, Rubinstein SM, de Boer MR, Ostelo R, Underwood M, Hayden JA, Buffart LM, van Tulder MW, Bronfort G, Foster NE, Maher CG, Hartvigsen J, Balthazard P, Cecchi F, Ferreira ML, Gudavalli MR, Haas M, Hidalgo B, Hondras MA, Hsieh CY, Learman K, McCarthy PW, Petersen T, Rasmussen-Barr E, Skillgate E, Verma Y, Vismara L, Walker BF, Xia T, Zaproudina N (2021) The effect of spinal manipulative therapy on pain relief and function in patients with chronic low back pain: an individual participant data meta-analysis. Physiotherapy. https://doi.org/10.1016/j.physio.2021.03.006

44. Hee SW, Mistry D, Friede T, Lamb SE, Stallard N, Underwood M, Patel S, Repository G (2021) Identification of subgroup effect with an individual participant data meta-analysis of randomised controlled trials of three different types of therapist-delivered care in low back pain. BMC Musculoskelet Disord 22:191. https://doi.org/10.1186/s12891-021-04028-8

45. Buchbinder R, Underwood M, Hartvigsen J, Maher CG (2020) The Lancet Series call to action to reduce low value care for low back pain: an update. Pain 161(Suppl 1):S57–S64. https://doi.org/10.1097/j.pain.0000000000001869

46. Cook CLK, Showalter C, Kabbaz V, O'Halloran B (2013) Early use of thrust manipulation versus non-thrust manipulation: a randomized clinical trial. Man Ther 18:191–198. https://doi.org/10.1016/j.math.2012.08.005

47. Wilkey AGM, Byfield D, McCarthy PW (2008) 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 14:465–473. https://doi.org/10.1089/acm.2007.0796

48. Hidalgo BPL, Hall T, Detrembleur C, Nielens H (2015) Short-term effects of Mulligan mobilization with movement on pain, disability, and kinematic spinal movements in patients with non-specific low back pain: a randomized placebo-controlled trial. J Manip Physiol Ther 38:365–374. https://doi.org/10.1016/j.jmpt.2015.06.013