SPINE An International Journal for the study of the spine, Publish Ahead of Print

DOI: 10.1097/BRS.0000000000003814

Moderators of 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, DC[1], Michiel R de Boer, PhD[1], Sidney M Rubinstein, DC, PhD[1], Maurits W van Tulder, PhD[1,2], Martin Underwood, MD,PhD[3,4], Jill A Hayden, DC, PhD[5], Laurien M Buffart, PhD[6,7], Raymond Ostelo, PT, PhD[1,7]

International IPD-SMT group:

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.

[1] Department of Health Sciences, Faculty of Science, Amsterdam Movement Science research institute, Vrije Universiteit, Amsterdam, The Netherlands
[2] Department Physiotherapy & Occupational Therapy, Aarhus University Hospital, Aarhus, Denmark
[3] Warwick Clinical Trials Unit, Warwick Medical School, The University of Warwick, Coventry CV4 7AL, UK
[4] University Hospitals of Coventry and Warwickshire, Coventry, CV2 2DX, UK
[5] Department of Community Health & Epidemiology, Dalhousie University, Halifax, Nova Scotia B3H 1V7, Canada
[6] Radboud UMC, Nijmegen, the Netherlands.
[7] Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Epidemiology & Biostatistics, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands.

Corresponding author
Annemarie de Zoete
Department of Health Sciences
Vrije Universiteit
De Boelelaan 1085, room WN U-454, 1081 HV Amsterdam, The Netherlands.
a.de.zoete@vu.nl

The manuscript submitted does not contain information about medical device(s)/drug(s).
European Chiropractic Union Research Fund (Contract No A.14.03) funds were received in support of this work.
Relevant financial activities outside the submitted work: grants, stocks, payment for lecture, travel/accommodations/meeting expenses.

Mini- Abstract
This individual participant data meta-analysis assessed the effect of spinal manipulative therapy for chronic low back pain. Based on this review, there is no evidence to suggest that specific patients characteristics are associated with a clinically better response to SMT as compared with other recommended treatments for chronic LBP.

Abstract
Study design: Individual participant data (IPD) meta-analysis
Objective. To identify which participant characteristics moderate the effect of spinal manipulative therapy (SMT) on pain and functioning in chronic LBP.
Summary of Background. The effects of SMT are comparable to other interventions recommended in guidelines for chronic low back pain (LBP); however, it is unclear which patients are more likely to benefit from SMT compared to other therapies.
Methods. IPD were requested from randomized controlled trials (RCTs) examining the effect of SMT in adults with chronic LBP for pain and function compared to various other therapies (stratified by comparison). Potential patient moderators (n=23) were a-priori based on their clinical-relevance. We investigated each moderator using a one-stage approach with IPD and investigated this interaction with the intervention for each time point (1, 3, 6 and 12 months).
Results. We received IPD from 21/46 RCTs (n= 4223)). The majority (12 RCTs, n=2249) compared SMT to recommended interventions. The duration of LBP, baseline pain (confirmatory), smoking and previous exposure to SMT (exploratory) had a small moderating effect across outcomes and follow-up points; these estimates did not represent minimally relevant differences in effects; for example, patients with less than one year of LBP demonstrated more positive point estimates for SMT vs recommended therapy for the outcome pain, (mean differences ranged from 4.97 (95% CI: -3.20 to 13.13) at three months, 10.76 (1.06 to 20.47) at six months to 5.26 (-2.92 to 13.44) at twelve months in patients with over a year LBP. No other moderators demonstrated a consistent pattern across time and outcomes. Few moderator analyses were conducted for the other comparisons because of too few data.
Conclusion. We did not identify any moderators that enable clinicians to identify which patients are likely to benefit more from SMT compared to other treatments.

Key words: Individual participant data, meta-analysis, spinal manipulative therapy, manipulation, low back pain, chronic pain, moderators, randomized clinical trial, subgroup analysis, mobilization
Level of Evidence: 2

INTRODUCTION

Low back pain (LBP) is the world's leading cause of disability.[1] Non-pharmacological approaches are the first choice of treatment.[2] The treatment options include spinal manipulation and mobilization which are used by a variety of heath care providers such as osteopaths, chiropractors and physiotherapists. These approaches can be used together or alone to treat patients with chronic LBP, and collectively defined as spinal manipulative therapy (SMT).

Many systematic reviews and meta-analyses have found that SMT is an effective treatment for patients with chronic LBP with a modest mean effect compared to other interventions.[3-8] Whilst SMT can relieve LBP in some patients, it is not effective for everyone; the number needed to treat is in the range of five to ten.[9] One potential explanation is that patients with 'non-specific' chronic LBP have different characteristics that influence the intervention effect, while another explanation can be the variation in duration, number and type of SMT. Relevant subgroups of patients with chronic LBP may exist that might benefit more or less from SMT.[10] The first step in identifying these subgroups is to examine which participant or treatment characteristics moderate the treatment effect (e.g. age, duration of LBP).[10-12] These moderators are typically not presented in the traditional meta-analyses[13] because aggregate data on relevant patient characteristics are often not available, are poorly reported, or derived and presented differently across studies. More importantly, if the results of subgroup analyses are reported, group averages or proportions are presented, which can result in ecological bias. The result is patient level intervention-covariate interactions are usually not examined or reported, even though they have the potential to better target the intervention.[14] Whilst some authors have presented appropriate analyses of treatment moderation, few trials are large enough to exclude important moderator effects.[10]

One way to test interactions of these characteristics with the intervention is to use Individual Participant Data (IPD). IPD provides much increased statistical power and allows for standardized analyses across studies, using direct derivation of information desired on an individual level, independent of whether and how it was reported in original publications. Therefore, IPD potentially allows identification of clinical characteristics of patients with chronic LBP that may moderate treatment effects.

The specific objective of this IPD meta-analysis is to:

- Identify individual participant characteristics measured at baseline that moderate the effect of SMT for pain and function at one, three, six and twelve months in adults with chronic LBP versus 1) recommended interventions; 2) non-recommended interventions; 3) sham SMT; 4) SMT + other intervention versus SMT only and 5) mobilization versus manipulation.

METHODS

This study was conducted and reported according to the Preferred Reporting Items of Systematic Reviews and Meta-Analyses for IPD (PRISMA-IPD) guidelines[15] (appendix eTable1, http://links.lww.com/BRS/B675). 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 protocol has been published previously.[16] The methodology presented in the current paper gives an overview of the moderator selection and analysis, while the eligibility criteria, search methods for identification of new trials (appendix eTable 11, http://links.lww.com/BRS/B675), risk of bias assessment, funnelplots (appendix eFig 1 and 2, http://links.lww.com/BRS/B675)collection and extraction of IPD are fully described in the published protocol.[16]

Types of outcome measures

Primary outcomes were pain (reported on a 0-10 or 0-100 NRS or VAS scale) and back-specific function on any scale e.g. Roland Morris Disability Questionnaire, Oswestry disability Index. All outcomes were self-reported and converted following decision rules (appendix eTable 4, http://links.lww.com/BRS/B675).

Moderators of treatment effect

Candidate moderators of treatment response were identified by the research team *a-priori* based on consensus (appendix eTable 5, http://links.lww.com/BRS/B675). In short, the selection of patient moderators was based on a specific rationale (e.g. understanding behavioral and sociocultural mechanisms by which response is modified or from prognostic research (treatment effect modification studies or prognostic factor research))[10-12] (see protocol[16]). Of 23 potential moderators identified; six were not analyzed, because data were insufficient, unavailable or only available at study level (patient preference/expectancy, comorbidities, alcohol use, income, non-specific LBP and for all treatment characteristics). The number and frequency of SMT treatments were measured at study level and not at patient level in the vast majority of the trials. The same accounts for type of SMT technique used. Therefore, contrary to the description in our protocol, we could not analyze moderating effects by these types of variables.

For psychological factors, analyses were only performed for combined depression scales. For other psychological scales, there were insufficient data.

Moderator analyses were classified into confirmatory or exploratory. Moderators in confirmatory analyses are those related to specific theory or evidence, while moderators in exploratory analyses relate to moderators for which no empirical evidence exists or for which a specific theory or mechanism is lacking. Our potential confirmatory moderators were age, gender, duration of LBP, psychological factors, treatment preference/expectation and baseline pain, function and quality of life. Other moderator analyses were exploratory (appendix eTable 5, http://links.lww.com/BRS/B675).[17] In both cases, the analytical technique was the

same, but for the interpretation of those confirmatory moderators the evidence was considered to be stronger.

Preparing data for moderator analyses

If data on a variable of interest were not available, we attempted to extract this information based on other data in the trial (e.g. information about employment was missing, but there was a variable on sick leave). Whenever possible, we used continuous data as presented, unless dichotomizing facilitated the translation of findings to clinical practice or was needed to meaningfully combine data across trials. The cut-off points were determined by consensus of the steering committee (ADZ, MRdB, SMR, MWT and RO)(Appendix eTable 5, http://links.lww.com/BRS/B675). For age, we used 65 years as a cut-off point. Additionally, duration of LBP was dichotomized into less than one year versus greater than one year. Similarly, physical activity was categorized into low (1 or less exercise sessions per week), medium (2-3 exercise sessions per week) and high (>3 exercise sessions per week) and subsequently dichotomized into low-medium or high. The choice of cut-off for physical activity was further evaluated in a sensitivity analysis (low vs medium-high).

For the outcome pain, all pain scores were converted to a 0-100 points scale following a decision rule (appendix eTable 4, http://links.lww.com/BRS/B675). 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 nominator. ($Z\ score = \frac{xi-\bar{x}}{SD}$). Analyzing 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, by multiplying the SMD with the population standard deviation (SD) of the studies measuring Roland Morris Disability Questionnaire ($SD_{pooled} = \sqrt{\sum_{i=0}^{n} \frac{(n_i-1)*S^2}{(n_i-1)}}$ n_i = sample size for each trial; S = standard deviation for each trial).

Data analysis

We studied moderators of intervention effects when three or more trials within a comparison had data on the moderator and the outcome at a specific time point. We used the following comparisons: 1) SMT versus recommended interventions including non-pharmacological treatment (e.g. exercise) and pharmacological treatment (e.g. NSAIDs, analgesics); 2) SMT versus non-recommended interventions (e.g. light massage, diathermy, ultrasound; 3) sham 'placebo' SMT; 4) SMT + intervention versus intervention alone; 5) high-velocity low-amplitude SMT versus low-velocity low-amplitude SMT (i.e. manipulation vs. mobilization).[16]

Potential moderators were analyzed using a one-stage random effect IPD meta-analyses. The baseline outcome, treatment, potential moderator and interaction between treatment and moderator were included as fixed effects. Study specific intercepts were also included as fixed effects. Random treatment and interaction effects were added to the model. (see protocol[16] and equation 2 in appendix eTable 6, http://links.lww.com/BRS/B675).[18] We performed these analyses for each time point and each moderator 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.[19] The within-study interaction explained the patient-level variation in treatment response, while the across-study interaction represented the moderator 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 effect of SMT vs comparison for the index group compared to a / the reference group (e.g. females compared to males).

We refrained from presenting stratified results for subgroups of moderator variables, because these include a combination of within- and across-study information because of differences in proportions of persons within the separate subgroups between studies.

Synthesis of evidence

Assessment of clinical relevance for the main effects analyses was defined as a small, medium or large difference and based upon the recommendations of the Cochrane Back and Neck group:[20]

In a consensus meeting with the project group, we discussed our results to determine whether a moderating effect was present. We considered a moderator effect to be present if the magnitude of the effect was at least half of our pre-specified clinically relevant main effects; i.e., more than 5-points (on a 100-point pain scale) or more than 0.25 for SMD on function and there was consistency in the direction of the moderators across three consecutive follow-up intervals for both pain and function. The arbitrary cut-offs of 5-points or 0.25 SMD were used to detect small differences within a moderator as SMT is a low intensity, low cost intervention.

As a crude method to guide interpretation and further synthesize evidence, we combined interaction effects with the main treatment effects, in case a moderator fulfilled the criteria for described above. Based on this, we assessed whether we could identify a clinically relevant treatment effect for potentially relevant moderators (e.g. interaction-effects around main effects near zero usually do not imply clinically meaningful effects within subgroups, while interaction effect reaching ten points on a 100 point scale does). We assumed that subgroup effects based on these interaction effects lie symmetrically around the main effects. For example, consider a moderator effect of ‘gender’ of -6 points on a 0-100scale, and a main effect of SMT of -8 points.

This would result in an approximated estimate of main treatment effect for men of -5 points (-8+3) and for women of 11 points (-8-3). These subgroup effects might indicate minimally relevant treatment effect for women, but not for men.

RESULTS

Characteristics of trials are presented in (Table 1), Risk of bias criteria and assessment are presented in the appendix eTable 2 and 3, http://links.lww.com/BRS/B675. For more details, see appendices and protocol.[16]

Identification of trials

In total, forty-three RCTs met the inclusion criteria, of which 21 (50%) provided data[21-41] (fig.1) from 4223 participants. Baseline characteristics were compared to the published results of the individual trials. In two trials, the results differed from the published results: one trial provided only data from participants who gave consent to share their data[37] while for the second trial, all relevant baseline moderator data of the participants were lost.[35]

Characteristics of study participants

Participant characteristics were fairly similar across all comparisons (Table 2 and appendix eTable 7, http://links.lww.com/BRS/B675). All trials except one[35] provided data on sex and age. The average age of the participants was 46.1 (SD 13.78) years, 54.4% were women.

For employment and BMI, moderator data were missing in 9.6% and 5.9 % of the participants respectively, while for all other moderators, data were missing in less than 2% of the participants.

Moderators of SMT for primary outcomes: pain and function

SMT vs recommended interventions

For most moderators, no moderating effects were identified except for the moderators described below. (Table 3, 4).

Confirmatory moderator analysis

For pain and/or function, we found a consistent moderator effect for duration of LBP. Patients with less than one year LBP showed more positive/less negative point estimates for SMT vs recommended therapy on pain, with MD of 4.97 (95% CI: -3.20 to 13.13) at three months, 10.76 (CI: 1.06 to 20.47) at six months and 5.26 (CI -2.92 to 13.44) at twelve months; for function: SMD were 0.07 (-0.29; 0.43) at one month, 0.02 (-0.30; 0.34) at three months; 0.19 (-0.02; 0.15) at six months to 0.13 (-0.25; 0.52) at twelve months) (Table 3, 4). These effects were small, except for pain at 6 months, which showed a moderate effect. Converted to a MD for the 24-point Roland Morris Disability Questionnaire, these moderating effects amount to 0.35 at one month, 0.10 at three months, 1.06 at six months and 0.78 at twelve months (Table 3, 4).

The direction of the main treatment effect of SMT versus recommended interventions for the outcome pain was in favor of SMT (e.g. six months: MD -5.56, 95% CI -9.63 to -1.50) (see appendix eTable 10, http://links.lww.com/BRS/B675). When adding the moderator effect of duration of LBP to the main treatment effect, the results may indicate minimally relevant effects, meaning that patients with shorter duration of LBP may benefit from SMT. For those with longer duration, SMT has similar benefit compared to recommended interventions.

For function, patients with pain score over 50 showed more positive/less negative point estimates for SMT vs recommended therapy on pain, with SMD of -0.20 (-0.36; -0.04) at one month, -0.20 (-0.37; -0.03) at three months, -0.22 (-0.39; -0.06) at six months and -0.14 (-0.33; 0.04) at twelve months. These effects were small.

Exploratory moderator analysis

For pain and/or function, we found a consistent moderator effect for smoking. Non-smokers showed more positive/less negative point estimates for SMT vs recommended therapy on pain, with MD of 3.19 (-3.20; 9.58) at one month, 2.45 (-3.68; 8.58) at three months, 6.02 (0.12; 11.92) at six months to 4.85 (-1.33; 11.03) at twelve months; for function: SMD were 0.24 (0.00; 0.48) at one month, 0.22 (-0.02; 0.47) at three months; 0.14 (-0.11; 0.38) at six months to 0.29 (0.02; 0.56) at twelve months) (see for conversion to the 24-point Roland Morris Disability Questionnaire Table 3, 4). These effects were small.

For pain and/or function, we found a consistent moderator effect for previous SMT for LBP. Patients that had no previous SMT showed more positive/less negative point estimates for SMT vs recommended therapy on pain, with MD of -3.97 (-13.12; 5.19) at one month, 5.34 (-3.25; 13.95) at three months, 6.86 (-1.63; 15.35) at six months to 15.59 (6.18; 24.99) at twelve months; for function SMD were -0.11 (-0.49; 0.26) at one month, -0.07 (-0.46; 0.32) at three months; 0.14 (-0.20; 0.49) at six months to 0.52 (0.09; 0.84) at twelve months) (Table 3, 4). This effect (i.e. patients that had no previous SMT improved more than patients that had previous SMT for the outcomes pain and function with SMT compared to other recommended treatments) was small, except for pain and functional status at twelve months, which showed a moderate effect. (Table 3, 4).

SMT vs non-recommended interventions, SMT as adjuvant therapy and Manipulation vs Mobilization

Ninety percent of the moderator analyses for these comparisons were not performed due to too few data. In the analyses performed (mainly age, sex and BMI), we found no consistent effect for any moderator. (Appendix eTables 8-9, http://links.lww.com/BRS/B675)

DISCUSSION

This study is the first large-scale IPD meta-analysis which attempted to identify potential moderators for those undergoing SMT for chronic LBP. In short ,the results suggest no substantive moderation in the effect of SMT compared to other interventions. We did, however identify (possible) small moderation effects for the following confirmatory moderators: duration of LBP and greater pain at baseline, and for the exploratory moderators smoking and previous exposure to SMT. However these effects were too small to be clinically-relevant. This suggests that targeting SMT, based upon individual characteristics examined in this study is not warranted at this time.

Analyses of moderator effects of SMT have rarely been performed and have largely been restricted to aggregate meta-analytic approaches, These different approaches make it difficult to compare those results to ours. Results from an earlier systematic review[10] indicated moderating effects of psychosocial and belief factors, expectations and baseline pain and disability.

Two earlier IPD studies evaluated moderators for other types of treatment for LBP and identified small effects for the following moderators: age, gender, BMI, no heavy physical demands, psychosocial factors, back pain disability, pain severity and medication use.[42 43] However, our analyses suggest only a weak moderating effect of baseline pain for functional

status in our confirmatory analysis. There are a number of reasons why the results of our moderator analyses might differ from the other studies.[10 43-45] Most importantly, these earlier studies examined various types of conservative treatments for LBP (e.g. cognitive behavioral therapy) and the comparisons were chosen differently than in our study.

An important difference of our IPD analysis compared to traditional aggregate meta-analyses is that we could adjust for covariates and were not dependent upon how these data were reported in the study publications. IPD allowed investigation for moderators in a more sophisticated and valid way. In the IPD analysis one can separate the between-study and the within-study interaction. The between-study interaction describes the moderation effects at study level. This is what is analyzed in a meta-regression or subgroup analyses in traditional aggregate meta-analysis. Results of these analyses can be severely affected by ecological biases.[19] The real interest lies in the within-study interaction, which describes the effects of covariates on the treatment effectiveness at the patient level.

Strengths and limitations: The most important strength is our large data set (i.e. 21 RCTs) from various countries resulting in a dataset which included many different moderators and thousands of patients of which over 2000 were in the SMT vs recommended intervention comparison. Most moderator analyses in this comparison included more than 500 patients provided from at least three trials. It has been suggested that this might be robust.[10] Far fewer patients were included in the moderator analyses for the other comparisons. Therefore, our (exploratory) moderator results should serve as a guide for future research only.

We collected a wide variety of moderators, but many moderators were measured differently across trials or were not measured at all. For example: duration of LBP was measured as a continuous variable in some trials and as a categorical variable. Only age, sex and BMI were measured similarly. This meant that in many instances we had to compromise our best detailed measures by categorizing the data, which led to loss of information. Importantly, there was a large diversity in frequency (1 to 6 times a week), duration (2 to 12 weeks) and number of treatments (2 to 36 (average of 8)) in included trials and these characteristics were measured at study-level in most trials. Therefore, the moderator analyses with treatment characteristics were not possible in contrast to what we planned in our protocol. A better understanding of the etiology of chronic LBP and key mechanisms involved in the effects of SMT would help to identify moderators.

We did not assess the effects of imputing missing data on outcomes and moderators. Methodology for imputing missing values in IPD meta-analysis is still in the developmental phase.[46 47] To our knowledge standard imputation methods for IPD meta-analysis of moderator effects have not been described in the literature. These models are especially challenging as they should result in valid estimates for the one-stage models we used that distinguish within-study and between-study interaction effects.

Additionally, we did not investigate multiple moderators in the same analysis as no evident clinically important moderators were found, although others[43] looked at multiple moderator at a lower level of statistical significance. At this moment our study clearly presents exploratory results to inform future studies.

Another challenge we encountered is the definition of clinical-relevance of the treatment moderator effects. For main effects, three-levels of clinical relevance (small, medium and large) are broadly used across systematic reviews, and are recommended by the Cochrane Back and Neck Review Group.[20 48] However, for moderator analysis, we think clinical relevance should not be defined by the same criteria as for the main treatment effects. Importantly, we interpreted the moderator effects considering hypothesized mechanisms and consistency of results across time and outcome measures. In summary, we used a consensus approach for the arbitrary cut-off points for drawing our conclusions to detect small difference within a moderator for low intensity, low cost intervention comparisons. This is subjective, but in our view the best method currently available.

Another potential limitation is selection bias. We included only 50% of the eligible trials, whch is comparable to other IPD studies.[49 50] However, the effect sizes, methodological quality and range of publication dates of studies where IPD was collected was comparable with the studies where no IPD was present. We also missed the data of the most recent trials as we only included trials until 2016, because collection of data for an IPD is time consuming as also seen in other IPD studies.[42] It took four years to collect and analyze the data, which is comparable to IPD meta-analyses in other fields.[43 51] When we updated our search May 4th, 2018, we found that the most recent trials were small in size, had few data on patient characteristics and were considered to have a high risk of bias.[52-56] Therefore, it is not likely that these most recent trials or the studies where IPD was not provided, would materially change our results.

The clinical implication of this IPD study is that based on the evidence to date there is no justification for using specific patient characteristics to target SMT for chronic LBP patients.

In addition to more detailed study of the etiology of chronic LBP and mechanism(s) of SMT, future initiatives should focus on standardizing the manner in which inclusion and exclusion criteria, outcomes and moderators are measured and reported.[54 57-59] This will facilitate an effective comparison of interventions across trials. Additionally, our wish is to form an international IPD repository of RCTs which have examined the effect of conservative treatment for LBP. This will provide an excellent resource for researchers with advantages such as the potential for future network meta-analysis and to standardize, safeguard and store data centrally. To facilitate this, we encourage researchers in future grant applications to obtain permission to share their data and to include costs of uploading their final data into a repository as well as permission from Research Ethical Committees and participants for sharieng these data. However three large IPD meta-analyses of non-pharmacological treatments for low back pain have failed to find any consistent and clinically important moderation effects indicates that this line of research is very unlikely to generate important finding to improve patient care.

CONCLUSION

Based on the current IPD analyses, there is no evidence for moderating effects of specific patient characteristics that enable clinicians to identify which patients are likely to benefit more from SMT compared to other treatments. Future research dealing with the effectiveness of SMT would benefit from shared procedures for including important treatment effect modifiers.

ABBREVIATIONS

IPD: individual participant data

LBP: low back pain

MD: mean difference

PRISMA-P: Preferred Reporting Items of Systematic Reviews and Meta-Analyses Protocol

RCT: randomized clinical trial

RR: relative risk

SD: standard Deviation

SMD: standard mean difference

SMT: spinal manipulative therapy

ACKNOWLEGDEMENTS

We thank Martin Roosenberg for the advice in writing the article and Lothar Kuijper and Trynke Hoekstra for helping with the analyses. We thank the World Federation of Chiropractic for receiving the NCMIC New Investigator Award at the WFC congress 2019 in Berlin.

Key points

- The effects of SMT are comparable to other interventions recommended in guidelines for chronic low back pain (LBP); however, it is unclear which patients are more likely to benefit from SMT compared to other therapies.
- Based on this review, there is no evidence to suggest that specific patients or treatment characteristics are associated with clinically better response to SMT as compared to other (recommended) treatments for chronic low back pain.
- This may well be a result of the great variation in reporting of potential treatment modifiers.
- 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.
-

REFERENCES

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017;390(10100):1211-59. doi: 10.1016/S0140-6736(17)32154-2

2. Foster NE, Anema JR, Cherkin D, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet* 2018;391(10137):2368-83. doi: 10.1016/S0140-6736(18)30489-6

3. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatments for Low Back Pain. Rockville (MD)2016.

4. Coulter ID, Crawford C, Hurwitz EL, et al. Manipulation and mobilization for treating chronic low back pain: a systematic review and meta-analysis. *Spine J* 2018;18(5):866-79. doi: 10.1016/j.spinee.2018.01.013

5. 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. doi: 10.1186/1471-2474-15-286

6. Rubinstein SM, van Middelkoop M, Assendelft WJ, et al. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database Syst Rev* 2011(2):CD008112. doi: 10.1002/14651858.CD008112.pub2

7. Ruddock JK, Sallis H, Ness A, et al. Spinal Manipulation Vs Sham Manipulation for Nonspecific Low Back Pain: A Systematic Review and Meta-analysis. *J Chiropr Med* 2016;15(3):165-83. doi: 10.1016/j.jcm.2016.04.014

8. Hidalgo B, Detrembleur C, Hall T, et al. The efficacy of manual therapy and exercise for different stages of non-specific low back pain: an update of systematic reviews. *J Man Manip Ther* 2014;22(2):59-74. doi: 10.1179/2042618613Y.0000000041

9. Froud R, Eldridge S, Lall R, et al. Estimating the number needed to treat from continuous outcomes in randomised controlled trials: methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. *BMC Med Res Methodol* 2009;9:35. doi: 10.1186/1471-2288-9-35

10. Gurung T, Ellard DR, Mistry D, et al. Identifying potential moderators for response to treatment in low back pain: A systematic review. *Physiotherapy* 2015;101(3):243-51. doi: 10.1016/j.physio.2015.01.006

11. Pincus T, Burton AK, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)* 2002;27(5):E109-20.

12. Pincus T, Kent P, Bronfort G, et al. Twenty-five years with the biopsychosocial model of low back pain-is it time to celebrate? A report from the twelfth international forum for

primary care research on low back pain. *Spine (Phila Pa 1976)* 2013;38(24):2118-23. doi: 10.1097/BRS.0b013e3182a8c5d6 [published Online First: 2013/08/24]

13. Rubinstein SM, de Zoete A, van Middelkoop M, et al. Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2019;364:l689. doi: 10.1136/bmj.l689

14. Debray TP, Moons KG, van Valkenhoef G, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods* 2015;6(4):293-309. doi: 10.1002/jrsm.1160 [published Online First: 2015/08/20]

15. Stewart LA, Clarke M, Rovers M, 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. doi: 10.1001/jama.2015.3656

16. de Zoete A, de Boer MR, van Tulder MW, et al. Rational and design of an individual participant data meta-analysis of spinal manipulative therapy for chronic low back pain-a protocol. *Systematic reviews* 2017;6(1):21. doi: 10.1186/s13643-017-0413-y

17. Pincus T, Miles C, Froud R, et al. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. *BMC Med Res Methodol* 2011;11:14. doi: 10.1186/1471-2288-11-14

18. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:c221. doi: 10.1136/bmj.c221

19. Hua H, Burke DL, Crowther MJ, et al. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. *Stat Med* 2017;36(5):772-89. doi: 10.1002/sim.7171

20. Furlan AD, Pennick V, Bombardier C, et al. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2009;34(18):1929-41. doi: 10.1097/BRS.0b013e3181b1c99f

21. Balthazard P, de Goumoens P, Rivier G, et al. 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. doi: 10.1186/1471-2474-13-162

22. Bronfort G, Hondras MA, Schulz CA, et al. 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. doi: 10.7326/M14-0006

23. Bronfort G, Maiers MJ, Evans RL, 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. doi: 10.1016/j.spinee.2011.01.036

24. Cecchi F, Molino-Lova R, Chiti M, 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. doi: 10.1177/0269215509342328

25. Cook C, Learman K, Showalter C, et al. Early use of thrust manipulation versus non-thrust manipulation: a randomized clinical trial. *Man Ther* 2013;18(3):191-8. doi: 10.1016/j.math.2012.08.005

26. Ferreira ML, Ferreira PH, Latimer J, 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. doi: 10.1016/j.pain.2006.12.008

27. Gudavalli MR, Cambron JA, McGregor M, 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. doi: 10.1007/s00586-005-0021-8

28. Haas M, Vavrek D, Peterson D, et al. 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. doi: 10.1016/j.spinee.2013.07.468

29. Hondras MA, Long CR, Cao Y, et al. 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. doi: 10.1016/j.jmpt.2009.04.012

30. Hsieh CY, Adams AH, Tobis J, 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.

31. 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. doi: 10.1097/AJP.0b013e31805593d8

32. 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.

33. Vismara L, Cimolin V, Menegoni F, et al. Osteopathic manipulative treatment in obese patients with chronic low back pain: a pilot study. *Man Ther* 2012;17(5):451-5. doi: 10.1016/j.math.2012.05.002

34. Walker BF, Hebert JJ, Stomski NJ, et al. Short-term usual chiropractic care for spinal pain: a randomized controlled trial. *Spine (Phila Pa 1976)* 2013;38(24):2071-8. doi: 10.1097/01.brs.0000435032.73187.c7

35. Wilkey A, Gregory M, Byfield D, et al. 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. doi: 10.1089/acm.2007.0796

36. Xia T, Long CR, Gudavalli MR, 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. doi: 10.1097/BRS.0000000000001373

37. Zaproudina N, Hietikko T, Hanninen OO, et al. Effectiveness of traditional bone setting in treating chronic low back pain: a randomised pilot trial. *Complement Ther Med* 2009;17(1):23-8. doi: 10.1016/j.ctim.2008.08.005

38. 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.

39. Petersen T, Larsen K, Nordsteen J, et al. 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. doi: 10.1097/BRS.0b013e318201ee8e

40. 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. doi: 10.1136/bmj.38282.669225.AE

41. Hidalgo B, Pitance L, Hall T, et al. 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. doi: 10.1016/j.jmpt.2015.06.013

42. Hayden JA, Wilson MN, Stewart S, 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 doi: 10.1136/bjsports-2019-101205

43. Patel S, Hee S, Mistry D, et al. Identifying back pain subgroups: developing and applying approaches using individual patient data collected within clinical trials. *Programme Grants Appl Res* 2016;4(10)

44. Beneciuk JM, Hill JC, Campbell P, et al. Identifying Treatment Effect Modifiers in the STarT Back Trial: A Secondary Analysis. *J Pain* 2017;18(1):54-65. doi: 10.1016/j.jpain.2016.10.002

45. Gilpin HR, Keyes A, Stahl DR, et al. Predictors of Treatment Outcome in Contextual Cognitive and Behavioral Therapies for Chronic Pain: A Systematic Review. *J Pain* 2017;18(10):1153-64. doi: 10.1016/j.jpain.2017.04.003

46. Jolani S, Debray TP, Koffijberg H, et al. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. *Stat Med* 2015;34(11):1841-63. doi: 10.1002/sim.6451

47. Resche-Rigon M, White IR. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. *Stat Methods Med Res* 2018;27(6):1634-49. doi: 10.1177/0962280216666564

48. Cohen J. Statistical power analysisfor the behavioral sciences. 2nd edn ed: Lawrence Earlbaum Associates, Hillsdale 1988.

49. Buffart LM, Kalter J, Sweegers MG, 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. doi: 10.1016/j.ctrv.2016.11.010

50. Panagioti M, Bower P, Kontopantelis E, et al. Association Between Chronic Physical Conditions and the Effectiveness of Collaborative Care for Depression: An Individual Participant Data Meta-analysis. *JAMA Psychiatry* 2016;73(9):978-89. doi: 10.1001/jamapsychiatry.2016.1794

51. Kalter J, Verdonck-de Leeuw IM, Sweegers MG, et al. Effects and moderators of psychosocial interventions on quality of life, and emotional and social function in patients with cancer: An individual patient data meta-analysis of 22 RCTs. *Psychooncology* 2018;27(4):1150-61. doi: 10.1002/pon.4648

52. Waqqar S, Shakil-Ur-Rehman S, Ahmad S. McKenzie treatment versus mulligan sustained natural apophyseal glides for chronic mechanical low back pain. *Pak J Med Sci* 2016;32(2):476-9. doi: 10.12669/pjms.322.9127

53. Ulger O, Demirel A, Oz M, et al. The effect of manual therapy and exercise in patients with chronic low back pain: Double blind randomized controlled trial. *J Back Musculoskelet Rehabil* 2017;30(6):1303-09. doi: 10.3233/BMR-169673

54. Samir SM, Zaky LA, Soliman MO. Mulligan versus Maitland mobilizations in patients with chronic low back dysfunction. *Int J Pharmtech Res* 2016;9(6):92-99.

55. Krekoukias G, Gelalis ID, Xenakis T, et al. Spinal mobilization vs conventional physiotherapy in the management of chronic low back pain due to spinal disk degeneration: a randomized controlled trial. *J Man Manip Ther* 2017;25(66-73) doi: 0.1080/10669817.2016.1184435

56. Sarker KK, Mohanty U, Sethi J. Effect of spinal manipulation on postural instability in patients with non specific low back pain. *Int J Pharm Bio Sci* 2017;7:992-9

57. Chiarotto A, Boers M, Deyo RA, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain* 2018;159(3):481-95. doi: 10.1097/j.pain.0000000000001117

58. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J* 2015;24(6):1127-42. doi: 10.1007/s00586-015-3892-3

59. Amundsen PA, Evans DW, Rajendran D, et al. Inclusion and exclusion criteria used in non-specific low back pain trials: a review of randomised controlled trials published between 2006 and 2012. *BMC Musculoskelet Disord* 2018;19(1):113. doi: 10.1186/s12891-018-2034-6

Figure 1 Flow diagram of study inclusion

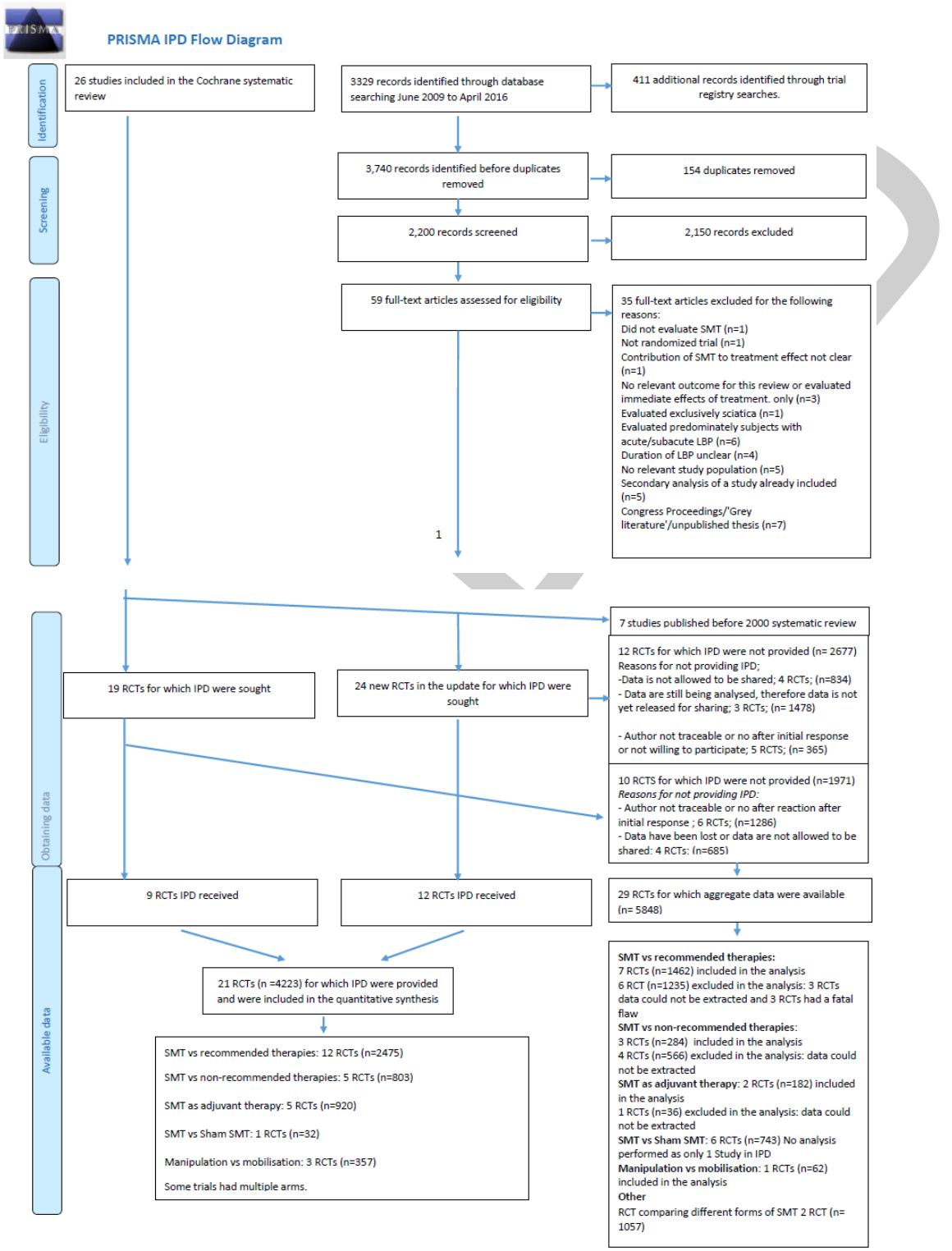

Table 1 Descriptives of studies evaluating the effects of SMT on outcomes included in the database (n=21) in alphabetical order of first author.

Author (year) *Acronym*	Country	N	Interventions	Duration of LBP According to inclusion criteria	Type of manipulator	Type of manipulation	Max no. treatments allowed and duration of treatment
Balthazard (2012)	Switzerland	42	1. Spinal manipulation Therapy plus active exercise (n = 22) 2. Detuned ultrasound plus active exercise (n = 20)	> 12 and < 26 weeks	Physiotherapist (n=1)	Manipulation and mobilization	8 over 4-8 weeks
Bronfort (2011)	USA	301	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
Bronfort (2014)	USA	192	1. Spinal manipulative therapy plus home exercise and advice (n=96) 2. Home exercise and advice (n=96)	> 4 weeks	Chiropractor (n=11)	Manipulation and mobilization	as many as 20 over 12 wks
Cecchi (2010)	Italy	210	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
Cook (2013)	USA	154	1. Thrust manipulation(n=77) 2. Non-thrust manipulation (n=77)	No restriction	Physiotherapist (n=17)	Manipulation or mobilization (depending upon grp. assignment)	1st 2 visit only afterwards clinician was allowed to choose technique they felt most beneficial for the patient
Ferreira (2007)	Australia	240	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	235	1. Flexion distraction mobilization (n=123) 2. Exercise therapy (n=112)	> 3 mo	Chiropractor (n = ?)	Mobilization (flexion-distraction)	16 over 4 wks
Haas	USA	400	1. 0 SMT (spinal manipulative therapy) + 18 LM (light	> 3 mo	Chiropractor	Manipulation or	18 over 6 wks

(2014)			massage) (n=100) 2. 6 SMT + 12 LM (n=100) 3. 12 SMT + 6 LM (n=100) 4. 18 SMT + 0 LM (n=100)		(n=12)	mobilization	
Hidalgo (2015)	Belgium	32	1. Spinal manipulative therapy (n=16) 2. Sham spinal manipulative therapy (n=16)	No restriction	Physiotherapist (n=1)	Mobilization	1 over 2 wks
Hondras (2009)	USA	240	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	206	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	350	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	47	1. Stabilizing training group (n = 24) 2. Manual therapy group (n = 23)	> 6 wks	Manual therapist (n = ?)	MOB	6 over 6 wks
Skillgate (2007)	Sweden	409	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	1334	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 =	(Essentially) > 3 wks	Chiropractor, osteopath or physiotherapist (n = 84)	Manipulation or mobilization	8 over 12 wks

			172) 6. Best care plus NHS manipulation plus exercise (n = 161)				
Verma (2013)	India	30	1. Exercise (n=15) 2. Lumbar mobilisation and exercise (n=15)	> 3 mo	Physiotherapist (n=?)	Mobilization	8 over 4 wks
Vismara (2012)	Italy	21	1. Osteopathic manipulation and Specific exercise (n=10) 2. Specific exercise (n=11)	> 6 mo	Osteopath (n=1)	Manipulation or mobilization	10 over 10 wks?
Walker (2013)	Australia	183	1. Sham group (n=91); 2. Usual chiropractic care group (n=92)	> 1wk	Chiropractor (n=8)	Manipulation or mobilization	2 over 2 wks
Wilkey (2008)	UK	63	1) Hospital pain clinic (n = 33) 2) Chiropractic treatment (n = 30)*	>3 mo	Chiropractor (n = ?)	Manipulation	16 over 8 wks
Xia (2015)	USA	192	1. Thrust spinal manipulation (n=72) 2. Non thrust spinal manipulation (n=72) 3. Control (n=48)	> 4 wks	Chiropractor (n=4)	Manipulation or mobilization	4 over 2 weeks
Zaproudina (2009)	Finland	73	1. Traditional bone setting (n = 36) 2. Physical therapy (n = 37)**	> 3 mo	Bone-setter (n = 8)	Mobilization	5 over 10 wks

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

*More patients data provided than published

** Only patient data used if patient consented to be included in our database, therefore less patients than published

Table 2: Patient characteristics at baseline for groups receiving SMT vs groups receiving recommended interventions

	SMT vs recommended interventions ***(m= 12; n=2475)***	
Demographic data	SMT	Recommended interventions
Age, mean (SD) years(m=11, n=2409)	47.13 (13.63)	47.18 (13.99)
Sex, n (%) female (m=11, n=2412)	667 (56.9)	684 (55.2)
BMI, mean (SD) (m=8, n=1434)	26.85 (5.12)	26.79 (5.10)
Ethnicity, n (%) white (m= 5, n=861)	409 (90.9)	388 (88.1)
Lifestyle factors		
Physical activity, n (%) (m= 6, n=824)		
Low (1 or less than once a week)	115 (31.9)	166 (35.9)
Medium (2-3x a week)	146 (40.4)	166 (35.9)
High (more than 3x a week)	100 (27.7)	131 (28.3)
Smoker, n (% non-smokers)	451 (79.5)	453 (74.8)
(m=6, n=1173)		
Alcohol use (%)	*	*
Socio-demographics		
Marital Status, n (%) married; living with a partner (m= 6, n=1173)	397 (69.0)	404 (67.6)
Level of Education, n (%) low/ middle (m= 7, n= 1672)	600 (68.0)	534 (67.6)
Income, n (%)	*	*
Employment status, n (%) at work (m= 9, n= 2126)	818 (77.9)	770 (71.6)
Nature and severity of LBP		
Duration of LBP, n (%) less than 12 months (m=7, n=1252)	121 (20.3)	149 (22.9)
Leg pain, n (%) (m= 5, n=1038)	320 (59.0)	281 (56.7)
Previous LBP treatment received, n (%) (m=5, n=930)	258 (27.7)	218 (23.4)
Previous physiotherapy for low back pain received, n (%) (m=5, n=771)	64 (8.3)	72 (9.3)
*Previous SMT for low back pain receive*d, n (%) (m=6, n=988)	209 (21.2)	111 (11.2)
Used medication for low back, n (%) (m=6, n=1018)	200 (19.6)	269 (26.4)
Non-specific, n (%)	*	*
Comorbidities	*	*
Type of treatment	*	*
Psychological factors	SMT	Control
Depression, n (%) (m= 5, n=1297)	43 (6.2)	75 (12.5)
Treatment preference/expectations	*	*
Primary outcomes		

Pain	SMT	Recommended interventions
Combined pain score at baseline, mean (SD), (m=12, n=2441)	49.47 (22.27)	49.75 (21.59)
Combined pain score at one month, mean (SD), (m=10, n=1948)	34.19 (22.95)	35.81 (23.91)
Combined pain score at three months, mean (SD), (m=9, n=1673)	27.92 (23.03)	32.12 (24.25)
Combined pain score at six months, mean (SD), (m=8, n=1321)	27.35 (23.12)	32.31 (23.90)
Combined pain score at 12 months, mean (SD), (m=10, n=*1816)*	31.80 (25.81)	33.32 (25.38)
Function		
RMDQ sum score at baseline, mean (SD), (m=9, n=2174)	8.99 (4.96)	10.07 (5.44)
RMDQ sum score at one month, mean (SD), (m=8, n=1760)	5.62 (5.02)	6.65 (5.37)
RMDQ sum score at three months, mean (SD), (m=8, n=1648)	4.81 (5.14)	5.52 (5.34)
RMDQ sum score at six months, mean (SD), (m=8, n=1348)	4.99 (5.44)	6.26 (5.95)
RMDQ sum score at 12 months, mean (SD), (m=7, n=1575)	5.44 (5.67)	6.16 (5.92)
Secondary outcomes		
SF36 Physical Component Scale of SF36 at baseline, mean (SD), (m=5, n= 1362)	40.69 (7.15)	41.06 (7.59)
SF36 Mental Component Scale of SF36 at baseline, mean (SD), (m=5, n= 1362)	43.83 (9.05)	45.08 (9.60)
Medication use at baseline, n (% medication use) (m =3, n= 668)	145 (21.7)	216 (32.3)

SD= standard deviation; m = number of studies; n = number of participants; * less than 3 studies or combining categories was not meaningful

Table 3: Moderator effects of SMT vs recommended interventions for pain. Within-study interaction (•) and 95% confidence intervals (CI) with the intervention effects of random-effect models adjusted for baseline using REML separating, between-study and within-study variation are presented.

Combined Pain scale	**Follow-up**			
	1 month	**3 months**	**6 months**	**12 months**
	• (95% CI)	**• (95% CI)**	**• (95% CI)**	**• (95% CI)**
	m; n	**m; n**	**m; n**	**m; n**
Demographic moderators				
Sex:	-1.63 (-5.02; 1.75)	-2.76 (-6.81; 1.28)	-5.11 (-9.69; -0.54)	-6.69 (-10.71; -2.67)
(reference: male)	9; 1859	8; 1592	7; 1270	9; 1740
Age *:	-1.74 (-6.03; 2.55)	-0.56 (-7.96; 6.84)	-9.42 (-17.65; -1.20)	2.60 (-5.27;10.47)
(reference: <65 years old)	9; 1859	8; 1590	7; 1268	6; 1042
Body mass index:	-1.96 (-7.30; 3.38)	3.78 (-3.10; 10.65)	4.36 (-2.28; 11.00)	6.01 (-0.84; 12.86)
(reference: <30)	6; 1060	5; 924	6; 1056	6; 1042
Ethnicity:	-1.96 (-10.08; 6.15) 4; 699	5.12 (-7.37; 17.60)	-0.08 (-13.39; 13.22) 3; 465	2.77 (-9.48; 15.02)
(reference: other than white)		3; 492		3; 469
Lifestyle factors				
Physical activity:	-0.46 (-7.45; 6.54)	-4.84 (-11.59; 1.92)	-4.67 (-11.47; 2.14)	-0.02 (-7.58; 7.54)
(reference: 3 or less a week)	4; 533	7; 721	4; 681	5; 661
(reference: 1 or less a week)				
	-3.96 (-10.07; 2.16)	-2.06 (-8.31; 4.18)	1.00 (-5.89; 7.60)	-3.22 (-10.35; 3.90)
	4; 533	7; 721	4; 681	5; 661
Smoker: smoker	3.19 (-3.20; 9.58)	2.45 (-3.68; 8.58)	6.02 (0.12; 11.92)	4.85 (-1.33; 11.03)
(reference: non-smoker)	5; 886	4; 866	5; 891	4; 806
Alcohol use	‡	‡	‡	‡
Socio-demographics				
Marital Status: (reference: not involved in relation)	2.60 (-3.11; 8.31)	1.39 (-4.47; 7.35)	1.84 (-4.19; 7.87)	-0.28 (-6.11; 5.57)
	6; 1076	4; 802	5; 834	4; 764
Employment status:	5.69 (0.32; 11.06)	0.88 (-5.72; 7.49)	5.12 (-2.94; 13.19)	0.96 (-4.93; 6.86)
(reference: not employed)	7; 1622	7; 1440	6; 1191	8; 1572
Level of Education: (reference: low or middle)	-0.42 (-4.62; 3.77)	-1.52 (-7.51; 4.48)	-3.31 (-14.50; 7.88)	-2.54 (-9.27; 4.19)
	6; 1377	5; 1117	4; 625	5; 1066
Income	‡	‡	‡	‡
Nature and severity of LBP				
Duration of LBP *:	-1.69 (-10.37; 7.00)	4.97 (-3.20; 13.13)	10.76 (1.06; 20.47)	5.26 (-2.92; 13.44)
(reference: < 1 year)	5; 876	4; 700	5; 875	5; 854
Radiation:	-4.13 (-9.68; 1.43)	-2.56 (-9.18; 4.06)	0.43 (-6.56; 7.41)	-4.41 (-11.33; 2.52)
(reference: no leg pain)	3; 649	4; 716	3; 636	4; 682

Previous LBP treatment received:	0.23 (-5.06; 5.53)	2.42 (-9.09; 4.25)	-0.36 (-7.18; 6.45)	-0.95 (-8.15; 6.25)
	5; 878	4; 675	3; 626	3; 622
(reference: no)				
Previous Physio for LBP:	-10.83 (-19.42; -2.25)	-3.38 (-11.51; 4.74)	2.41 (-6.64; 11.46)	-1.55 (-10.13; 7.02)
(reference: no)	3; 448	4; 679	4; 639	4; 624
Previous SMT for LBP:	-3.97 (-13.12; 5.19)	5.34 (-3.25; 13.95)	6.86 (-1.63; 15.35)	15.59 (6.18; 24.99)
(reference: no)	4; 629	4; 676	4; 636	4; 621
Previous medication for LBP:	1.32 (-3.47; 6.13)	-2.99 (-8.44; 2.46)	-2.43 (-8.22; 3.37)	-1.69 (-11.33; 7.96)
(reference: no)	5; 887	6; 902	4; 790	6; 862
Comorbidities	‡	‡	‡	‡
Psychological factors				
Depression *:	1.45 (-6.78; 9.68)	-1.06 (-11.23; 9.11)	‡	-1.20 (-13.23; 10.83)
(reference: no depression)	4; 1053	4; 860		4; 817
Treatment preference/expectations *	‡	‡	‡	‡
Primary/secondary outcomes at baseline as moderator				
Baseline pain score	-0.20 (-1.2; 0.80)	-0.70 (-2.10; 0.80)	-0.40 (-0.18; 1.10)	-1.1 (-2.90; 0.70)
per 10 points change *	10; 1922)	9; 1922)	8; 1922)	9; 1791)
Baseline function scales combined (z-score) *	-1.25 (-3.90; 1.39)	-0.77 (-3.43; 1.87)	-1.73 (-4.27; 0.82)	-0.90 (-4.07; 2.26)
	10; 1914	9; 1641	8; 1313	10; 1783
MCS	-0.09 (-0.40; 0.21)	-0.01 (-0.28; 0.26)	0.17 (-0.16; 0.49)	-0.19 (-0.48; 0.11)
	5; 1190	4; 1121	4; 681	4; 1046
PCS	0.19 (-0.36; 0.73)	-0.07 (-0.52; 0.38)	-0.40 (-1.07; 0.28)	0.09 (-0.50; 0.69)
	5; 1190	4; 1121	4; 681	4; 1046

A negative interaction coefficient indicates a more positive / less negative effect of SMT versus recommended therapies for the index group (e.g. females) as compared to the reference group (e.g. males).

• (95% CI) = within-study interaction and confidence interval: the mean difference in pain score for the specific moderator for SMT vs recommended therapies on scale from 0-100; m=number of studies; n= number of participants

* confirmatory moderator analysis

‡ not enough data

Table 4 Moderator effects of SMT vs recommended interventions for function. Within-study interaction (•) and 95% confidence intervals (CI) with the intervention effects of random-effect models adjusted for baseline using REML, separating between-study and within-study variation are presented.

Standardized mean difference of combined function scales	**Follow-up**			
	1 month	**3 months**	**6 months**	**12 months**
	• (95% CI)	**• (95% CI)**	**• (95% CI)**	**• (95% CI)**
	m; n	**m; n**	**m; n**	**m; n**
Demographic moderators				
Sex:	-0.03 (-0.16; 0.10)	0.06 (-0.09; 0.21)	0.01 (-0.15; 0.18)	-0.02 (-0.17; 0.14)
(reference: male)	9; 1876	10; 1837	8; 1439	9; 1775
Age *:	-0.03 (-0.17; 0.12)	-0.06 (-0.25; 0.13)	-0.23 (-0.45; -0.01)	-0.32 (-0.57; -0.07)
(reference: <65 years old)	9; 1876	10; 1835	8; 1437	9; 1773
Body mass index:	-0.07 (-0.26; 0.12)	-0.18 (-0.42; 0.07)	-0.06 (-0.30; 0.17)	0.15 (-0.14; 0.44)
(reference: <30)	6; 1047	7; 1139	7; 1225	6; 1055
Ethnicity: white vs other	-0.19 (-0.55; 0.17)	0.07 (-0.39; 0.53)	-0.23 (-0.65; 0.19)	0.03 (-0.38; 0.44)
(reference: other than white)	4; 691	5; 707	4; 630	3; 469
Lifestyle factors				
Physical activity:				
(reference: 3 or less a week)	-0.002 (-0.008; 0.002)	-0.002 (-0.009; 0.004)	-0.13 (-0.40; 0.14)	0.00 (-0.01; 0.01)
(reference: 1 or less a week)	4; 511	6; 739	4; 659	5; 676
	-0.11 (-0.33; 0.11)	-0.05 (-0.27; 0.18)	-0.13 (-0.34; 0.09)	-0.10 (-0.33; 0.14)
	4; 511	6; 739	4; 659	5; 676
Smoker: smoker	0.24 (0.00; 0.48)	0.22 (-0.02; 0.47)	0.14 (-0.11; 0.38)	0.29 (0.02; 0.56)
(reference: non-smoker)	5; 873	6; 1075	6; 1050	4; 801
DZ converted to a MD on the 24 point RMDQ scale	1.08	1.11	0.72	1.54
Alcohol use	‡	‡	‡	‡
Socio-demographics				
Marital Status:	-0.12 (-0.29; 0.05)	-0.09 (-0.29; 0.12)	-0.08 (-0.26; 0.10)	-0.12 (-0.33; 0.10)
(reference: not involved in relation)	6; 1066	6; 1016	6; 1005	4; 777
Employment status:	0.10 (-0.08; 0.27)	-0.17 (-0.39; 0.04)	0.06 (-0.16; 0.27)	-0.06 (-0.36; 0.23)
(reference: not employed)	7; 1657	8; 1665	7; 1377	8; 1606
Level of Education:	-0.17 (-0.37; 0.01)	-0.11 (-0.29; 0.08)	-0.07 (-0.27; 0.13)	-0.14 (-0.35; 0.05)
(reference: low or middle)	6; 1398	7; 1363	5; 796	5; 1106
Income	‡	‡	‡	‡
Nature and severity of LBP				

Duration of LBP *: >1 year	0.07 (-0.29; 0.43)	0.02 (-0.30; 0.34)	0.19 (-0.02; 0.39)	0.13 (-0.25; 0.52)
(reference: < 1 year)	5; 861	6; 910	6; 1031	5; 848
DZ converted to a MD on the 24 point RMDQ scale	0.35	0.10	1.06	0.78
Radiation: leg pain	-0.16 (-0.40; 0.80)	0.04 (-0.18; 0.27)	-0.01 (-0.22; 0.21)	-0.08 (-0.33; 0.17)
(reference: no leg pain)	3; 660	5; 911	4; 820	4; 681
Previous LBP treatment received: (reference: no)	0.09 (-0.11; 0.29)	0.11 (-0.12; 0.34)	0.02 (-0.21; 0.15)	0.12 (-0.14; 0.39)
	5; 884	5; 871	5; 812	4; 637
Previous Physio for LBP:	-0.08 (-0.63; 0.48)	0.04 (-0.33; 0.40)	0.18 (-0.19; 0.56)	0.11(-0.27; 0.50)
(reference: no)	3; 426	5; 696	4; 615	4; 634
Previous SMT for LBP:	-0.11 (-0.49; 0.26)	-0.07 (-0.46; 0.32)	0.14 (-0.20; 0.49)	0.52 (0.09; 0.95)
(reference: no)	4; 616	6; 885	5; 795	4; 631
DZ converted to a MD on the 24 point RMDQ scale	-0.49	-0.36	0.74	2.93
Previous medication for LBP: (reference: no)	-0.02 (-0.22; 0.19)	-0.11 (-0.33; 0.09)	-0.20 (-0.41; 0.004)	-0.11 (-0.39; 0.17)
	5; 887	6; 903	4; 790	6; 877
Comorbidities	‡	‡	‡	
Psychological factors				
Depression *:	0.28 (-0.03; 0.58)	0.33 (-0.02; 0.69)	0.13 (-0.26; 0.51)	0.02 (-0.39; 0.43)
(reference: no depression)	4; 1075	5; 1066	3; 505	4; 825
Treatment preference/expectations *	‡	‡	‡	‡
Primary/secondary outcomes at baseline as moderator				
Baseline function scales combined (z-score	-0.04 (-0.17; 0.09)	+0.01 (-0.10; 0.13)	-0.04 (-0.17; 0.08)	-0.05 (-0.17; 0.07)
	10; 1939	11; 1892	9; 1490	10; 18266
Baseline pain dichotomized *: (reference: baseline painless than 50)	-0.20 (-0.36; -0.04)	-0.20 (-0.37; -0.03)	-0.22 (-0.39; -0.06)	-0.14 (0.33; 0.04)
	10; 1932	11; 1882	9; 1506	10; 1806
	-0.90	-0.92	-1.00	-0.64
MCS	-0.00 (-0.10; 0.10)	-0.03 (-0.16; 0.08)	0.09 (-0.07; 0.25)	0.00 (-0.12; 0.12)
per 10 points change	5; 1204	6; 1358	5; 864	4; 1069
PCS	0.03 (-0.17; 0.23)	-0.06 (-0.18; 0.06)	-0. 13 (-0.32; 0.07)	0.00 (-0.22; 0.21)
per 10 points change	5; 1204	6; 1358	5; 864	4; 1069

A negative interaction coefficient indicates a more positive / less negative effect of SMT versus recommended therapies for the index group (e.g. females) as compared to the reference group (e.g. males).

• (95% CI) = within-study interaction and confidence interval of the interaction term: the difference in Z score for the specific moderator for function for SMT vs recommended therapies; m=number of studies; n= number of participants

* confirmatory moderator analysis

‡ not enough data