BMC Complementary Medicine and Therapies

SYSTEMATIC REVIEW

Open Access

Baseline individual factors associated with clinical outcomes in adults with non-specific low back pain following manual therapy: a systematic review

Gaetan Barbier[1,2,3*], Mathieu Picchiottino[3,4], Arnaud Delafontaine[1,2,5,6,8], Guillaume Goncalves[3,4,7], André Bussières[5], François Cottin[9] and Arnaud Lardon[1,2,3]

Abstract

Background Primary care providers consider the identification of patient subgroups as a high research priority. Unfortunately, evidence to support the benefit of treatments targeting subgroups of patients with NSLBP remains inconsistent. Specifically, little is known about baseline individual patient characteristics associated with optimal clinical improvement from manual therapy. This systematic review aims to identify baseline individual factors (BIFs), including patient characteristics, self-reported questionnaires, clinical examination, and ancillary test factors associated with clinical improvement (or lack of) among adult patients with Non-Specific Low Back Pain (NSLBP) following manual therapy.

Methods A systematic review of published evidence in Medline, Embase, Cochrane, Index To Chiropractic Literature, and CINAHL was conducted until April 2024. Studies included participants aged 18 years and over with NSLBP and without radiculopathy. Participants received manual therapies, including musculoskeletal manipulation/mobilization (spinal and extremities) and soft tissue therapy. We excluded mechanically assisted manipulations and interventions mainly involving exercise, education, and/or advice. Two independent assessors screened studies for inclusion, extracted data, and assessed risks of bias using the Quality In Prognosis Studies (QUIPS) Tools. A qualitative synthesis of findings was undertaken. BIFs were synthesized according to patient-reported outcomes measure domains: 1) pain intensity measures, 2) disability measures, 3) global perceived effect, and 4) other factors (e.g., satisfaction with care, total number of visits).

Results Data from 19 studies (reported in 21 articles) involving 4,689 participants were analyzed. Twelve studies reported pain intensity, 18 reported disability outcomes, and 4 reported patient's global perceived effect. Over 70% of the included studies had a high risk of confounding bias.

Included studies explored the potential association between clinical outcomes and 172 BIFs. BIFs were categorized into patient characteristics ($n=40$), self-reported questionnaire ($n=31$), clinical examination ($n=82$), and ancillary tests ($n=20$). Fourteen multivariate models explored the association with clinical improvement, and four others investigated the association with non-improvement. Findings were inconsistent across studies.

*Correspondence:
Gaetan Barbier
gbarbier@ifec.net
Full list of author information is available at the end of the article

Conclusion Using BIFs in clinical practice to predict clinical outcomes following manual therapy treatment appears to be premature. Future studies should aim to replicate the results and differentiate prognostic factors from treatment effect modifiers.

Trial registration CRD42019131416.

Keywords Low back pain, Musculoskeletal manipulations, Systematic review, Clinical outcomes, Population characteristics, Physical examination, Surveys and questionnaires, Prognosis factors, Treatment effect modifier

Introduction

In 2019, low back pain represented the leading cause of years lived with disability among all age groups and was one of the top ten causes of disability-adjusted life-year [1]. Most cases of low back pain are classified as non-specific (NSLBP) [2, 3]. Several contributors influence both individuals' pain experiences and disabilities associated with NSLBP, including biopsychosocial factors, genetic factors, comorbidities, and pain-processing mechanisms [2].

Despite the availability of high-quality clinical practice guidelines supporting multimodal management for NSLBP, many patients do not receive optimal care [4–7]. While previous research emphasized identifying subgroups of patients who may benefit from specific interventions [8], recent consensus has shifted. According to Dionne et al. (2022) [9], although stratification and personalized care remain relevant, the identification of patient subgroups is no longer considered a top research priority. Instead, the focus has moved toward improving self-care strategies and stimulating self-reliance or identifying the best strategies for treating LBP. In this context, understanding which baseline individual factors predict outcomes following manual therapy appears essential to advancing therapeutic decision-making and the development of individualized treatment strategies.

Moderators of a favorable response from care delivered by manual therapists (e.g., manipulation, exercise, cognitive behavioral therapy, and acupuncture) to NSLBP patients include younger age, having a sedentary occupation, a higher pain intensity at baseline, a greater positive expectation from care, and having completed over ten years of schooling [10]. In contrast, users of narcotic medication may benefit less from manual therapy [10]. A recent meta-analysis by de Zoete et al. [11] also showed that spinal manipulative therapy is more effective for patients with recent onset low back pain (< 1 year) compared to other therapies. In addition, a review by Vidal et al. (2024) identified over 200 prediction models developed to estimate outcomes in patients with LBP receiving conservative treatment, including manual therapy [12]. While a few models demonstrated acceptable discrimination and calibration, the overall risk of bias was high, primarily due to inadequate variable selection strategies and insufficient reporting of performance metrics. Consequently, the authors concluded that existing models are not yet recommended for clinical decision-making and emphasized the need to prioritize external validation of promising models over the development of new ones.

Although these reviews explored treatment effect modifiers and prediction models, no systematic review of the literature has comprehensively explored Baseline Individual Factors (BIF) associated (i.e., prognostic factor and treatment effect modifiers) with response to manual therapy treatment only (i.e., manipulation, mobilization, and soft tissue techniques) in people with NSLBP. According to Hayden et al., in the framework to determine prognostic factors, three consecutive phases should be considered [13]: i) an exploration phase, ii) a confirmation phase, and iii) an understanding phase. The first step to be able to identify a treatment-based subgroup is to explore baseline factors associated with a favorable or unfavorable clinical response [14, 15]. These factors may be prognostic factors, defined as factors that can predict the course of a specific condition over time, or treatment effect modifiers that can predict treatment response [16].

Aim

This study aimed to determine if baseline individual factors, including patient characteristics, self-reported questionnaires, clinical examination, and ancillary tests, are associated with clinical improvement (or lack of) among adult patients with NSLBP following manual therapy.

Methods

This systematic review was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Statement [17]. The review protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) on August 7, 2019 – CRD42019131416.

Eligibility criteria

Inclusion criteria followed the Participants, Exposure, Comparator, and Outcomes (PECO) framework [18], and study design, intervention, and language.

Participants

Adults (≥ 18 years) with NSLBP without radiculopathy. Studies had to specify that people with radiculopathy were excluded.

Exposure

BIFs included patient characteristics (e.g., demographics, socioeconomic status, education), self-reported questionnaires (e.g., disability, pain, quality of life, return to work), clinical examination (e.g., tests), and ancillary test factors such as multifidus activation or stiffness. In this review, the term 'baseline individual factor(s)' refers to both possible prognostic factors (defined as factors predictive of patient outcomes) and treatment effect modifiers (defined as factors predictive of treatment effects) [16].

Comparators

We included dichotomic variables as defined by the study authors (e.g., female vs. male; Visual Analogue Scale (VAS) >2/10 vs. VAS ≤2/10) and continuous variables used to estimate the dose–effect of the potential association.

Outcomes

All relevant clinical outcomes, including pain, disability, and global perceived effect, were eligible for inclusion.

Interventions

Interventions with a focus on manual therapy alone (including musculoskeletal manipulation, mobilization, and soft tissue therapy) or in combination with other conservative therapies. Manual therapy is commonly delivered by physiotherapists, osteopaths, chiropractors, massage therapists, and occasionally occupational therapists [19] to manage NSLBP [5]. Manual therapy can be defined as "the use of hands-on technique with therapeutic intent" [20] and includes a range of techniques such as manual traction, manipulation, mobilization, and soft tissue techniques, used alone or a combination of these.

Design

Cohort studies (prospective or retrospective), randomized controlled trials, secondary analysis of randomized controlled trials, or systematic reviews.

Language

Studies published in the English or French languages.

We excluded cross-sectional studies, editorials, letters, and commentaries. We also excluded studies on specific spinal pathologies (e.g., vertebral fracture, malignancy, spinal infection, axial spondylarthritis, cauda equina syndrome) and conditions known to cause LBP (e.g., pregnancy-related LBP). Mechanically assisted manipulations were excluded, as they were not performed directly by the therapist's hand. Interventions involving mainly exercise, education, and/or advice were also excluded.

Information sources

Five electronic databases were systematically searched from inception to June 2019, including Medline, Embase, Cochrane Central Register of Controlled Trials, Index To Chiropractic Literature (ICL), and CINAHL. We updated the search on April 24, 2024.

Search strategy

The search strategy was developed in MEDLINE in consultation with a health sciences librarian, adapted for the other databases, and reviewed by a second librarian using the Peer Review of Electronic Search Strategies (PRESS) Checklist [21, 22]. Search terms consisted of subject headings specific to each database (e.g., MeSH in PubMed) and free text words relevant to manual therapy, low back pain, outcome and BIFs related to patient characteristics, self-reported questionnaires, clinical examination, and ancillary tests (Additional Table S1).

Selection process

All citations identified by the search were exported into EndNote, and automatization tools and reviewers removed duplicates. In two phases, two independent reviewers (GB-AL) screened articles using a standardized pre-piloted Excel spreadsheet. In Phase 1, pairs of reviewers independently screened titles and abstracts for possibly relevant citations based on the eligibility criteria. In phase 2, four pairs of reviewers (GB-AL, GB-MP, GB-GG, GB-AD) applied the same eligibility criteria for full-text articles. Any disagreement was resolved by discussion, and a third reviewer was involved if a consensus could not be reached. Study authors were contacted if the information about the participants or the treatments was unclear.

Study risk of bias assessment

The same four pairs of reviewers independently applied the QUality In Prognosis Studies tool (QUIPS) [23] recommended by the Cochrane Prognosis Methods Group to assess the risk of bias in prognosis studies. The QUIPS tool is composed of six main domains that should be critically appraised: (1) study participation, (2) study attrition, (3) prognostic factor measurement, (4) outcome measurement, (5) study confounding, and (6) statistical analysis and reporting. An initial pilot screening of two non-included papers was used to prepare the reviewers for using QUIPS. To determine the risk of bias for each sub-item and item, we proceeded according to Grooten

et al.'s proposed framework in rehabilitation domain [24]. Original articles describing the protocol or the primary studies were consulted for secondary analyses of RCTs or cohorts.

Data extraction, management and analysis

The lead author (GB) extracted data from the included articles and constructed evidence tables. A second reviewer (AL) independently checked the extracted data. Any disagreements were resolved through discussion and a third reviewer if needed.

Extracted information included general study characteristics (first author, year of publication, geographical location, study design, and statistical approach), sample size, age, intervention (treatment modality and therapist), outcome measures for pain, disability, global perceived effect, and others. We extracted unadjusted and adjusted estimates for each variable (RRs, ORs, Coefficient β Value, Student's t-test) and for each model (stepwise logistic regression, positive likelihood ratio, backward stepwise logistic regression modeling, stepwise hierarchical linear regression, multivariable logistic regression, linear mixed-effects regression models), and measures of precision (confidence intervals, *p*-values, standard deviations, standard errors). We contacted the authors of the studies if any information or data was missing.

All BIF-associated and all clinical outcomes were extracted from studies. No time restrictions were retained.

Data synthesis

A qualitative synthesis of findings was carried out following Synthesis Without Meta-analysis (SWiM) recommendations [25]. Each BIFs (i.e., possible prognostic factors and treatment effect modifiers) was synthesized according to four patient-reported outcome measures (PROMs) domains: 1) pain intensity; 2) disability; 3) Global perceived effect (questionnaires related to the patient's perception such as Global Perceived Effect (GPE)); and 4) others (questionnaires related to other domains not covered in the above domains such as number of visits, or Fear Avoidance Belief Questionnaire (FABQ)).

We also report eligible studies according to the phases described by Hayden et al. [13]: i) *an exploration phase* to provide hypothesis-generating evidence by measuring the existence of a relationship between potential(s) factor(s) and the clinical outcome. ii) *A confirmation phase testing independent associations and providing evidence supporting a prognostic factor's independent effect by measuring the strength of the prognostic relationship* between a factor and a given clinical outcome while controlling for alternative explanations. iii) *An understanding phase* to better understand prognostic pathways and provide evidence supporting the prognostic factor's mechanism(s) of action on the outcome by examining the role of each prognostic factor and the process by which it impacts a given clinical outcome.

If we observed a significant association in both univariate and multivariate analyses, we only reported the result from the multivariate analysis. We did not perform a meta-analysis due to the heterogeneity of scientifically admissible studies.

All studies meeting the inclusion criteria were included in the synthesis and were summarized according to PROMs domains as described above. No data conversion was performed.

Protocol deviations from PROSPERO registration

We elected to use the QUIPS tool designed to assess the risk of bias in prognostic studies instead of the AMSTAR-2 and SIGN checklists conceived for various types of reviews.

Results

Study selection

Our search yielded 17,107 citations, with two additional citations found by citation tracking. Of those, 4437 duplicates were removed, leaving 12,672 citations screened for title and abstract (Phase I) and 258 full text reviews (Phase II). Nineteen studies reported in 21 articles were critically appraised and included in the qualitative analysis [26–46] (Fig. 1).

Study characteristics

We included 13 secondary analyses of Randomized Controlled Trials (RCT) [29, 30, 33–39, 42, 43, 45, 46], six reports of cohort studies [26–28, 32, 41, 44], one non-randomized controlled study [40] and one retrospective study outcome-based analysis [31]. 4,689 participants were included, with drop-out rates reported across studies ranging from 0% to 25.4%. The participants' ages ranged between 18 and 81 years old; two reports did not specify the age of the participants [31, 34]. All included articles were published in English between 2000 and 2021. Studies originated from USA (n =13) [26–30, 36–42, 46], Canada (n= 2) [31, 43], Australia (n= 1) [35], Denmark (n= 1) [45], Italy (n= 1) [34], Netherlands (n= 1) [44], UK (n= 1) [33], and Switzerland (n= 1) [32]. Eight studies had a follow-up time of less than one month [26, 27, 29, 41, 43, 45], seven between one and three months [28, 30, 32, 33, 35, 42, 44, 46], three had a follow-up between six months and one year [31, 34, 36], and three studies followed participants to discharge [37–39]. In two studies [37, 39], patients were discharged once the clinician felt the patient had met their maximal improvement within the current treatment program. There were

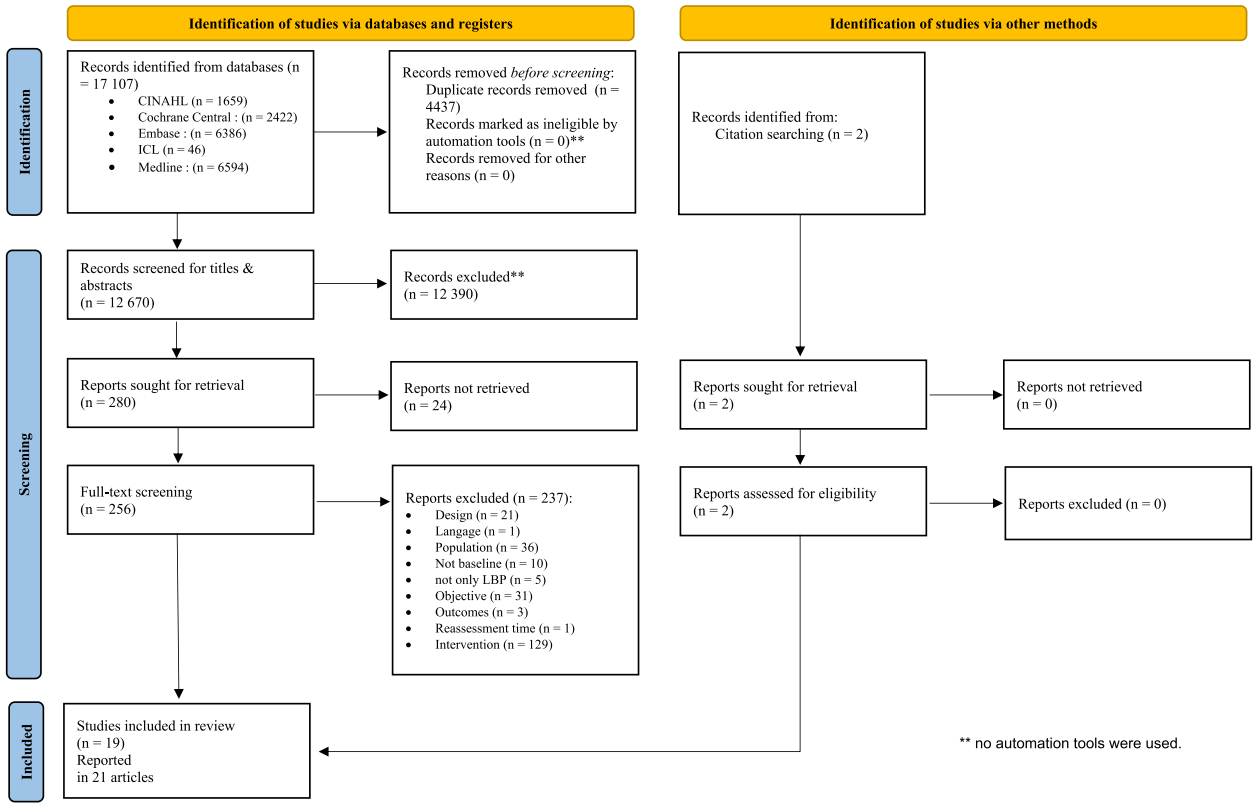

Fig. 1 Flow diagram

no restrictions on total visits for each patient enrolled in this trial. Discharge was not specified in one other study [38].

Twenty of the 21 reports related to phase 1 explanatory information [26–40, 42–46], and one report related to phase 3 understanding and testing of the prognostic pathway [41].

Most included studies ($n=$ 18) focused on clinical change measuring disability using self-reported questionnaires, while less than 25% ($n=$ 5) focused on the overall perception of perceived change. For Patient Reported Outcomes Measures (PROMS), 12 studies reported results on pain intensity using a VAS ($n=$ 5) [29–31, 35, 42] or a Numerical Pain Rating Scale (NPRS, $n=$ 7) [32, 36, 37, 39, 44–46]. Eighteen studies reported results on disability using the Oswestry disability index (ODI, $n=$ 8) [26, 27, 31, 37, 39, 41, 43, 45], the modified ODI (mODI, $n=$ 1) [40], Roland Morris disability questionnaire (RMDQ, $n=$ 7) [28, 33–36, 42, 46], Bournemouth questionnaire (BQ, $n=$ 1) [32], Pain Disability Inventory (PDI, $n=$ 1) [44] and the Patient Specific Functional Scale (PSFS, $n=$ 1) [35]. Four studies reported patients' global perceived effect outcomes, two measured by perception of extent of recovery [37, 39], one by global perceived effect [35], and one by mODI/NPRS/GRoC score [38]. Studies also report other outcome measures using the work subscale of fear-avoidance beliefs questionnaire [37], short form 36 health survey general (SF36) [42], Satisfaction with back care using a five-point Likert scale [42], use of co-treatment [42], the number of lost workdays because of LBP [42], days in care [37] and total visits and days in care [37]. Key characteristics of the studies are described in Table 1.

Risk of bias in studies

The risk of bias assessment for each study is reported in Table 2 and Fig. 2.

Figure 2 reports the overall risk of bias assessment for each of the six domains across all included studies. Over 90% of the studies had a low to moderate risk of bias for study participation, study attrition, prognostic factor measurement, outcome measurement, statistical analysis, and reporting. In contrast, over 70% of the included studies had a high risk of bias for study confounding.

Results of individual studies

In total, 172 BIFs were studied, of which 41 related to patient characteristics, 31 to the self-reported questionnaires, 82 to the clinical examination, and 18 to the ancillary tests category. See Additional Table S2.

Table 1 Study characteristics

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Annen M. (2018—Switzerland) [32]	Prospective Cohort	1 week 1 month 3 months	112 (0.0%)	18–81	Chiropractic treatment—specific high-velocity, low-amplitude spinal manipulation at the level of pain for most patients and in the application of passive therapies, such as cold application and muscular techniques. [Chiropractic students (final year)]	BQ, Change score (baseline—follow-up score) NPRS, Change score (baseline—follow-up score)
Burns S. (2018—USA) [38]	Secondary analysis—RCT [58]	Discharge	90 (20.0%)	≥ 18	The LBP group received only pragmatic LBP treatment, which reflected guideline-oriented care for LBP. The LBP + HIP group received both pragmatic LBP treatment plus a prescriptive set of manual therapy and therapeutic exercises targeting both hips. Participants in both groups received an LBP-oriented home exercise program that was pragmatically derived from their primary impairments but did not focus on the hips. [Physical therapist]	mODI/NPRS/GROC score To be considered as recovered, a participant needed to score ≤ 10% on the mODI and ≤ 2 on the NPRS at discharge and record a GROC score ≥ + 4 at both 2 weeks and discharge (average of 7.95 [± 4.68]). Participants had to achieve all 4 of these variables to be considered recovered
Cecchi F. (2012—Italy) [34]	Secondary analysis—RCT [59]	1 year (discharge)	210 (2.4%)	/	Spinal manipulation was performed according to the manual medicine approach described by Robert Maigne. The whole spine was examined by static and dynamic assessment; treatment consisted of vertebral direct and indirect mobilization and manipulation, with associated soft tissue manipulation, aimed at restoring the physiological movement in the dysfunctional vertebral segment(s). Patients assigned to spinal manipulation received 4–6 (as needed) weekly sessions of 20 min each for a total of 4–6 weeks of treatment (80–120 min of treatment altogether) [Physiotherapist]	RMDQ Non-responders: RMDQ score improved less than 2.5 on discharge compared to their baseline score

Table 1 (continued)

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Cook C. (2013—USA) [39]	Secondary analysis—RCT [47]	Discharge	154 (3.2%)	≥ 18	A comprehensive rehabilitation approach that included 1 of 2 forms of manual therapy (thrust or nonthrust manipulation) for the first 2 visits only, followed by physical therapist–directed care after the initial 2 visits. Both groups in the study received a manual therapy approach. [Physical therapist]	NPRS ≥ 2,5 points change A linear change score ODI 50% reduction of the ODI Total visits total number of visits used to measure the quantity of treatment (cutoff: ≤ 6 visits) Perception of the extent of recovery Self-report of the extent of recovery (0%–100%) was used to measure the perception of recovery. The question was stated as "What percent, 0 percent (meaning not at all) to 100 percent (meaning totally recovered) do you feel that you have recovered at this point?"; Self-reported extent of recovery of ≥ 75%
Donaldson M. (2013—USA) [37]	Secondary analysis—RCT [47]	Discharge	154 (3.2%)	≥ 18	A comprehensive rehabilitation approach that included 1 of 2 forms of manual therapy (thrust or nonthrust manipulation) for the first 2 visits only, followed by physical therapist–directed care after the initial 2 visits. Both groups in the study received a manual therapy approach. [Physical therapist]	NPRS A linear change score ODI A linear change score and 50% reduction of the ODI Total visits – used to measure the quantity of treatment Days in care – used to capture the temporal duration of care FABQ–Work subscale Perception of the extent of recovery—"What percent, 0 percent (meaning not at all) to 100 percent (meaning totally recovered), do you feel that you have recovered at this point?"
Ferreira M. (2009—Australia) [35]	Secondary analysis—RCT [60]	8 weeks 12 sessions	240 (20.4%)	18–80	Joint mobilization or manipulation techniques were applied to the subject's spine or pelvis. The choice of technique was at the discretion of the treating physiotherapist, according to the subject's physical examination findings. [Physiotherapist]	VAS-Pain – Average pain intensity over the last 24 h GPE PSFS RMDQ

Table 1 (continued)

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Flynn T. (2002—USA) [26]	Prospective Cohort	8 days max 3 sessions	75 (5.3%)	18—60	1/Manipulation technique with the patient supine (2 sessions) 2/Supine pelvic tilt range of motion exercise (10 repetitions, 3–4 times daily) 3/Instruction to maintain usual activity level within the limits of pain. [Physical therapist]	ODI > 50% improvement
Fritz J. (2004—USA) [27]	Prospective Cohort	8 days max 3 sessions	75 (5.3%)	18—60	1/Manipulation technique with the patient supine (2 sessions) 2/Supine pelvic tilt range of motion exercise (10 repetitions, 3–4 times daily) 3/Instruction to maintain usual activity level within the limits of pain. [Physical therapist]	ODI ≤ 5 points or less of improvement on the ODI by the time of the 3rd treatment
Fritz J. (2011—USA) [41]	Prospective cohort	6–8 days	51 (2.0%)	19–60	SMT during sessions 1 and 2. The technique provides a posterior-inferior thrust at the patient's pelvis. Thrusts were applied to each side of the pelvis during each session. [Physical therapist or chiropractor]	ODI improvement from baseline to session 3 ((ODIinitial—ODIfinal)/ODIinitial * 100%)
Goldstein (2002—USA) [36]	Secondary analysis—RCT	2 weeks 6 weeks 6 months	169 (0.0%)	> 18	Spinal manipulation or other spinal adjusting techniques, instruction in proper back care, and strengthening and flexibility exercises—chiropractic care only (DC group) [Chiropractor]	NPRS RMDQ Disability score Most severe pain/average pain
Hadizadeh M. (2020—Canada) [43]	Secondary analysis—RCT [61]	1 week	241 (1.2%)	18–60	High-Velocity, Low-Amplitude (HVLA) thrust to the anterior superior iliac spine in a posterior/inferior direction. [?]	ODI ≥ 30% improvement
Hallegraeff J.M. (2020—Netherlands) [44]	Prospective Cohort	12 weeks	225 (9.3%)	18–60	Standard care physiotherapy according to the Dutch clinical practice guideline for low back pain was carried out (Physiotherapists working in primary care and who were studying for a Master of Science degree in Manual Therapy) [Physiotherapist]	NPRS—greater than 'very mild' pain that resulted in a pain score ≥ 3/10 on the Numerical Pain Rating Scale PDI—the outcome of pain-related disability at 12 weeks, a score ≥ 19/70 on the PDI

Table 1 (continued)

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Kizhakkeveettil A. (2019—USA) [46]	Secondary analysis—RCT [62]	60 days	28 (0.0%)	≥ 18	Doctors of chiropractic provided SMT using specific contact points on vertebral processes to improve the biomechanics of their associated joints. The segments to be treated were determined based on static and motion palpation and reports of tenderness to palpation. Adjunctive therapies included passive articulatory mobilization of the lumbosacral spinal joints, paraspinal soft tissue stretching, digital pressure on tender points, and postisometric muscular relaxation procedures. If necessary, physical therapeutic modalities such as heat, cold, ultrasound, electrical muscle stimulation, and active care exercises were also used. The typical visit lasted 15 to 30 min. [Chiropractor]	NPRS: • changes in pain (2 + points) from baseline to 60 days—current LBP • changes in pain (2 + points) from baseline to 60 days—Typical LBP last week • changes in pain (2 + points) from baseline to 60 days—lowest LBP last week • changes in pain (2 + points) from baseline to 60 days—Highest LBP last week RMDQ—changes in disability (3 + points) from baseline to 60 days

Table 1 (continued)

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Licciardone J. (2013—USA) [42]	Secondary analysis—RCT [63]	12 weeks	230 (21.7%)	21–69	Osteopathic manual treatment included high-velocity, low-amplitude thrusts; moderate-velocity, moderate amplitude thrusts; soft tissue stretching, kneading, and pressure; myofascial stretching and release; positional treatment of myofascial tender points; and muscle energy techniques. These techniques were aimed primarily at the lumbosacral, iliac, and pubic regions. [Osteopathic physicians]	VAS-Pain ≥ 50% pain reduction to determine substantial improvement at week 12 RMDQ score reduction ≥ 5 points represents a minimally important change in a patient SF-36 GH (General Health) An increase of ≥ 6 points represents a minimally important change on this scale Satisfaction with back care five-point Likert scale was dichotomized by combining"very satisfied"and"satisfied"responses vs. all others Use of co-treatment Exercise programming Non-Prescription medication Prescription medication Physical Therapy CAM therapies Work Disability (the number of lost work days because of LBP)
Licciardone J. (2014—USA) [30]	Secondary analysis—RCT [63]	12 weeks	230 (21.7%)	21–69	Osteopathic manual treatment targeted the lumbosacral, iliac, and pubic regions and consisted primarily of HVLA thrusts. Moderate-velocity, moderate amplitude thrusts; soft-tissue stretching, kneading, and pressure; myofascial stretching and release; positional treatment of myofascial tender points (counterstain); and muscle energy techniques. [Osteopathic physicians]	VAS-Pain ≥ 30% pain reduction from baseline to week 12
McMorland. (2000—Canada) [31]	Retrospective outcome-based analysis	1 year	58 (NA)	/	spinal manipulation (diversified technique) and various soft-tissue techniques [Chiropractor]	VAS-Pain mild pain, 0 to 3 points; moderate pain, 4 to 7 points; severe pain, 8 to 11 points ODI minimal disability, 0 to 20 points; moderate disability, 21 to 40 points; severe disability, > 40 points

Table 1 (continued)

Author (year-Country)	Study design	Follow-up	Baseline Sample size (% dropout)	Age population	Intervention [Therapist]	Outcome measures
Nim C. (2021—Denmark) [45]	Secondary analysis—RCT [64]	4 weeks	132 (6.8%)	18–60	The spinal manipulation was provided in a standardized manner with the participant in a side-lying position. A high velocity, low amplitude thrust targeted the randomized segment (the stiffest or the most pain-sensitive) in a posterior to anterior direction. [Chiropractor]	NPRS – Linear change score ODI – Linear change score
Sheets C. (2012—USA) [29]	Secondary analysis—RCT [65, 66]	3 weeks 6 sessions max	148 (6.1%)	20—80	McKenzie method, in addition to the first-line care (advice to remain active and avoid bed rest, reassurance of the favorable prognosis of acute LBP, And instruction to take paracetamol on a time-contingent basis) [Physical therapist]	VAS-Pain 1 point difference in pain score
Underwood M.R. (2007—UK) [33]	Secondary analysis—RCT [67, 68]	3 month 12 months	1334 (25.4%)	18–64	"active management"of back pain, and they provided patients with copies of"the back book"+ Spinal manipulation package (8 sessions over up to 12 weeks) [Physical therapist or chiropractor or osteopath]	RMDQ
Wong A. (2015—USA) [40]	Non-Randomized controlled study	1 week	32 (0.0%)	18–60	Standardized application of SMT. This technique applies a posteroinferior thrust to the patient's pelvis, and a maximum of 2 thrusts are delivered to each side of the subject during each session. [?]	mODI – SMT responders (a cutoff of ≥ 30% reduction in baseline mODI scores) OR nonresponders (< 30% reduction in baseline mODI scores)
Xia T. (2017 – USA [28])	Prospective Cohort	6 weeks	80 (15.0%)	21–65	High-velocity, low-amplitude thrust spinal manipulation is a diversified chiropractic technique to treat LBP (or modified side-lying thrust SM). The clinician determined the number of manipulations and specific target joints at each visit. [Chiropractor]	RMDQ

Table 2 Risk of bias assessment of included studies

Author (Year—Country)	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting
Annen M. (2018—Switzerland) [32]	M	L	L	L	H	L
Burns S. (2018—USA) [38]	M	M	M	M	H	L
Cecchi F. (2012—Italy) [34]	L	M	L	L	H	L
Cook C. (2013—USA) [39]	M	M	M	M	H	L
Donaldson M. (2013—USA) [37]	M	M	M	M	H	M
Ferreira M. (2009—Australia) [35]	M	M	M	L	H	L
Flynn T. (2002—USA) [26]	M	M	M	M	H	L
Fritz J. (2004—USA) [27]	M	L	M	M	H	L
Fritz J. (2011—USA) [41]	M	M	L	L	M	M
Goldstein. (2002—USA) [36]	M	H	M	L	M	L
Hadizadeh M. (2020—Canada) [43]	L	L	L	L	M	L
Hallegraeff JM. (2020—Netherlands) [44]	L	M	L	L	M	L
Kizhakkeveettil A. (2019—USA) [46]	M	L	L	L	H	L
Licciardone J. (2013—USA) [42	L	M	M	L	H	L
Licciardone J. (2014—USA) [30]	L	M	H	L	M	L
McMorland. (2000—Canada) [31]	H	H	M	L	H	H
Nim C. (2021—Denmark) [45]	M	L	M	L	H	L
Sheets C. (2012—USA) [29]	L	M	L	L	H	L
Underwood MR. (2007—UK) [33]	M	M	L	L	H	L
Wong A. (2015—USA) [40]	M	L	M	L	H	L
Xia T. (2017—USA) [28]	M	M	M	L	L	L

L Low risk of bias, *M* Medium risk of bias, *H* High risk of bias

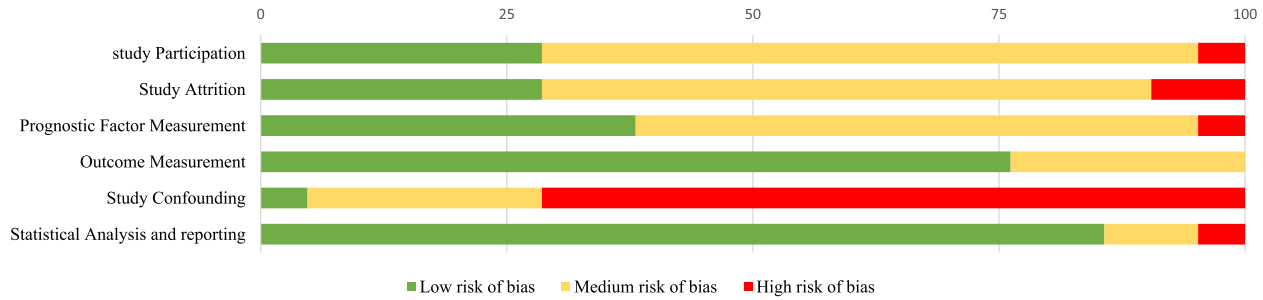

Fig. 2 Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies

Association with an improvement in pain intensity

Twelve studies reported associations between BIFs and measures of pain intensity [29–32, 35–37, 39, 42, 44–46]. Among these, six studies reported on the *patient characteristics* category [29–31, 37, 39, 46], four on the *self-reported questionnaires* category[29, 30, 36, 39], four on the *clinical examination* category [30, 35, 39, 47], and three on the *ancillary tests* category [30, 32, 45].

In multivariate analysis, participants who had completed college education were 3.26 times more likely to respond favorably to manual therapy at 12 weeks than participants with lower educational attainment (adjusted Odds Ratio (aOR): 3.26, 95%CI: 1.72–6.16). Two models using the same data reported a significant association with pain score change. The first model showed an association between a positive clinical prediction rule (CPR) where participants meet at least 4 out of 5 criteria from a clinical prediction rule reported by Flynn rule [26], shorter duration of symptoms, lower NPRS score at the first visit and at least a 2.5-point improvement in NPRS (Nagelkerke, $R^2 = 20.2$; $p < 0.01$ [39]). The second model showed an association between NPRS score at the first visit, ODI score at the first visit, positive CPR, shorter duration of symptoms, and NPRS change score (F value/

χ^2 Value: 57.4; adjusted R^2/Pseudo R^2: 61.9; $p < 0.01$) [39] (Table 3).

Inconsistent results were found for ODI baseline scores, pain ratings, and duration of symptoms.

Non-significant associations ($p > 0.05$) between BIFs and an improvement in pain are listed in Additional Table S2.

Association with non-improvement in pain intensity

Only one study reported results on non-improvement in pain [44]. This study explored potential associations between 'state' anxiety and 'trait' anxiety measured using the State-Trait Anxiety Inventory (STAI-S—STAI-T) at baseline (*self-reported questionnaires* category) and pain measured using the NPRS at 12 weeks.

The adjusted multivariate regression of 'state' and 'trait' anxiety showed an increase of 10% in the probability of having pain scores ≥ 3/10 on the NPRS at 12 weeks after manual therapy, with every unit increase in the score on the STAI-S (β = 0.08, aOR:1.1 (1.1—1.1), $p = 0.00$) and the STAI-T (β = 0.05, aOR:1.1 (1.0—1.1), p = 0.01), respectively (Table 3).

Non-significant associations ($p > 0.05$) between BIFs and a lack of improvement in pain are listed in Additional Table S2.

Association with an improvement in disability

Fifteen studies reported associations with disability improvement outcomes [26, 28, 31–33, 35–37, 39–43, 45, 46], including eight in the *patient characteristics* category [26, 31, 33, 37, 39, 40, 43, 46], seven in the *self-reported questionnaires* category [26, 33, 36, 39, 40, 42, 43], four in the *clinical examination* category [26, 35, 39, 43] and six in the *ancillary tests* category [28, 32, 40, 41, 43, 45].

In multivariate analysis, among the BIFs studied, spinal stiffness based on the palpatory method was associated with at least 30% improvement from baseline on the RMDQ at six weeks, after adjusting for age, sex, educational level, baseline measures of work status, co-morbid osteoarthritis, low back pain duration, use of prescription and non-prescription medication for low back pain, co-treatment with either active or sham ultrasound therapy [28]. Spinal stiffness based on the palpatory method was significantly related to responder status (linear mixed-effects regression models F1,63 = 5.38, p = 0.02, adjusting for sex, age, and body mass index (BMI). The adjusted mean Newton/millimeter (Standard Error) baseline stiffness of responders (adj. mean: 4.6; SE: 0.2) was less than that of non-responders (5.3 (0.3)) [28].

Five other models showed an association between BIFs and disability change scores. The first model, including duration of symptoms (< 16 days), FABQ work subscale (< 19 out of 42), lumbar hypomobility, hip internal rotation (> 35°), symptoms not distal to the knee was associated with a minimum improvement of 50% of the ODI score model ($\chi^2 = 48.5$, df = 5, P < 0.001, Nagelkerke $R^2 = 0.67$) [26]. The second model included Met CPR, younger age, and strains and sprains diagnosis was associated with a minimum improvement of 50% of the ODI score (Nagelkerke $R^2 = 23.1$, $P < 0.01$) [39]. The third model, including ODI score on the first visit, met CPR, symptoms duration, age, strains, and sprains diagnosis, was associated with linear ODI change score (Model F value/χ^2 Value: 20.3, Model adjusted R^2/Pseudo R^2: 41.1, $p < 0.01$) [39]. The fourth model, including Age, Sex (male), BMI, Met CPR, initial terminal stiffness (and Immediate change in global stiffness), was associated with improvement of ODI (Adjusted R^2: 0.21; Standardized β: −0.32; $p = 0.047$) [41]. Finally, the fifth model, including Height, gender, neck or upper back pain, pain frequency in the past six months, patient's expectation of medication, patient's expectation of strengthening exercises, the score of STarT Back Screening Tools, extension status, was associated with an improvement of 30% of the ODI [43] (Table 3).

Inconsistent results were found in univariate association for participants'age, symptoms duration, expectation/preference which intervention and gender in the *patient characteristics* category and for the duration of symptoms, work subscale of FABQ, the ODI baseline score, the pain intensity rating, and the pain location diagram (or pain distribution) of the *self-reported questionnaires* category. Inconsistent results were also found for the total extension range of motion in the *clinical examination* category (Additional Table S2).

Non-significant associations (p > 0.05) between BIFs and an improvement in disability are listed in Additional Table S2.

Association with a non-improvement of disability

Three studies reported results on disability non-improvement [27, 34, 44]. All studied BIFs in the *self-reported questionnaires* category [27, 34, 44], two in *patient characteristics* category [27, 34], one studied BIFs in the *clinical examination* category [27], and none studied BIFs in the *ancillary tests* category.

Univariate analysis revealed a statistically significant difference between responders and non-responders regarding the previous low back pain treatment [34], with non-responders showing a lower frequency of prior treatment. When this variable was tested in a multivariate analysis with baseline disability score and age, the association was not statistically significant.

In multivariate analysis, three models showed a significant association between BIFs and lack of improvement. The first model revealed a significant association between

Table 3 Multivariate association with improvement and non-improvement

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias	
Burns S. (2018—USA) [38]	Improvement (needed to score ≤ 10% on the mODI and ≤ 2 on the NPRS at discharge and record a GROC score ≥ + 4 at both 2 weeks and discharge)	90—Discharge	the stepwise logistic regression equation	BMI (continuous)	0.84 (0.75; 0.94)	< 0.01	A lower BMI score was associated with improved odds of recovery	$R^2 = 0.384$	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M M M H L
				Concurrent hip problem (binary)	5.34 (1.31; 21.8)	.02	Having a concurrent hip problem made an individual 5.34 times more likely to be defined as recovered			
				Irritability status (binary)	3.63 (1.16; 11.4)	.03	A clinical presentation that was deemed to be irritable by the treating physical therapist at baseline examination was also significantly associated with increased odds of recovery			
				Baseline Pain rating < 4 (dichotomized—binary)	4.99 (1.41; 17.7)	.01	A baseline NPRS score of 4 points or less was significantly associated with improved odds of achieving recovery			
				Groupe allocation	0.5 (0.18; 1.44)	.20				
Cook C. (2013—USA) [39]	Improvement—ODI (≥ 50% improvement)	149—Discharge	Backward stepwise logistic regression modeling	Met CPR (binary)	2.9 (1.4; 6.2)	<.01	Individuals who met the CPR for manipulation were 2.9 times more likely to respond favorably compared with those who did not meet the CPR	Nagelkerke R^2 = 23.1 p < 0.01	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M M M H L
				Age (binary)	1.04 (1.01; 1.06)	<.01	Younger age was significantly associated with a 50% reduction in ODI scores			
				Strains and sprains – diagnosis (binary)	2.6 (1.2; 5.5)	.01	Individuals with a diagnosis of strains and sprains in the lower back were 2.6 times more likely to respond favorably compared with those without this diagnosis			

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias
	Improvement -NPRS ≥ 2.5 points change	147—Discharge		Met CPR (binary)	4.8 (1.8; 10.4)	<.01	Individuals who met the CPR for manipulation may have been 4.8 times more likely to respond favorably compared with those who did not meet the CPR	Nagelkerke R^2 = 20.2 $p < 0.01$	
				Duration of symptoms	2.4 (1.4; 4.1)	<.01	Individuals with short duration of symptoms maybe 2.4 times more likely to respond favorably compared with those with longer duration of symptoms		
				Lower NPRS score on the first visit	1.2 (0.99; 1.6)	.06			
	Improvement -Total visits ≤ 6	141—Discharge		Met CPR (binary)	3.7 (1.7; 7.6)	<.01	Individuals who met the CPR for manipulation were 3.7 times more likely to respond favorably compared with those who did not meet the CPR	Nagelkerke R^2 = 15.4 $p < 0.01$	
				Strains and sprains – diagnosis (binary)	1.2 (1.0; 4.4)	.05	Individuals diagnosed with strains or sprains in the lower back were 1.2 times more likely to respond favorably compared with those who did not meet the CPR		
				Irritability	0.51 (0.23; 1.1)	.10			
	Improvement -Extent of recovery ≥ 75%	142—Discharge		Met CPR (binary)	4.0 (1.6; 9.8)	<.01	Individuals who met the CPR for manipulation were 4.0 times more likely to respond favorably compared with those who did not meet the CPR	Nagelkerke R^2 = 17.5 $p < 0.01$	
				Irritability	0.35 (0.15; 0.82)	.02	No irritability at baseline was associated with a positive outcome		

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias
	Improvement -ODI change score	149—Discharge		ODI score on the first visit	0.48 (0.38; 0.59)[§]	<.01	The ODI score at baseline was associated with ODI change scores	Model F value/χ^2 Value: 20.3 Model adjusted R^2/ Pseudo R^2: 41.1 $p < 0.01$	
				Met CPR (binary)	−4.2 (−7.7; −0.69)[§]	.02	Meeting the clinical prediction rule for manipulation at baseline is associated with a decrease in the ODI score		
				Duration of symptoms	0.35 (−0.06; −0.01)[§]	.01	A shorter duration of symptoms was associated with ODI change scores		
				Age	−0.16 (−0.27; −0.04)[§]	.01	Age was negatively associated with ODI change score		
				Strains and sprains – diagnosis (binary)	−3.39 (−6.7; −0.08)[§]	.04	Diagnosing strains or sprains for manipulation at baseline is associated with a decrease in the ODI score		
	Improvement -NPRS change score Improvement -Total visits	147—Discharge		NPRS score on the first visit	0.92 (0.79; 1.05)[§]	<.01	NPRS score on the first visit was associated with the NPRS change score	F value/χ^2 Value: 57.4; adjusted R^2/ Pseudo R^2: 61.9; $p < 0.01$	
				ODI score on the first visit	−0.02 (−0.04; −0.001)[§]	.04	The ODI score at baseline was negatively associated with NPRS change scores		
				Met CPR (binary)	−0.98 (−1.5; −0.47)[§]	<.01	Meeting the clinical prediction rule for manipulation at baseline is associated with decreased pain		
				Duration of symptoms	−0.004 (−0.009; 0.0)[§]	.05			
		141—Discharge		Met CPR (binary)	0.32 (0.19; 0.45)[§]	<.01	Meeting the clinical prediction rule for manipulation at baseline is associated with total visits	Model F value/χ^2 Value: 39.5 Model adjusted R^2/Pseudo R^2: 4.4 $p < 0.01$	
	Improvement -Extent of recovery (0%—100%)			Strains and sprains – diagnosis (binary)	0.19 (0.07; 0.32)[§]	<.01	Diagnosing strains and sprains in the lower back was associated with a decrease in the total number of visits		
				Irritability	−0.19 (−0.32; −0.05)[§]	<.01	No irritability at baseline was associated with a positive outcome		

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias	
		142—Discharge		Met CPR (binary)	−10.8 (−18.3; −3.1)[§]	<.01	Meeting the clinical prediction rule for manipulation at baseline is associated with improvement	Model F value/χ^2 Value: 6.2 Model adjusted R^2/Pseudo R^2: 7.1 $p < 0.01$		
				Irritability	7.1 (−1.1; 15.4)[§]	.09				
Flynn T. (2002—USA) [26]	Improvement - ODI (> 50% improvementI)	75—8 days Max—3 sessions	Stepwise logistic regression	Duration of symptoms (dichotomized – binary: < 16 days)	4.39 (1.83; 10.51)[£]		Having a symptom duration of less than 16 days made an individual 4.39 times more likely to be improved	model $\chi^2 = 48.5$, df = 5, Nagelkerke $R^2 = 0.67$ $p < 0.001$	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M M M H L
				FABQ work subscale (dichotomized – binary: < 19)	1.65 (1.17; 2.31)[£]		Having a FABQ work subscale of less than 19 made an individual 1.65 times more likely to be improved			
				Lumbar hypomobility (binary: One or more lumbar levels with manual spring testing)	1.26 (1.05; 1.51)[£]		Having a lumbar hypomobility made an individual 1.26 times more likely to be improved			
				Hip internal rotation (binary: At least one hip internal rotation range of motion > 35°)	3.25 (1.44; 7.33)[£]		Having a hip internal rotation greater than 35° made an individual 1.65 times more likely to be improved			
				Distribution of symptoms (binary: Symptoms not distal to the knee)	1.36 (1.04; 1.79)[£]		Having symptoms not distal to the knee made an individual 1.65 times more likely to be improved			
				compression/distraction sacroiliac test (binary)	1.22 (0.94; 1.58)[£]					
				Frequency of low back pain episodes (binary)	1.33 (0.95; 1.87)[£]					
				Lumbar spring testing (binary: Pain at one or more lumbar levels with spring testing)	1.11 (0.97; 1.27)[£]					
				Distribution of symptoms (binary: Symptoms in the low back only)	1.76 (0.87; 3.58)[£]					

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias	
				Does not peripheralize with lumbar single movement testing (binary)	1.27 (0.97; 1.65)[£]					
				Standing is not ranked as worst position (binary)	1.31 (1; 1.74)[£]					
Fritz J. (2011—USA) [41]	Improvement ODI	50—6–8 days	Stepwise hierarchical linear regression (Step 1: Age; Sex (male); BMI; Met CPR | Step 2: Immediate change in global stiffness.)	Initial terminal stiffness			Less initial TS was associated with greater ODI improvement over one week	Adjusted R^2: 0.21; Standardized β: −0.32 p = 0.047	Study participation Study attrition Prognostic factor measurement Outcome measuremen Study confounding Statistical analysis and reporting	M M L L M M
Hadizadeh M. (2020—Canada) [43]	Improvement - ODI(≥ 30% improvement)	241–1 week	Logistic regression	Height	0.75 (0.65–0.86)	0.00	Shorter, more improvement	PLR: 4.6 (58.1–83.1)	Study participation Study attrition Prognostic factor measuremen Outcome measurement Study confounding Statistical analysis and reporting	L L L L M L
				Gender	0.42 (0.24–0.73)	0.00	Male, more improvement			
				Neck or upper back pain	0.53 (0.35–0.80)	0.00	No neck or upper back pain, more improvement			
				Pain frequency in the past 6 months	2.25 (1.58–3.20)	0.00	More pain frequency, more improvement			
				Patient's expectation of medication	0.49 (0.33–0.72)	0.00	Lower expectations, more improvement			
				Patient's expectation on strengthening exercises	2.47 (1.24–4.93)	0.01	Higher expectations, more improvement			
				The STarT Back Tool	0.74 (0.60–0.90)	0.00	Lower score, more improvement			
				Extension status	1.48 (1.04–2.11)	0.03	Peripheralized pain with extension, more improvement			
				Current pain duration	1.00 (1.00–1.00)	0.01	No changes			
				Depression	0.68 (0.44–1.03)	0.07	Not significant			
Licciardone J. (2014 – USA) [30]	Improvement - VAS-Pain	230–12 weeks	Logistic regression	completed college education*	3.26 (1.72–6.16)		Participants with completed college education are 3.26 times more likely to respond favorably to manual therapy at 12 weeks than participants with lower educational attainment		Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	L M H L L L

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias	
Xia T. (2017—USA) [28]	Improvement - RMDQ	6 weeks	linear mixed-effects regression models	Spinal stiffness based on the palpatory method**			The adjusted mean (SE) baseline stiffness of responders [4.6 (0.2)] was less than that of nonresponders [5.3 (0.3)]	$F_{1,63} = 5.38$, $p = 0.02$	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M M L L L
Cecchi F. (2012—Italy) [34]	Non-improvement - RMDQ (RMDQ score improved less than 2.5 on discharge compared to their baseline score.)	68—1 year (discharge)/ individual physiotherapy	Multivariable logistic regression	Disability score	0.81 (0.71–0.92)	0.00	Patients with the lowest disability score tertile (Roland Morris less than 6) were at higher risk of non-recovery	LR $\chi^2 = 15.72$; Prob > $\chi^2 < 0.001$; Pseudo $R^2 = 0.172$	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M L L M M
				Age	1 (0.97–1.04)	0.78				
				Previous treatment	1.51 (0.43–5.23)	0.52				
		69—1 year (discharge)/spinal manipulation		Age	1.01 (0.94–1.08)	0.76		LR $\chi^2 = 5.75$; Prob > $\chi^2 < 0.125$; Pseudo $R^2 = 0.108$		
				Previous treatment	0.34 (0.07–1.59)	0.17				
				Disability score	0.8 (0.62–1.04)	0.10				
Fritz et al. (2004—USA) [27]	Non-improvement - mODI - (≤ 5 points or less of improvement on the ODI by the time of the 3rd treatment)	75—8 days Max—3 sessions	Logistic Regression	Average total hip rotation ROM	0.95 (0.90–1.00)	0.04	Less total hip rotation ROM was associated with no improvement	Hosmer–Lemeshow $\chi^2 = 4.87$, $R^2 = 0.63$ $p = 0.77$,	Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	M M M M H L
				Duration of symptoms	1.03 (1.01–1.06)	0.01	Longer duration of symptoms was associated with no improvement			
				Gaenslen sign	0.11 (0.02–0.68)	0.02	A negative provocation test (Gaenslen test) was more common in subjects who did not improve with manipulation			
				Hip medial rotation ROM discrepancy	0.68 (0.51–0.9)	0.01	Reduced hip medial rotation ROM discrepancy was associated with no improvement			
				Any hypomobility in the lumbar spine with spring testing	0.092 (0.01–0.84)	0.03	The absence of hypomobility in the lumbar spine with spring testing was associated with no improvement			
				Low back pain only	0.14 (0.014–1.46)	0.10				

Table 3 (continued)

Author (Year—Country)	Outcome measures (result criteria)	Number of participants—Follow-up	Statistical tests	Studied potential associated factors in the final model (variable type)	Result ORs (95%CI)	P value	Interpretation of original author	Result of the final model	Risk of bias	
Hallegraeff JM. (2019—Netherlands) [44]	Non-improvement - NPRS ≥ 3/10	225–12 weeks	Multivariate regression	State anxiety about pain	1.1 (1.1; 1.1)	0.00	An increase of 10% in the probability of developing CLBP with every unit increase in the score on the STAI-S (state anxiety)		Study participation Study attrition Prognostic factor measurement Outcome measurement Study confounding Statistical analysis and reporting	L M L L M L
				Trait anxiety for pain	1.1 (1.0; 1.1)	0.01	an increase of 10% in the probability of developing CLBP with every unit increase in the score on the STAI-T.(trait anxiety)			
				State anxiety for pain ***	1.1 (1.0; 1.2)	0.00	An increase of 10% in the probability of developing CLBP with every unit increase in the score on the STAI-S (state anxiety)			
				Trait anxiety for pain***	1.0 (0.9; 1.1)	0.26	Not significant			

§ Unstandardized Coefficient β Value | L: Low risk of bias; M: Moderate risk of bias; H: High risk of bias | 95%CI: Confidence interval of 95% | ODI: Oswestry Disability Index | RMDQ: Roland-Morris Disability Questionnaire | VAS: Visual Analogue Scale | CPR: Clinical Prediction Rule | NPRS: Numeric Pain Rating Scale | GROC: Global Rating of Change | BMI: Body Mass Index | PDI: Pain Disability Inventory | LR: Likelihood Ratio | $^{£}$: Positive Likelihood Ratio (PLR) | P* Fully adjusted model: age, sex, and educational level; baseline measures of work status, co-morbid osteoarthritis, low back pain duration, and use of prescription and non-prescription medication for low back pain; and co-treatment with either active or sham ultrasound therapy | ** adjusted for sex, age and BMI | ***Adjusted for pain, pain-related disability, duration, physical workload and widespread pain.

a low RMDQ disability score (< 2.5) at baseline and a non-recovery at one-year follow-up (LR $\chi^2 = 15.72$; Prob > $\chi^2 < 0.001$; Pseudo $R^2 = 0.172$) [34]. The second model (including average total hip rotation, duration of symptoms, Gaenslen's sign, hip medial rotation discrepancy, and any hypomobility in the lumbar spine with manual spring testing) showed a significant association with ODI disability change score of ≤ 5 points or less at eight days (Hosmer–Lemeshow $\chi^2 = 4.87$, $p = 0.77$, $R^2 = 0.63$) [27]. The third model showed a significant association between a high anxiety state about the pain reported at baseline and no disability improvement at 12 weeks, as measured by a PDI score above 19/70, after adjusting for pain, pain-related disability, duration, physical workload, and widespread pain [44] (Table 3).

Inconsistent results were found regarding participants'age in the *patient characteristics* category and for the pain distribution of the *self-reported questionnaires* category.

Non-significant associations ($p > 0.05$) between BIFs and a lack of improvement in disability are listed in Additional Table S2.

Association with an improvement of global perceived effect

Five studies reported results on global perceived effect outcomes [35, 37–39, 42], three studied BIFs in the *patient characteristics* category [37–39], three studied BIFs in the *self-reported questionnaires* category [38, 39, 42], three studied BIFs in the *clinical examination* category [35, 38, 39], and one studied BIFs in the *ancillary tests* category [39].

In multivariate analysis, three models showed an association between BIFs and improved global perceived effect. The first model (including BMI, concurrent hip problem, irritability status, and baseline pain rating <4/10) showed a significant association with improvement in a composite score (needed to score ≤10% on the mODI and ≤2 on the NPRS at discharge and record a GRoC score ≥ + 4) ($R^2 = 0.384$) [38]. The second model (Met CPR and irritability) was associated with an extent of recovery ≥75% (Nagelkerke $R^2 = 17.5$, $p < 0.01$) [27]. The third model showed an association between met CPR and extent of recovery (Model F value/χ^2 Value: 6.2, Model adjusted R^2/Pseudo R^2: 7.1, $p < 0.01$) [39] (Table 3).

Inconsistent findings were found for LBP irritability status [48] based on the characteristics that guide classification (i.e., time and vigor of activity aggravating pain, severity of pain, and persistence of pain after aggravating activities are stopped) and BMI in the *patient characteristics* category.

Non-significant associations ($p > 0.05$) between BIFs and an improvement in global perceived effect are listed in Additional Table S2.

Association with non-improvement of global perceived effect

No included studies investigated the association between BIFs and non-improvement in the global perceived effect outcome category.

Association with an improvement with other outcomes.

Three studies reported potential associations between BIFs and other outcomes improvement [37, 39, 42]. Two of these studies assessed potential association with BIFs in the *patient characteristics* category [37, 39]; Two studies assessed potential association with BIFs in the *self-reported questionnaires* category [39, 42], and only one assessed potential association with BIFs in the *clinical examination and ancillary tests* categories [39].

In multivariate analysis, two models showed a significant association between BIFs and a decreased number of total visits. The first model (including met CPR and diagnosis of strains and sprains) was associated with a total number of visits ≤6 (Nagelkerke $R^2 = 15.4$, $p < 0.01$) [39]. The second model (including met CPR, diagnosis of strains and sprains, and irritability [48, 49]) was associated with low number of patient visits (Model F value/χ^2 Value: 39.5, Model adjusted R^2/Pseudo R^2: 4.4, $p < 0.01$) [39] (Table 3).

Inconsistent results were reported for the variable "preference for the intervention received"of the *patient characteristics* category and irritability of the *clinical examination* category.

Non-significant associations ($p > 0.05$) between BIFs and an improvement in the other category are listed in Additional Table S2.

Association with non-improvement

No included studies investigated the association between BIFs and non-improvement in the other outcomes category.

Discussion

To our knowledge, this is the first systematic review to summarize the available evidence on baseline individual factors (BIFs) associated with clinical outcomes following manual therapy interventions in managing NSLBP without radiculopathy.

Across the included studies, several BIFs showed significant associations with clinical improvement following manual therapy, such as higher educational level, shorter symptom duration, specific clinical signs (e.g., lumbar hypomobility, spinal stiffness, irritability), and positive treatment expectations. Conversely, psychological factors (notably state and trait anxiety), low baseline disability, and absence of physical impairment (e.g., normal mobility or reduced stiffness) were associated with

non-improvement. However, findings were inconsistent across studies.

Subgrouping strategies for patients with NSLBP have evolved significantly over the past decades. Earlier efforts, as described by Kent et al. (2005), frequently attempted to classify patients according to putative pathoanatomical sources [50]. To date, clinical prediction rules or stratification tools are multicomponent, and include a biopsychosocial approach [51, 52]. A recent systematic review by Vidal et al. (2025), which identified over 200 prognostic models for conservative care in LBP highlighted methodological flaws (e.g., selection of predictor, intervention characteristics) and the absence of external validation [12]. Together, these findings emphasize that although subgrouping remains conceptually appealing, its clinical implementation is not yet supported by evidence.

Our findings are consistent with another review, published in 2015, reporting that no multivariate models have been validated for clinical practice subgrouping low back pain patients [53]. Another approach to sub-grouping is based on treatment modalities. Bastos and al., in 2022, in a systematic review and meta-analysis, concluded that a stratification model using clinical signs and symptoms as treatment effect modifiers for chronic NSLBP was no better than the comparison groups without classification. Nonetheless, they reported that a subgrouping approach may be helpful for acute cases, sciatica, or stenosis. Still, the quality of evidence at this time was low, with effect sizes ranging from small to moderate [54]. Further, a systematic review by Saragiotto et al. in 2016 concluded that claims of a subgroup effect are often overstated in primary studies [55].

Given our results and the complex nature of NSLBP with multifactorial contributors, it is currently impossible to identify patients who will respond (or not) to manual therapy. According to Hayden et al.'s framework [13], the evidence of factors associated with clinical improvement after manual therapy remains in the exploration phase. The current review did not identify any article differentiating prognostic factors and treatment effect modifiers.

Furthermore, as emphasized by Maissan et al. (2018), the assessment of prognostic factors requires a clearly articulated clinical reasoning process in which the intervention is appropriately matched to a clinically relevant impairment [56]. When no such impairment is present at baseline, for example, when the range of motion is within normal limits, the likelihood of observing a meaningful treatment effect is inherently limited, which may obscure the identification of prognostic factors.

Strengths and limitations

There are several strengths to this review. First, the search strategy was carried out by an independent librarian in four electronic databases reviewed by a second librarian using the Peer Review of Electronic Search Strategies (PRESS) Checklist. Second, in contrast to other review on this topic [11], we only selected studies that included NSLBP without radiculopathy. Third, we used a specific and validated risk-of-bias scoring tool for prognostic studies (QUIPS) [23].

The main limitation of this review is that it is not possible to conclude if the BIFs may be considered as predictors of clinical course or treatment effect modifiers. Indeed, only randomized clinical trials allow for the conclusion of treatment effect modifiers, while prospective/retrospective cohort studies could be used to determine prognosis. Further, only articles in English or French were included, possibly leaving out some relevant studies. Over 75% of the included studies were classified as high risk of bias for confounding variables. Indeed, most of the included studies conducted univariate analyses corresponding to an exploration phase. The results of these univariate analyses should be interpreted cautiously as any significant association may be due to another non-explored confounding variable.

We selected interventions with a focus on manual therapy alone or in combination with other therapies such as advice, ultrasound therapy, and exercises. The inclusion of studies with limited information on intervention content necessitated discussion and consensus among the authors. An association between a factor and clinical outcome may be linked to another complementary therapy. It should also be noted that our review focuses only on BIFs; thus, not all factors related to the therapeutic relationship have been explored. An expert consensus on manual therapies suggests that the therapeutic relationship may play an important role in the clinical results of spinal manipulation treatments [57]. However, our review did not find significant associations between "patient confidence in the practitioner", "treatment confidence" and pain or disability outcomes in two included studies.

Implications of the results for practice

Manual therapy providers should be aware that current evidence does not support the classification of patients for choosing therapeutic tools based on BIFs to improve clinical outcomes in NSLBP.

Implications of the results for future research

This review highlights the need for well-designed confirmatory studies based on factors associated in the exploratory phase, using a design that can distinguish prognostic factors from a treatment effect modifier. To determine if these associations are valid, the factor should be included in confirmatory phase studies

according to the Hayden et al. framework in multivariate models controlling for confounders.

Qualitative studies exploring the therapeutic relationship and patients'expectations would also be beneficial in identifying factors associated with clinical outcomes following manual therapy to improve tailored care.

Conclusions

The use of patient BIFs in clinical practice to predict clinical outcomes following manual therapy treatment appears to be premature at this time. Studies included in this review were exploratory in nature, and most had a high risk of confounding bias. Future studies should aim to differentiate prognostic factors and treatment effect modifiers.

Abbreviations

BIFs	Baseline Individual Factors
BMI	Body Mass Index
BQ	Bournemouth Questionnaire
CPR	Clinical Prediction Rule
FABQ	Fear Avoidance Belief Questionnaire
GPE	Global Perceived Effect
GroC	Global Rating of Change
NSLBP	Non-Specific Low Back Pain
NPRS	Numerical Pain Rating Scale
ODI	Oswestry Disability Index
mODI	modified Oswestry Disability Index
PDI	Pain Disability Inventory
PSFS	Patient-Specific Functional Scale
PRISMA	Preferred Reporting Items for Systematic review and Meta-Analyses
PROMs	Patient-Reported Outcomes Measures
QUIPS	Quality In Prognosis Studies
RCT	Randomized controlled trials
RMDQ	Roland Morris Disability Questionnaire
STAI-S	State Anxiety Inventory
STAI-T	Trait Anxiety Inventory
VAS	Visual Analogue Scale

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12906-025-04975-y.

Supplementary Material 1. Table S1 – Search strategy.

Supplementary Material 2. Table S2 - Synthesis of the associations.

Supplementary Material 3. Table S3 – PRISMA Checklist.

Acknowledgements

We thank Florian Barbier-Cazorla for his logistical support. We gratefully acknowledge the assistance of M. Colin Sidre at the Bibliothèque Interuniversitaire de Santé (Université Paris Cité) and M. Stéphane Benais at the Bibliothèque Universitaire d'Orsay and the librarian team from Ontario Tech University.

Authors' contributions

GB, MP, AL, AB, and FC contributed to the conception or design of the work. GB, MP, AL, AD, and GG contributed to selecting articles, analyzing, or interpreting data for the work. GB, AB, and AL drafted the manuscript. MP, AD, GG, and FC critically revised the manuscript. All gave final approval and agreed to be accountable for all aspects of the work, ensuring integrity and accuracy. GB is the guarantor for the paper.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Gaetan Barbier, Mathieu Picchiottino, Guillaume Goncalves, and Arnaud Lardon are researchers-chiropractors who practice manual therapy techniques.

Author details

[1]CIAMS, Université Paris-Saclay, 91405 Orsay Cedex, France. [2]CIAMS, Université d'Orléans, 45067 Orléans, France. [3]Institut Franco-Européen de Chiropraxie, 94200 Ivry-Sur-Seine, France. [4]Société Franco-Européenne de Chiropraxie, Lille, France. [5]Université du Québec À Trois-Rivières, Trois-Rivières, Québec, Canada. [6]Laboratoire D'Anatomie Fonctionnelle, Faculté des Sciences de la Motricité, Université Libre de Bruxelles, route de Lennik 808, CP 619 - 1070 Bruxelles, Belgique. [7]Faculty of Allied Health & Wellbeing, University of Central Lancashire (UCLan), Preston, UK. [8]Laboratoire d'Anatomie, de Biomécanique et d'Organogenèse, Faculté de Médecine, Université Libre de Bruxelles, route de Lennik808, CP 619 - 1070 Bruxelles, Belgique. [9]Université Paris-Saclay, Inria, CIAMS, Gif-Sur-Yvette 91190, France.

Received: 14 November 2024 Accepted: 12 June 2025
Published online: 18 September 2025

References

1. Diseases GBD, Injuries C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204–22. https://doi.org/10.1016/S0140-6736(20)30925-9.
2. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. Lancet. 2018;391(10137):2356–67. https://doi.org/10.1016/S0140-6736(18)30480-X.
3. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet. 2017;389(10070):736–47. https://doi.org/10.1016/S0140-6736(16)30970-9.
4. Buchbinder R, Underwood M, Hartvigsen J, et al. The Lancet Series call to action to reduce low value care for low back pain: an update. Pain. 2020;161(Suppl 1):S57–64. https://doi.org/10.1097/j.pain.0000000000001869.
5. Beyera GK, O'Brien J, Campbell S. Health-care utilisation for low back pain: a systematic review and meta-analysis of population-based observational studies. Rheumatol Int. 2019;39(10):1663–79. https://doi.org/10.1007/s00296-019-04430-5.
6. Kamper SJ, Logan G, Copsey B, et al. What is usual care for low back pain? A systematic review of health care provided to patients with low back pain in family practice and emergency departments. Pain. 2020;161(4):694–702. https://doi.org/10.1097/j.pain.0000000000001751.
7. 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. https://doi.org/10.1016/S0140-6736(18)30489-6.
8. Henschke N, Maher CG, Refshauge KM, et al. Low back pain research priorities: a survey of primary care practitioners. BMC Fam Pract. 2007;8:40. https://doi.org/10.1186/1471-2296-8-40.
9. Dionne CE, Rossignol M, Deyo RA, et al. Back to the Future: A Report From the 16th International Forum for Back and Neck Pain Research in Primary Care and Updated Research Agenda. Spine (Phila Pa 1976) 2022;47(19):E595-E605. https://doi.org/10.1097/BRS.0000000000004408

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. https://doi.org/10.1016/j.physio.2015.01.006.
11. de Zoete A, de Boer MR, Rubinstein SM, et al. 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. Spine (Phila Pa 1976) 2021;46(8):E505-E17. https://doi.org/10.1097/BRS.0000000000003814
12. Vidal R, Grotle M, Johnsen MB, et al. Prediction models for outcomes in people with low back pain receiving conservative treatment: a systematic review. J Clin Epidemiol. 2024;177: 111593. https://doi.org/10.1016/j.jclinepi.2024.111593.
13. Hayden JA, Cote P, Steenstra IA, et al. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. J Clin Epidemiol. 2008;61(6):552–60. https://doi.org/10.1016/j.jclinepi.2007.08.005.
14. Kent P, Cancelliere C, Boyle E, et al. A conceptual framework for prognostic research. BMC Med Res Methodol. 2020;20(1):172. https://doi.org/10.1186/s12874-020-01050-7.
15. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ. 2013;346: e5595. https://doi.org/10.1136/bmj.e5595.
16. Kamper SJ, Maher CG, Hancock MJ, et al. Treatment-based subgroups of low back pain: a guide to appraisal of research studies and a summary of current evidence. Best Pract Res Clin Rheumatol. 2010;24(2):181–91. https://doi.org/10.1016/j.berh.2009.11.003.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev. 2021;10(1):89. https://doi.org/10.1186/s13643-021-01626-4.
18. Morgan RL, Whaley P, Thayer KA, et al. Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ Int. 2018;121(Pt 1):1027–31. https://doi.org/10.1016/j.envint.2018.07.015.
19. National Guideline C. NICE Evidence Reviews Collection. Evidence review for manual therapy for chronic primary pain: Chronic pain (primary and secondary) in over 16s: assessment of all chronic pain and management of chronic primary pain: Evidence review I. London: National Institute for Health and Care Excellence (NICE). Copyright © NICE 2021. 2021.
20. Lederman E. The Science & Practice of Manual Therapy. 2 ed2005.
21. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016;75:40–6. https://doi.org/10.1016/j.jclinepi.2016.01.021.
22. Sampson M, McGowan J, Cogo E, et al. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol. 2009;62(9):944–52. https://doi.org/10.1016/j.jclinepi.2008.10.012.
23. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280–6. https://doi.org/10.7326/0003-4819-158-4-201302190-00009.
24. Grooten WJA, Tseli E, Ang BO, et al. Elaborating on the assessment of the risk of bias in prognostic studies in pain rehabilitation using QUIPS-aspects of interrater agreement. Diagn Progn Res. 2019;3:5. https://doi.org/10.1186/s41512-019-0050-0.
25. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368: l6890. https://doi.org/10.1136/bmj.l6890.
26. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. Spine (Phila Pa 1976) 2002;27(24):2835–43. https://doi.org/10.1097/00007632-200212150-00021.
27. Fritz JM, Whitman JM, Flynn TW, et al. Factors related to the inability of individuals with low back pain to improve with a spinal manipulation. Phys Ther 2004;84(2):173–90.
28. Xia T, Long CR, Vining RD, et al. Association of lumbar spine stiffness and flexion-relaxation phenomenon with patient-reported outcomes in adults with chronic low back pain - a single-arm clinical trial investigating the effects of thrust spinal manipulation. BMC Complement Altern Med. 2017;17(1):303. https://doi.org/10.1186/s12906-017-1821-1.
29. Sheets C, Machado LA, Hancock M, et al. Can we predict response to the McKenzie method in patients with acute low back pain? A secondary analysis of a randomized controlled trial. Eur Spine J. 2012;21(7):1250–6. https://doi.org/10.1007/s00586-011-2082-1.
30. Licciardone JC, Kearns CM, Crow WT. Changes in biomechanical dysfunction and low back pain reduction with osteopathic manual treatment: results from the OSTEOPATHIC Trial. Man Ther. 2014;19(4):324–30. https://doi.org/10.1016/j.math.2014.03.004.
31. McMorland G, Suter E. Chiropractic management of mechanical neck and low-back pain: a retrospective, outcome-based analysis. J Manipulative Physiol Ther 2000;23(5):307–11.
32. Annen M, Peterson C, Humphreys BK. Comparison of Treatment Outcomes in Nonspecific Low-Back Pain Patients With and Without Modic Changes Who Receive Chiropractic Treatment. J Manipulative Physiol Ther. 2018;41(7):561–70. https://doi.org/10.1016/j.jmpt.2018.01.008.
33. Underwood MR, Morton V, Farrin A, et al. Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. Rheumatology (Oxford). 2007;46(8):1297–302. https://doi.org/10.1093/rheumatology/kem113.
34. Cecchi F, Negrini S, Pasquini G, et al. Predictors of functional outcome in patients with chronic low back pain undergoing back school, individual physiotherapy or spinal manipulation. Eur J Phys Rehabil Med 2012;48(3):371–8.
35. Ferreira ML, Ferreira PH, Latimer J, et al. Relationship between spinal stiffness and outcome in patients with chronic low back pain. Man Ther. 2009;14(1):61–7. https://doi.org/10.1016/j.math.2007.09.013.
36. Goldstein MS, Morgenstern H, Hurwitz EL, et al. The impact of treatment confidence on pain and related disability among patients with low-back pain: results from the University of California, Los Angeles, low-back pain study. Spine J 2002;2(6):391–9; discussion 99–401. https://doi.org/10.1016/s1529-9430(02)00414-x.
37. Donaldson M, Learman K, O'Halloran B, et al. The role of patients' expectation of appropriate initial manual therapy treatment in outcomes for patients with low back pain. J Manipulative Physiol Ther. 2013;36(5):276–83. https://doi.org/10.1016/j.jmpt.2013.05.016.
38. Burns SA, Cleland JA, Cook CE, et al. Variables Describing Individuals With Improved Pain and Function With a Primary Complaint of Low Back Pain: A Secondary Analysis. J Manipulative Physiol Ther. 2018;41(6):467–74. https://doi.org/10.1016/j.jmpt.2017.11.006.
39. Cook CE, Learman KE, O'Halloran BJ, et al. Which prognostic factors for low back pain are generic predictors of outcome across a range of recovery domains? Phys Ther. 2013;93(1):32–40. https://doi.org/10.2522/ptj.20120216.
40. Wong AY, Parent EC, Dhillon SS, et al. Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? Spine (Phila Pa 1976) 2015;40(17):1329–37. https://doi.org/10.1097/BRS.0000000000000981.
41. Fritz JM, Koppenhaver SL, Kawchuk GN, et al. Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. Spine (Phila Pa 1976) 2011;36(21):1772–81. https://doi.org/10.1097/BRS.0b013e318216337d.
42. Licciardone JC, Kearns CM, Minotti DE. Outcomes of osteopathic manual treatment for chronic low back pain according to baseline pain severity: results from the OSTEOPATHIC Trial. Man Ther. 2013;18(6):533–40. https://doi.org/10.1016/j.math.2013.05.006.
43. Hadizadeh M, Kawchuk GN, Prasad N, et al. Predicting who responds to spinal manipulative therapy using a short-time frame methodology: Results from a 238-participant study. PLoS ONE. 2020;15(11): e0242831. https://doi.org/10.1371/journal.pone.0242831.
44. Hallegraeff JM, Kan R, van Trijffel E, et al. State anxiety improves prediction of pain and pain-related disability after 12 weeks in patients with acute low back pain: a cohort study. J Physiother. 2020;66(1):39–44. https://doi.org/10.1016/j.jphys.2019.11.011.
45. Nim CG, Weber KA, Kawchuk GN, et al. Spinal manipulation and modulation of pain sensitivity in persistent low back pain: a secondary cluster analysis of a randomized trial. Chiropr Man Therap. 2021;29(1):10. https://doi.org/10.1186/s12998-021-00367-4.
46. Kizhakkeveettil A, Rose KA, Kadar GE, et al. An Exploratory Analysis of Gender as a Potential Modifier of Treatment Effect Among Patients in a Randomized Controlled Trial of Integrative Acupuncture and Spinal Manipulation for Low Back Pain. J Manipulative Physiol Ther. 2019;42(3):177–86. https://doi.org/10.1016/j.jmpt.2018.11.005.

47. Cook CE, Showalter C, Kabbaz V, et al. Can a within/between-session change in pain during reassessment predict outcome using a manual therapy intervention in patients with mechanical low back pain? Man Ther. 2012;17(4):325–9. https://doi.org/10.1016/j.math.2012.02.020.
48. Barakatt ET, Romano PS, Riddle DL, et al. The Reliability of Maitland's Irritability Judgments in Patients with Low Back Pain. J Man Manip Ther. 2009;17(3):135–40. https://doi.org/10.1179/jmt.2009.17.3.135.
49. Barakatt ET, Romano PS, Riddle DL, et al. An Exploration of Maitland's Concept of Pain Irritability in Patients with Low Back Pain. J Man Manip Ther. 2009;17(4):196–205. https://doi.org/10.1179/106698109791352175.
50. Kent P, Keating JL. Classification in nonspecific low back pain: what methods do primary care clinicians currently use? Spine (Phila Pa 1976) 2005;30(12):1433–40. https://doi.org/10.1097/01.brs.0000166523.84016.4b.
51. Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. Clin J Pain. 2003;19(2):80–6. https://doi.org/10.1097/00002508-200303000-00002.
52. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum. 2008;59(5):632–41. https://doi.org/10.1002/art.23563.
53. Haskins R, Osmotherly PG, Rivett DA. Validation and impact analysis of prognostic clinical prediction rules for low back pain is needed: a systematic review. J Clin Epidemiol. 2015;68(7):821–32. https://doi.org/10.1016/j.jclinepi.2015.02.003.
54. Bastos RM, Moya CR, de Vasconcelos RA, et al. Treatment-based classification for low back pain: systematic review with meta-analysis. J Man Manip Ther. 2022;30(4):207–27. https://doi.org/10.1080/10669817.2021.2024677.
55. Saragiotto BT, Maher CG, Moseley AM, et al. A systematic review reveals that the credibility of subgroup claims in low back pain trials was low. J Clin Epidemiol. 2016;79:3–9. https://doi.org/10.1016/j.jclinepi.2016.06.003.
56. Maissan F, Pool J, de Raaij E, et al. The clinical reasoning process in randomized clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Musculoskelet Sci Pract. 2018;35:8–17. https://doi.org/10.1016/j.msksp.2018.01.011.
57. Innes S, Beynon A, Hodgetts C, et al. Predictors of instanteous relief from spinal manipulation for non-specific low back pain: a delphi study. Chiropr Man Therap. 2020;28(1):39. https://doi.org/10.1186/s12998-020-00324-7.
58. Bade M, Cobo-Estevez M, Neeley D, et al. Effects of manual therapy and exercise targeting the hips in patients with low-back pain-A randomized controlled trial. J Eval Clin Pract. 2017;23(4):734–40. https://doi.org/10.1111/jep.12705.
59. 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. https://doi.org/10.1177/0269215509342328.
60. 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. https://doi.org/10.1016/j.pain.2006.12.008.
61. Fritz JM, Sharpe JA, Lane E, et al. Optimizing treatment protocols for spinal manipulative therapy: study protocol for a randomized trial. Trials. 2018;19(1):306. https://doi.org/10.1186/s13063-018-2692-6.
62. Kizhakkeveettil A, Rose KA, Kadar GE, et al. Integrative Acupuncture and Spinal Manipulative Therapy Versus Either Alone for Low Back Pain: A Randomized Controlled Trial Feasibility Study. J Manipulative Physiol Ther. 2017;40(3):201–13. https://doi.org/10.1016/j.jmpt.2017.01.002.
63. Licciardone JC, Minotti DE, Gatchel RJ, et al. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. Ann Fam Med. 2013;11(2):122–9. https://doi.org/10.1370/afm.1468.
64. Nim CG, Kawchuk GN, Schiottz-Christensen B, et al. The effect on clinical outcomes when targeting spinal manipulation at stiffness or pain sensitivity: a randomized trial. Sci Rep. 2020;10(1):14615. https://doi.org/10.1038/s41598-020-71557-y.
65. Machado LA, Maher CG, Herbert RD, et al. The effectiveness of the McKenzie method in addition to first-line care for acute low back pain: a randomized controlled trial. BMC Med. 2010;8:10. https://doi.org/10.1186/1741-7015-8-10.
66. Machado LA, Maher CG, Herbert RD, et al. The McKenzie Method for the management of acute non-specific low back pain: design of a randomised controlled trial [ACTRN012605000032651]. BMC Musculoskelet Disord. 2005;6:50. https://doi.org/10.1186/1471-2474-6-50.
67. Brealey S, Burton K, Coulton S, et al. UK Back pain Exercise And Manipulation (UK BEAM) trial–national randomised trial of physical treatments for back pain in primary care: objectives, design and interventions [ISRCTN32683578]. BMC Health Serv Res. 2003;3(1):16. https://doi.org/10.1186/1472-6963-3-16.
68. Team UBT. 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. https://doi.org/10.1136/bmj.38282.669225.AE.

Publisher's Note