

# A systematic review of orthopaedic manual therapy randomized clinical trials quality

Sean P. Riley<sup>1,2</sup>, Brian Swanson<sup>3</sup>, Jean-Michel Brismée<sup>4</sup>, Steven F. Sawyer<sup>4</sup>

<sup>1</sup>Department of Rehabilitation Services and Sports Medicine, Uconn Health, Farmington, CT, USA, <sup>2</sup>Physical Therapy Program, Storrs, CT, USA, <sup>3</sup>Department of Physical Therapy, University of New England, Portland, ME, USA, <sup>4</sup>Department of Rehabilitation Sciences and Center for Rehabilitation Research, School of Health Professions, Texas Tech University Health Sciences Center, Lubbock, TX, USA

*Study Design:* Systematic review and meta-analysis. *Objectives:* To conduct a systematic review and meta-analysis of randomized clinical trials (RCTs) in the orthopaedic manual therapy (OMT) literature from January 2010 to June 2014 in order to determine if the CONSORT checklist and Cochrane Risk of Bias (RoB) assessment tools: (1) are reliable; (2) have improved the reporting and decreased the risk of bias in RCTs in the OMT literature; (3) differ based on journal impact factor (JIF); and (4) scores are associated with each other. *Background:* The CONSORT statement is used to improve the accuracy of reporting within RCTs. The Cochrane RoB tool was designed to assess the risk of bias within RCTs. To date, no evaluation of the quality of reporting and risk of bias in OMT RCTs has been published. *Methods:* Relevant RCTs were identified by a literature review from January 2010 to June 2014. The identified RCTs were assessed by two individual reviewers utilizing the 2010 CONSORT checklist and the RoB tool. Agreement and a mean composite total score for each tool were attained in order to determine if the CONSORT and RoB tools were reliable and varied by year and impact factor. *Results:* A total of 72 RCTs in the OMT literature were identified. A number of categories within the CONSORT and RoB tools demonstrated prevalence-adjusted bias-adjusted kappa (PABAK) scores of less than 0.20 and from 0.20 to 0.40. The total CONSORT and RoB scores were correlated to each other ( $r = 0.73$ ; 95% CI 0.60 to 0.82;  $p < 0.0001$ ). There were no statistically significant differences in CONSORT or RoB scores by year. There was a statistically significant correlation between both CONSORT scores and JIF ( $r = 0.64$ , 95% CI 0.47 to 0.76;  $p < 0.0001$ ), and between RoB scores and JIF ( $r = 0.42$ , 95% confidence interval 0.21–0.60;  $p < 0.001$ ). There was not a statistically significant correlation between JIF and year of publication. *Conclusion:* Our findings suggest that the CONSORT and RoB have a number of items that are unclear and unreliable, and that the quality of reporting in OMT trials has not improved in recent years. Improvements in reporting are necessary to allow advances in OMT practice.

**Level of Evidence:** 1A

**Keywords:** CONSORT, Risk of bias, Manual therapy, Randomized clinical trials

## Introduction

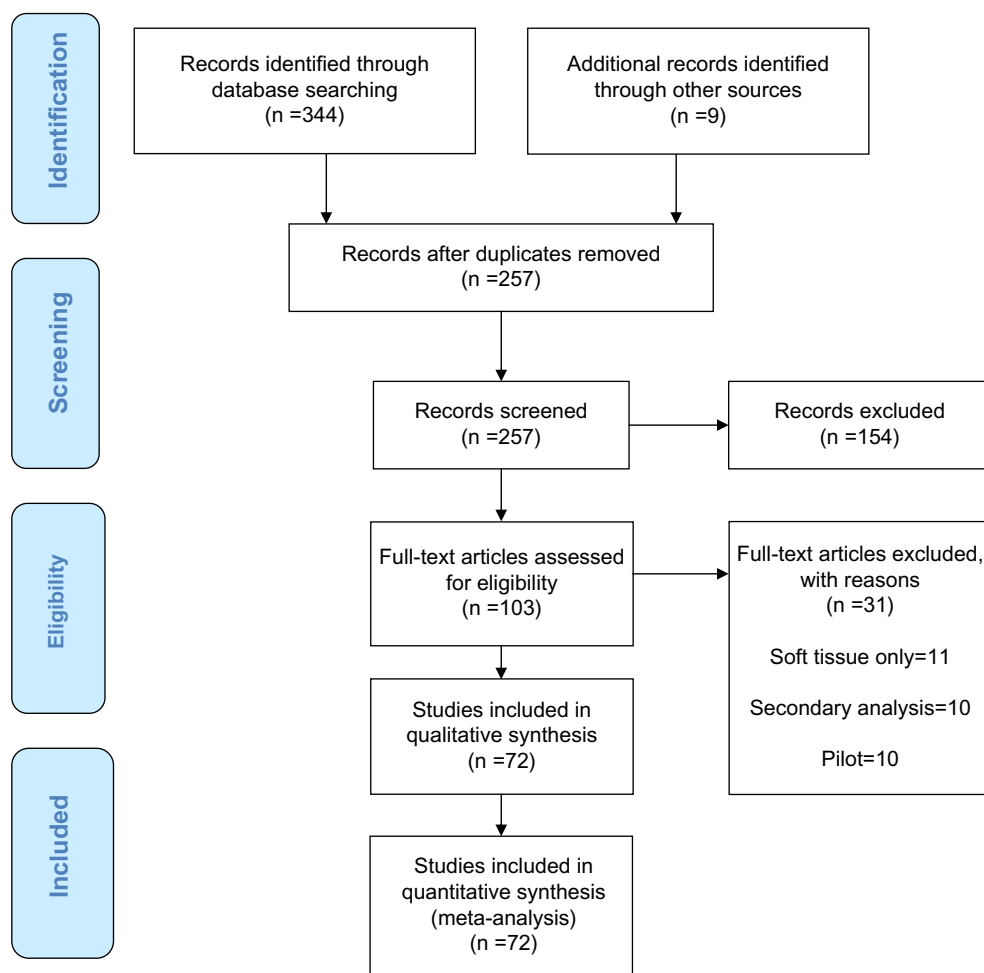
The CONSORT statement was born out of the recognition of inconsistencies of reporting in randomized controlled trials (RCTs). The CONSORT statement first made an appearance in the Journal of the American Medical Association in 1996<sup>1</sup> and the guidelines have been updated in 2001<sup>2</sup> and in 2010<sup>3</sup>. The purpose of the CONSORT statement was to facilitate complete and accurate reporting within RCTs in order to allow for critical appraisal and interpretation of findings.<sup>4</sup> These recommendations involve a checklist, a flow diagram and descriptive text.<sup>4</sup>

The Cochrane Risk of Bias (RoB) tool was designed to assess the risk of bias with seven items covering six domains within RCTs.<sup>5</sup> These domains include: generation

of the allocation sequence, concealment of the allocation sequence, blinding, incomplete outcome data, selective outcome reporting and other biases.<sup>5</sup> Unlike other assessment tools, each domain is assessed by a narrative explanation that is provided in order to assess each domain.<sup>5</sup> One of the purported features of the RoB tool is that it allows the RoB judgments to be displayed graphically across a number of studies that are included within a systematic review.<sup>5</sup> Armijo-Olivo *et al.*<sup>6</sup> found that the RoB tool demonstrated poor reliability when used to evaluate individual trials as well as meta-analyses. These findings were consistent with the low agreement that was found by Hartling and colleagues<sup>7</sup> when using the RoB tool to evaluate systematic reviews.

A Cochrane review of the utilization of the CONSORT statement in reporting RCTs published in medical journals

Correspondence to: Sean P. Riley, Musculoskeletal Institute 1st floor, 263 Farmington Ave, Farmington, Connecticut 06030, USA. Email: sriley@uchc.edu



**Figure 1 Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram studies included in quantitative synthesis (meta-analysis; n = 72).**

was completed in 2012.<sup>4</sup> This review found that the completeness of reporting of RCTs in the medical literature was still lacking in spite of the CONSORT statement being endorsed by the reviewed journals.<sup>4</sup> To date, no evaluation of the quality of orthopaedic manual therapy (OMT) RCTs has been published in journals requiring the use of CONSORT statement for RCTs.

The purposes of this study were to determine if: (1) the CONSORT checklist is reliable; (2) the RoB assessment tool is reliable; (3) the CONSORT checklist and RoB assessment tool have improved the reporting and risk of bias in RCTs in the OMT literature from January 2010 to June 2014; (4) the CONSORT checklist and RoB assessment tool scores are different based on rehabilitation-related journal impact factor; and (5) the CONSORT checklist score is associated with the RoB assessment tool score. In addition, an item by item analysis was used to allow recommendations on individual items within each of these tools. These questions were assessed by conducting a systematic review of RCTs in OMT from January 2010 to June 2014, and then grading each of the selected papers for their inclusion of the individual CONSORT checklist and RoB tool items.

## Methods

Our methodology is consistent with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines statement (Figure 1).<sup>8</sup> Relevant RCTs were identified by a literature review utilizing PubMed, CINHAL and Google Scholar from January 2010 to June 2014. The years of the search were selected based on the latest revision CONSORT statement and the endorsement of its use by journals that publish OMT trials. The search was narrowed by utilizing the search terms, “manual therapy” and “randomized”, to identify key terms within the title consistent with the 2010 CONSORT checklists. The abstracts were then screened to ensure that the study was a RCT. A professional librarian was consulted in order to ensure impartiality and accuracy of the search.

The definition of each individual CONSORT checklist item was clarified between reviewers to allow for accurate reporting. The individual checklist items were determined to be present, absent or unclear based on the information reported. All RCTs were identified based on the title and abstract of the manuscripts in accordance with the above described search strategy. The identified RCTs were assessed by two individual reviewers (SR and BS) without knowledge of each other’s work. Individual

**Table 1 Scoring and consensus method for the CONSORT checklist and Cochrane risk of bias (RoB) tool**

Item number	CONSORT checklist requirements
1a	Yes/no determination. The title of the RCTs was assessed by identifying if the term "randomized trial" is identified in the title
1b	Was the summary broken down into sections? Yes/No. The abstract was assessed by ensuring that it contained all 4 components as identified by the CONSORT statement, which included: "Trial Design", "Methods", "Results", and "Conclusions"
2a	Discussion, did background adequately explain the need for/purpose of the study? Yes/No. The background and objectives of the introduction were scored by ensuring the "scientific background and explanation of rationale"
2b	After discussion, could we identify the <i>specific</i> research questions? Yes/No. "Specific objectives or hypotheses" were identified
3a	If the study was a simple 2 group design, then designation of "randomized controlled trial" was deemed adequate. Other designs require more detail to meet the standard. The trial design was scored based on the "Description of trial design (such as parallel, factorial) including allocation ratio"
3b	Simple yes/no determination; information present or not. "Important changes to methods after trial commencement (such as eligibility criteria), with reasons"
4a	The study was deemed to meet this standard if the inclusion/exclusion criteria were included in a detailed fashion. The identification of the participants was scored by identifying "Eligibility criteria for participants"
4b	Could we determine the geographic location and study setting from the text? Yes/No. The identification of the participants was scored by identifying settings and locations where the data were collected
5	If the study was pragmatic in nature, then "yes" was determined for this study. For other trials, a yes was given if there was enough detail to allow a similarly trained researcher to reproduce the trial. The interventions were scored by using the definition, "The interventions for each group with sufficient details to allow replication, including how and when they were actually administered"
6a	Were the reported outcomes and their collection times listed in the methods section? The outcomes were scored using "Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"
6b	Simple yes/no determination; information present or not. The outcomes were scored using "Any changes to trial outcomes after the trial commenced, with reasons"
7a	Simple yes/no determination; information present or not. Was the sample size calculation performed using an accepted method? Sample size was evaluated by identifying "How sample size was determined"
7b	Simple yes/no determination; information present or not. "When applicable, explanation of any interim analyses and stopping guidelines"
8a	Discussion regarding the specific method of randomization was required. Randomization was graded by identifying "Method used to generate the random allocation sequence"
8b	If 2 groups only, then "subjects randomized by ..." was acceptable. More complex studies required appropriate details regarding sampling methodology. Randomization was graded by identifying "Type of randomisation; details of any restriction (such as blocking and block size)"
9	Specific details of sequence allocation were required. Allocation concealment was identified through "Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned"
10	Specific details regarding the study personnel generating sequence and enrolling subjects was required. Implementation of randomization was assessed by "Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions"
11a	Specific details regarding who was blinded, and their specific role in the study, were required. Blinding "If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how"
11b	For comparative trials with multiple actual interventions, the details of the interventions adequate to reproduce them were considered adequate. For placebo trials, the description of the similarity of the placebo to treatment was required. "If relevant, description of the similarity of interventions"
12a	Simple yes/no determination; information present or not. Statistical methods used were graded by "Statistical methods used to compare groups for primary and secondary outcomes"
12b	If a study was registered, did they perform the predetermined measures, and did they identify any additional analysis performed? If not registered, then the trial scored "no" as there is no way to determine exploratory analysis. "Methods for additional analyses, such as subgroup analyses and adjusted analyses"
13a	Was the "n" per group detailed in results or participant flow diagram, or could the "n"; per group be gleaned by reading the text? If not, then No. Results were assessed by identifying "For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome"
13b	If both "n" per group and details of reasons for loss were included, then yes. If did not meet both criteria, then No. "For each group, losses and exclusions after randomisation, together with reasons"
14a	Simple yes/no determination; information present or not. Recruitment "Dates defining the periods of recruitment and follow-up"
14b	If either detailed explanation of interim stopping guidelines or the study met full a priori sample size, then yes. If the study did not meet sample calculations, and no explanation given, then No. "Why the trial ended or was stopped"
15	Simple yes/no determination; information present or not. Baseline data were assessed for "A table showing baseline demographic and clinical characteristics for each group"
16	Simple yes/no determination; information present or not. "For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups"
17a*	If sample reported either Cohen's d or 95% CI for outcomes, then yes. All others, No. "For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)"
17b*	Yes was given for those studies reporting absolute and relative values, or not using binary outcome measures. No was given to studies using binary outcomes but not reporting relative values. "For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)"

Item number	CONSORT checklist requirements
18	If trial is registered, were outcomes the same as those in protocol? No was given to studies either (a) not registered, (b) including additional analysis, and (c) registered retrospectively. "Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory"
19	Simple yes/no determination; information present or not. "All important harms or unintended effects in each group"
20	Discussion of limitations, where yes was deemed to be relevant and applicable, while no represented either not present or not adequately addressed. "Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses"
21	Yes was given if reporting was appropriate to findings, no for studies not deemed appropriate to findings. "Generalizability (external validity, applicability) of the trial findings"
22	Yes was given if reporting was appropriate to findings, no for studies not deemed appropriate to findings. "Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence"
23	Simple yes/no determination; information present or not. "Registration number and name of trial registry"
24	Yes = authors able to find trial registration; no = unable to locate trial registration. "Where the full trial protocol can be accessed, if available"
25	Simple yes/no determination; information present or not. "Sources of funding and other support (such as supply of drugs), role of funders"
	<i>Cochrane risk of bias (RoB) tool requirement</i>
1	Was the sequence generation adequately reported, and could it be determined as meeting the low risk criteria?
2	Could the allocation method be determined? If so, was it low risk?
3	Could it be determined who was blinded? If blinded outcomes assessor reported, then trial given a "yes"
4	Did the authors report their handling of missing data, and if so, was the method appropriate?
5	Did the study report the pre-registered outcomes? A no was given if the study (a) reported extra outcomes beyond protocol, (b) was not registered, or (c) was registered retrospectively
6	If any points 1–5 had received a "no", then category 6 was determined to be "no". Additionally, if any sources of bias were noted in reporting, then "no" as well

\*If present, the effect size was recorded. If absent, an attempt was made to calculate the effect size based on the data presented in the manuscript. If the effect size was not reported and could not be calculated it was left blank.

items were scored as absent, unclear if present or present. This work was then sent to a third reviewer to assess agreement. Agreement was determined by original scoring (absent, unclear if present or present), as well as by collapsing the scoring of absent and unclear into one variable, absent. Final agreement was determined through the grading of absent or present. Following review by the third reviewer, the two initial reviewers were made aware of the discrepancies and subsequently reviewed and discussed each point of disagreement in an effort to reach consensus regarding the quality of each article (Table 1). Once agreement was determined, the two individual reviewers again provided the results of the individual reviews to the third reviewer. All reviewers had greater than 11 years of clinical experience, hold advanced certifications in orthopaedics and manual therapy and have attained terminal doctoral degrees.

The RCTs were scored using the 2010 CONSORT checklist. Thirty-seven individual items were assessed as outlined by the CONSORT checklist with a minimal possible score of 0 and a total maximal score if all items were included of 37. The reporting of the individual items is assessed as described in Table 1.

Risk of Bias was assessed utilizing the seven-item, six-domain RoB tool described by the Cochrane Collaboration.<sup>9</sup> The seven items included in the six domains are: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete outcome data, selective outcome reporting and other biases. In addition, the clarifying guidelines reported by Armijo-Olivo *et al.*<sup>6</sup> were incorporated. Consistent with the grading of the CONSORT individual items were scored as absent, unclear if present

or present. These scores were then assessed and collapsed as described for the CONSORT scoring. Seven individual items were assessed as outlined above with a minimal possible score of 0 and a total maximal score of 7.

Once the scores on the CONSORT and RoB were determined, a composite mean total score was calculated by year from January 2010 to June 2014 in order to determine if the scores had improved as a result of the progressive endorsement of the use of the CONSORT statement by journals publishing RCTs in OMT and to determine if the risk of bias had also improved. In addition, journals were assessed by impact factor as defined by Thomson Reuters in order to determine if the quality of reporting and risk of bias varied based on this variable.

Data management inferential statistical analyses and graphing were performed with Microsoft Excel 2010, IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, NY), MedCalc version 15.8 (Medcalc Software, Ostend, Belgium), GraphPad InStat version 3.06 for Windows (GraphPad Software, San Diego California), SigmaPlot version 12.5 (Systat Software, Inc., San Jose California), and the online Kappa calculator at [graphpad.com/quickcalcs/kappa1/](http://graphpad.com/quickcalcs/kappa1/).

## Results

### Literature review

A professional librarian was consulted to perform a complete and unbiased search of the literature regarding randomized trials in manual therapy from January 2010 to June 2014. The librarian conducted a search using the following databases: Pubmed, CINAHL complete and Google scholar. The Pubmed MeSH term, "musculoskeletal manipulations", was utilized since this is the medical

subject heading which contains the term “manual therapy”. The search was performed with the following filters: “randomized”, “clinical trial”, language: English, full text, human subjects, NOT “soft tissue” and NOT “applied kinesiology”. The date of the search was 06/23/2014: This search resulted in 243 articles. After titles’ screening, it was determined that 104 articles were possibly appropriate for this review. The CINAHL search of “manual therapy” with the filters, “randomized”, English language, human subject, peer reviewed, full text, NOT “myofascial release” and NOT “massage” resulted in 54 articles; after screening for duplication and appropriateness for this study, three additional articles were added to the review. An identical search of Google Scholar resulted in 47 total articles, after screening for duplication and appropriateness for this study, two additional articles were included in the review. The three search engines resulted in a total of 344 potential articles. Following screening by title, 109 articles were selected to screen by their abstract. After screening the abstracts, a total of 98 articles were included for review of the full text. Additionally, a hand review was undertaken to ensure that the scope was as comprehensive as possible. Using the ULRICH subject database, a search of “manual therapy” and “randomized trial” was performed. This yielded a possible total of 425 potential journals, limited to 37 titles following removal of duplicate titles, journals not available in English and those not appropriate to the topic. A hand search of the journals that were not duplicated in our previous searches yielded a total of 9 additional articles since 2010, 5 of which were retained after screening of the abstracts for a total of 103 possible titles. A total of 30 articles were excluded during the review of full text as they did not meet the inclusion criteria (Figure 1), with the subsequent review covering a total of 72 articles. (Appendix 1)

### **Stage 1 analysis: inter-rater agreement and reliability of independent ratings with three designations (“Present”, “Unclear if Present” or “Absent”)**

The analysis of the 72 published research reports involved the two independent raters assigning a designation of either “Present”, “Unclear if Present” or “Absent” to each of the 37 CONSORT categories and the 6 RoB categories for each report. For CONSORT guideline categories, 62% of the designations by the two raters were “Present”, 23% were “Unclear if Present” and 15% were “Absent.” For RoB, the designations were 40, 54 and 6%, respectively. The distribution of CONSORT designations for Rater 1 was 52% as “Present”, 31% as “Unclear if Present” and 17% as “Absent”, whereas the corresponding distribution for Rater 2 was 71, 15 and 13%. For RoB, the Rater 1 distributions were 50, 49 and 2% for these categories, and distributions for Rater 2 were 29, 61 and 10%, respectively.

Rater agreement on the “Present”, “Unclear if Present” or “Absent” designations was  $72 \pm 23\%$  (mean  $\pm$  S.D.) for all

CONSORT categories (median = 79%, range = 18–97%), and  $66 \pm 16\%$  for all RoB categories (median = 69%, range = 36–85%). Mean Kappa inter-rater reliability scores for CONSORT categories was  $0.33 \pm 0.27$  (median = 0.24, range = -0.04 to 0.89), which can be considered to be a collective “fair” degree of reliability.<sup>13</sup> For RoB categories, mean Kappa inter-rater reliability scores were  $0.30 \pm 0.22$  (median = 0.30, range = 0.04–0.66), which likewise is a “fair” degree of reliability.<sup>13</sup>

There were many instances of low kappa scores secondary to the manner by which agreements between the two raters accumulated, with a disproportionate number of “Present”, “Unclear if Present” or “Absent” designations for selected items; in other words, there was prevalence and/or bias for many categories.<sup>12</sup> For this reason, a prevalence-adjusted bias-adjusted kappa (PABAK) assessment in the reliability analysis was performed.<sup>10</sup> PABAK scores for the CONSORT categories averaged  $0.58 \pm 0.34$  (median = 0.69, range = -0.23 to 0.96), which can be considered to be a “moderate” to “substantial” levels of reliability.<sup>13</sup> PABAK scores for the six RoB categories averaged  $0.49 \pm 0.24$  (median = 0.53, range = 0.04–0.77), which can be considered to be “moderate” level of reliability.

### **Stage 2 analysis: inter-rater agreement and reliability of independent ratings with two designations (present, absent)**

The ratings assessment involved equating the “Unclear if Present” designation with the “Absent” designation. As such, there were two rater designations in this analysis: “Present” and “Absent”, which allowed tallying category scores for the CONSORT and RoB guidelines (1 point for each category if deemed present in a research report) to yield CONSORT and RoB scores for each research report. Rater 1 CONSORT and RoB scores averaged  $19.2 \pm 6.0$  (out of maximum of 37) and  $2.99 \pm 1.72$  (out of maximum of 6), respectively, whereas Rater 2 scores averaged  $26.6 \pm 5.8$  and  $1.78 \pm 1.26$ , respectively.

Agreement and reliability findings for the “Present” and “Absent” designations are summarized in Tables 2 and 3. Per cent agreement for raters on CONSORT categories was  $78 \pm 18\%$  (median = 82%, range = 33–97%), with Kappa scores of  $0.40 \pm 0.29$  (median = 0.32, range = -0.03 to 0.94) and PABAK scores of  $0.56 \pm 0.36$  (median = 0.64, range = -0.33 to 0.94). For RoB categories, per cent agreement for raters was  $73 \pm 15\%$  (median = 75%, range = 46–86%), with Kappa scores of  $0.30 \pm 0.25$  (median = 0.28, range = -0.03 to 0.68) and PABAK scores of  $0.46 \pm 0.31$  (median = 0.50, range = -0.08 to 0.72).

A number of categories within the CONSORT and RoB tools demonstrated PABAK scores indicative of slight (less than 0.20) to fair (0.2–0.40) degrees of reliability.<sup>11</sup> For CONSORT, the corresponding categories were 3a, 5, 8b, 10, 11a, 11b, 12b, 14b, 17a, 18 and 21; the RoB categories were “Allocation Concealment”, “Blinding of Participants and Personnel” and “Incomplete Outcome Data.”

**Table 2** Rater assessment of research reports based on the CONSORT guidelines.

Item number	CONSORT inter-rater agreement and reliability				CONSORT compliance		
	% Agreement (present or absent)	Kappa (95% CI)	Prevalence index	Bias index	PABAK (95% CI)	% Present rater 1	% Present rater 2
1a	96	0.82 (0.62, 1.00)	0.74	0.01	0.92 (0.82, 1.00)	86	88
1b	93	0.81 (0.66, 0.97)	0.51	0.07	0.86 (0.74, 0.98)	72	79
2a	96	Undefined	0.96	0.04	0.92 (0.82, 1.00)	96	100
2b	85	0.13 (-0.10, 0.37)	0.82	0.15	0.69 (0.53, 0.86)	83	99
3a	36	Undefined	0.36	0.64	-0.28 (-0.50, -0.06)	36	100
3b	89	0.16 (-0.16, 0.49)	0.86	0.08	0.78 (0.63, 0.92)	3	11
4a	97	Undefined	0.97	0.03	0.94 (0.87, 1.00)	97	100
4b	81	0.49 (0.26, 0.71)	0.50	0.08	0.61 (0.43, 0.79)	71	79
5	63	0.23 (0.09, 0.38)	0.40	0.38	0.25 (0.03, 0.47)	51	89
6a	89	0.18 (-0.13, 0.49)	0.86	0.11	0.78 (0.63, 0.92)	88	99
6b	94	-0.02 (-0.05, 0.01)	0.94	0.03	0.89 (0.78, 1.00)	1	4
7a	88	0.74 (0.58, 0.90)	0.24	0.10	0.75 (0.60, 0.90)	57	67
7b	92	0.53 (0.22, 0.85)	0.81	0.08	0.83 (0.71, 0.96)	6	14
8a	82	0.55 (0.34, 0.76)	0.46	0.10	0.64 (0.46, 0.82)	68	78
8b	49	0.14 (0.02, 0.26)	0.21	0.49	-0.03 (-0.26, 0.20)	36	85
9	76	0.51 (0.32, 0.70)	0.26	0.18	0.53 (0.33, 0.72)	54	72
10	67	0.38 (0.22, 0.55)	0.11	0.31	0.33 (0.12, 0.55)	29	60
11a	69	0.32 (0.11, 0.53)	0.36	0.17	0.39 (0.18, 0.60)	60	76
11b	67	0.23 (0.00, 0.05)	0.11	0.31	0.33 (-0.79, -0.43)	11	92
12a	94	Undefined	0.94	0.06	0.89 (0.78, 1.00)	94	100
12b	43	0.1 (-0.02, 0.22)	0.04	0.51	-0.14 (-0.37, 0.09)	22	74
13a	89	0.27 (-0.09, 0.64)	0.83	0.03	0.78 (0.63, 0.92)	90	93
13b	82	0.52 (0.31, 0.74)	0.51	0.15	0.64 (0.46, 0.82)	68	83
14a	92	0.83 (0.70, 0.96)	0.11	0.06	0.83 (0.71, 0.96)	53	58
14b	44	0.07 (-0.04, 0.19)	0.22	0.50	-0.11 (-0.34, 0.12)	14	64
15	90	0.64 (0.40, 0.88)	0.68	0.07	0.81 (0.67, 0.94)	81	88
16	79	0.19 (-0.07, 0.44)	0.71	0.13	0.58 (0.40, 0.77)	79	92
17a	61	0.3 (0.13, 0.46)	0.06	0.33	0.22 (0.00, 0.45)	31	64
17b	81	-0.03 (-0.08, 0.02)	0.81	0.17	0.61 (0.43, 0.79)	1	18
18	33	0.03 (-0.05, 0.10)	0.19	0.64	-0.33 (-0.55, -0.12)	8	72
19	93	0.86 (0.74, 0.98)	0.10	0.04	0.86 (0.74, 0.98)	53	57
20	89	0.55 (0.28, 0.81)	0.72	0.11	0.78 (0.63, 0.92)	81	92
21	68	0.18 (0.02, 0.34)	0.57	0.32	0.36 (0.15, 0.58)	63	94
22	76	0.29 (0.08, 0.50)	0.63	0.24	0.53 (0.33, 0.72)	69	93
23	97	0.94 (0.86, 1.00)	0.25	0.03	0.94 (0.87, 1.00)	39	36
24	78	0.42 (0.20, 0.65)	0.50	0.14	0.56 (0.36, 0.75)	18	32
25	93	0.86 (0.74, 0.98)	0.04	0.04	0.86 (0.74, 0.98)	50	54

### Stage 3 analysis: unified rater ratings with two designations (present and absent)

The next step in the inter-rater analysis involved the raters conferring with each other about their independent Stage 2 (“Present” and “Absent”) designations and seeking to find agreement for all items. When agreement was not attained, a third rater served as the deciding vote, yielding a singular set of designations for CONSORT and RoB categories for the 72 research reports. The statistical analysis described below pertains to this Stage 3 data-set.

### Summary of unified CONSORT and RoB scores

The unification of the two rater scores into a singular set of “Present” and “Absent” designations yielded a mean CONSORT score of  $24.6 \pm 6.2$  out of a maximum of 37 (median = 26.0, range = 6–35). RoB scores (maximum of 6) averaged  $2.69 \pm 1.32$  (median = 3.00, range 0–5). Averages for CONSORT categories ranged from 0.03 (mostly “Absent”) to 1.00 (all “Present”), with the following categories notable for averages less than 0.2 (i.e. 80% of research articles rated as being “Absent” for the

category): 6b, 7b, 12b, 17b and 18. For RoB, the “Selective Outcome Reporting” and “Other Potential Threats to Validity” categories were notable for low scores, averaging 0.07 and 0.00, respectively.

### Correlations of CONSORT and RoB scores

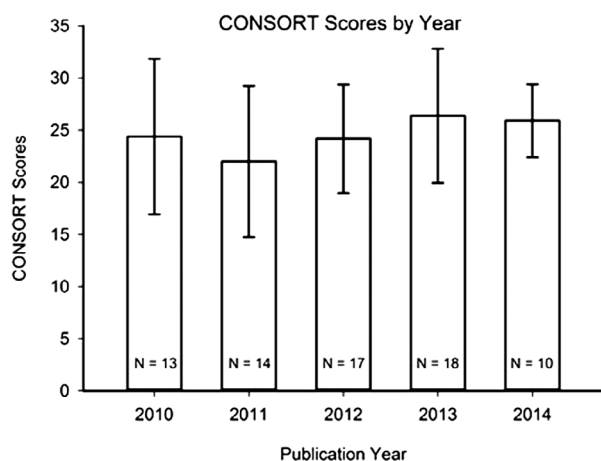
The scores for the CONSORT and RoB categories were summed to yield a CONSORT score and RoB score for each of the 72 research articles. These scores were correlated to each other in a statistically significant manner:  $r = 0.73$ ; 95% confidence interval 0.60–0.82;  $p < 0.0001$ .

### CONSORT and RoB scores for research reports published from 2010 to 2014

The research reports that were evaluated in this study were published between January 2010 and June 2014, inclusive. The research reports’ CONSORT and RoB scores were analysed for changes over this five-year time period; see Figures 2 and 3. An ANOVA revealed no statistically significant differences in CONSORT scores with respect to year of publication:  $F(4,67) = 1.11$ ;  $p = 0.36$ ; since this

**Table 3** Rater assessment of research reports based on the risk of bias guidelines.

Risk of bias item	Risk of bias inter-rater agreement and reliability				Risk of bias compliance		
	% Agreement (present or absent)	Kappa (95% CI)	Prevalence index	Bias index	PABAK (95% CI)	% Present rater 1	% Present rater 2
Sequence generation	86	0.68 (0.51, 0.86)	0.36	0.08	0.72 (0.56, 0.88)	72	64
Allocation concealment	69	0.44 (0.28, 0.60)	0.03	0.31	0.39 (0.18, 0.60)	67	36
Blinding of participants, personnel and outcome assessors	46	0.16 (0.06, 0.26)	0.04	0.54	-0.08 (-0.31, 0.15)	75	21
incomplete outcome data	69	0.39 (0.18, 0.60)	0.03	0.00	0.39 (0.18, 0.60)	51	51
Selective outcome reporting	81	0.16 (-0.07, 0.40)	0.75	0.17	0.61 (0.43, 0.79)	21	4
Other potential threats to validity	86	-0.03 (-0.07, 0.02)	0.86	0.11	0.72 (0.56, 0.88)	13	1



**Figure 2** CONSORT scores from Stage 3 of analysis (as described in Results section) of the 72 rated research reports, summarized by publication year. Mean  $\pm$  Standard Deviation. N = number of research reports in given year. There were no statistically significant differences in CONSORT scores across publication years.

data-set did not pass the normality test, a non-parametric Kruskal–Wallis ANOVA was run:  $H = 4.46$ ;  $p = 0.35$ . Likewise, RoB scores showed no differences across years:  $F(4,67) = 0.83$ ;  $p = 0.51$ ; since this data-set did not pass the normality test, a Kruskal–Wallis ANOVA was run:  $H = 2.43$ ;  $p = 0.66$ .

### Correlation of CONSORT and RoB scores with journal impact factor scores

Journal Impact Factor (JIF) values for the 72 research reports averaged  $2.04 \pm 2.04$  (median = 1.90). There was one outlier in JIF scores in the data-set of 72 research article with a JIF of 16.1; this outlier datum value was not included in the correlation analysis of JIF with CONSORT and RoB scores. For the remaining 71 articles, there was a statistically significant correlation between both CONSORT scores and JIF ( $r = 0.64$ , 95% confidence

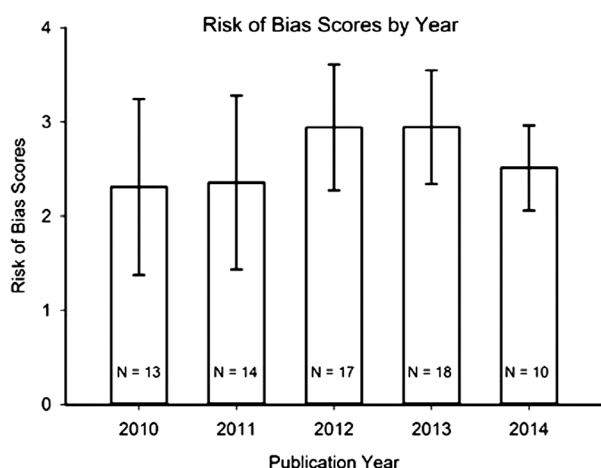
interval 0.47–0.76;  $p < 0.0001$ ), and between RoB scores and JIF ( $r = 0.42$ , 95% confidence interval 0.21–0.60;  $p < 0.001$ ). (Note: There was not a statistically significant correlation between JIF and Year of Publication [ $r = 0.12$ ,  $p = 0.30$ ; JIF outlier excluded], so JIF was not a confound in the one-factor ANOVAs discussed above that used CONSORT and RoB as dependent variables and Year of Publication as the independent variable; in other words, there was no need to perform ANCOVAs with JIF as a co-variate.)

### Discussion

The importance of accuracy in reporting and minimizing bias in research publications is related to the ability to report the research with sufficient detail to allow researchers to learn from the mistakes of their colleagues and to replicate positive findings. This basic premise of the scientific method has been recognized for hundreds of years in the basic sciences,<sup>12</sup> yet has not been fully embraced in the medical and health sciences literature until relatively recently.<sup>4,13–22</sup>

This is the first study to assess the quality of reporting and bias of published RCTs in the OMT peer-reviewed literature. The suboptimal levels of reporting and risk of bias found in OMT RCTs in this study have also been observed in RCTs published in medical journals across many disciplines.<sup>5,7,13–22</sup> These findings suggest that the current use of the CONSORT and RoB is less than optimal.

Beyond the quality of reporting and risk of bias, our results showed that items 3a (trial design), 5 (details about intervention allowing study replication), 8b (type of randomization), 10 (random allocation sequence generation), 11a and 11b (blinding information), 12b (methods of additional analyses), 14b (reason why the trial was stopped), 17a (effect size reporting of outcomes), 18 (subgroup analyses reporting) and 21 (external validity of trial findings) on the CONSORT statement displayed poor reliability



**Figure 3 Risk of Bias scores from Stage 3 of analysis (as described in Results section) of the 72 rated research reports, summarized by publication year. Mean  $\pm$  Standard Deviation. N = number of research reports in given year. There were no statistically significant differences in Risk of Bias scores across publication years.**

and that the quality of the reporting did not improve over time following the widespread adoption of CONSORT. For the RoB, the categories, “Allocation Concealment”, “Blinding of Participants and Personnel” and “Incomplete Outcome Data”, had similar findings suggesting the need for improvements in the reporting of published RCTs in the OMT peer-reviewed literature so the information can be used in systematic reviews and meta-analyses.

The present state of reporting, as well as the tools used to assess this reporting, limit advances in knowledge and practice across many disciplines in medical and health sciences. Our findings suggest that the CONSORT and Rob have a number of items that are unclear and unreliable in their definition and reporting in OMT literature. If these tools are not sufficiently reliable and valid, they are not useful for the purpose of reporting and reducing the risk of bias in the literature. One possible solution would be to remove or attempt to improve the individual items that were found to be unreliable. Such a modification may create a less subjective tool, with more stringently defined criteria for each category. A second possible solution lies within the publishing journals, which are in a position to ensure that the key CONSORT and RoB criteria are met prior to publication. If the journals that endorse and/or require the use of the CONSORT and RoB tools do not ensure that the key criteria are met during the review process, we cannot expect to see meaningful changes in the quality of reporting in these journals. A third potential solution may be to standardize the editorial and review process. Removing details in order to achieve a specific word count may influence compliance with the CONSORT and RoB as well as introduce selective reporting bias. When this bias is introduced, it may not be possible to compare the published RCT to the trial registry. In addition, statistical reporting of p values versus effect sizes

with 95% confidence intervals may be subject to the preferences and biases of the editor and reviewers during the review process making it difficult to compare RCTs in the OMT literature.

### Limitations

The major limitation of this study is the generalizability of our findings. In order to attain a reasonable number of research articles to allow a practical literature review and statistical analysis, we limited our literature search to RCTs focusing on OMT directed at articular structures. We therefore cannot generalize our findings to all journals and research topics that utilize the CONSORT and RoB.

### Conclusions

The accuracy of reporting and minimizing bias in research publications is critical to the advancement of OMT practice. If research articles are not published in sufficient detail to allow for replication and to assess bias, researchers will be unable to build upon the established foundation of literature. Our findings suggest that the CONSORT and RoB have a number of items that are unclear and unreliable. Without a strong, reliable foundation of detailed reporting in the literature, progress may not be possible. If the quality of reporting and risk of bias have not improved over time, it may be time to reassess the process. Peer-reviewed journals claiming to adhere to the CONSORT statement should ensure that authors are required to submit evidence of the reporting of each CONSORT item within their manuscript.

### Acknowledgement

We would like to thank Chad Cook, PT, PhD for his assistance with the conception and design of this study.

### Funding

No funding was provided for the development of this manuscript.

### References

- Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996;276:637–639.
- Moher D, Schulz KF, Altman D, Group C. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:1987–1991.
- Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010;8:18.
- Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev*. 2012;1:60.
- Savović J, Weeks L, Sterne JA, Turner L, Altman DG, Moher D, et al. Evaluation of the Cochrane collaboration’s tool for assessing the risk of bias in randomized trials: focus groups, online survey, proposed recommendations and their implementation. *Syst Rev*. 2014;3:37.
- Armijo-Olivo S, Ospina M, da Costa BR, Egger M, Saltaji H, Fuentes J, et al. Poor reliability between Cochrane reviewers and blinded external reviewers when applying the cochrane risk of bias tool in physical therapy trials. *PLoS ONE*. 2014;9:e96920.



- 7 Hartling L, Hamm MP, Milne A, Vandermeer B, Santaguida PL, Ansari M, et al. Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. *J Clin Epidemiol.* 2013;66:973–981.
- 8 Panic N, Leoncini E, de Belvis G, Ricciardi W, Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS ONE.* 2013;8:e83138.
- 9 Furlan AD, Pennick V, Bombardier C, van Tulder M, Editorial Board CBRG, 2009 updated method guidelines for systematic reviews in the Cochrane back review group. *Spine.* 2009;34:1929–1941.
- 10 Flight L, Julious SA. The disagreeable behaviour of the kappa statistic. *Pharm Stat.* 2015;14:74–78.
- 11 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
- 12 Straulino S. Reconstruction of Galileo Galilei's experiment: the inclined plane. *Phys Edu.* 2008;43:316–321.
- 13 Hajibandeh S, Hajibandeh S, Antoniou GA, Green PA, Maden M, Torella F. Reporting and methodological quality of randomised controlled trials in vascular and endovascular surgery. *Eur J Vasc Endovasc Surg.* 2015; 50:664–670.
- 14 Carmichael K, Nolan SJ, Weston J, Tudur Smith C, Marson AG. Assessment of the quality of harms reporting in non-randomised studies and randomised controlled studies of topiramate for the treatment of epilepsy using CONSORT criteria. *Epilepsy Res.* 2015;114:106–113.
- 15 Koch M, Riss P, Umek W, Hanzal E. CONSORT and the internal validity of randomized controlled trials in female pelvic medicine. *NeuroUrol Urodyn.* 2015. [Epub ahead of print].
- 16 Kloukos D, Papageorgiou SN, Doulis I, Petridis H, Pandis N. Reporting quality of randomised controlled trials published in prosthodontic and implantology journals. *J Oral Rehabil.* 2015. [Epub ahead of print].
- 17 Zhai X, Wang Y, Mu Q, Chen X, Huang Q, Wang Q, et al. Methodological reporting quality of randomized controlled trials in 3 leading diabetes journals from 2011 to 2013 following CONSORT statement: a system review. *Medicine.* 2015;94:e1083.
- 18 Jull A, Aye PS. Endorsement of the CONSORT guidelines, trial registration, and the quality of reporting randomised controlled trials in leading nursing journals: a cross-sectional analysis. *Int J Nurs Stud.* 2015;52:1071–1079.
- 19 Lu J, Gary KW, Copolillo A, Ward J, Niemeier JP, Lapane KL. Randomized controlled trials in adult traumatic brain injury: a review of compliance to CONSORT statement. *Arch Phys Med Rehabil.* 2015;96:702–714.
- 20 Rajasekharan S, Vandenbulcke J, Martens L. An assessment of the quality of reporting randomised controlled trials published in paediatric dentistry journals. *Eur Arch Paediatr Dent.* 2015;16:181–189.
- 21 Carlton DA, Kocherginsky M, Langerman AJ. A systematic review of the quality of randomized controlled trials in head and neck oncology surgery. *Laryngoscope.* 2015;125:146–152.
- 22 Peters JP, Hooft L, Grolman W, Stegeman I. Assessment of the quality of reporting of randomised controlled trials in otorhinolaryngologic literature – adherence to the CONSORT statement. *PLOS ONE.* 2015;10:e0122328.
4. Bicalho E, Setti JA, Macagnan J, Cano JL, Manffra EF. Immediate effects of a high-velocity spine manipulation in paraspinal muscles activity of nonspecific chronic low-back pain subjects. *Man Ther.* 2010;15:469–475.
5. Jowsey P, Perry J. Sympathetic nervous system effects in the hands following a grade III postero-anterior rotatory mobilisation technique applied to T4: a randomised, placebo-controlled trial. *Man Ther.* 2010;15:248–253.
6. Bennell KL, Matthews B, Greig A, Briggs A, Kelly A, Sherburn M, et al. Effects of an exercise and manual therapy program on physical impairments, function and quality-of-life in people with osteoporotic vertebral fracture: a randomised, single-blind controlled pilot trial. *BMC Musculoskelet Disord.* 2010;11:36.
7. Bennell K, Wee E, Coburn S, Green S, Harris A, Staples M, et al. Efficacy of standardised manual therapy and home exercise programme for chronic rotator cuff disease: randomised placebo controlled trial. *BMJ.* 2010;340:c2756.
8. Cecchi F, Molino-Lova R, Chiti M, Pasquini G, Paperini A, Conti AA, et al. Spinal manipulation compared with back school and with individually delivered physiotherapy for the treatment of chronic low back pain: a randomized trial with one-year follow-up. *Clin Rehabil.* 2010;24:26–36.
9. Leaver AM, Maher CG, Herbert RD, Latimer J, McAuley JH, Jull G, et al. A randomized controlled trial comparing manipulation with mobilization for recent onset neck pain. *Arch Phys Med Rehabil.* 2010;91:1313–1318.
10. Pool JJ, Ostelo RW, Knol DL, Vlaeyen JW, Bouter LM, de Vet HC. Is a behavioral graded activity program more effective than manual therapy in patients with subacute neck pain? Results of a randomized clinical trial. *Spine (Phila Pa 1976).* 2010;35:1017–1024.
11. Lilje S, Friberg H, Wykman A, Skillgate E. Naprapathic manual therapy or conventional orthopedic care for outpatients on orthopedic waiting lists?: A pragmatic randomized controlled trial. *Clin J Pain.* 2010;26:602–610.
12. Borusiak P, Biedermann H, Bosserhoff S, Opp J. Lack of efficacy of manual therapy in children and adolescents with suspected cervicogenic headache: results of a prospective, randomized, placebo-controlled, and blinded trial. *Headache.* 2010;50:224–230.
13. Cleland JA, Mintken PE, Carpenter K, Fritz JM, Glynn P, Whitman J, et al. Examination of a clinical prediction rule to identify patients with neck pain likely to benefit from thoracic spine thrust manipulation and a general cervical range of motion exercise: multi-center randomized clinical trial. *Phys Ther.* 2010;90:1239–1250.
14. Kumar D, Sandhu JS, Broota A. Efficacy of Mulligan concept (NAGs) on pain at available end range in cervical spine: a randomised controlled trial. *Indian Journal of Physiotherapy & Occupational Therapy.* 2011;5:154–158.
15. Bialoszewski D, Zaborowski G. Usefulness of manual therapy in the rehabilitation of patients with chronic rotator cuff injuries. Preliminary report. *Ortop Traumatol Rehabil.* 2011;13:9–20.
16. von Piekartz H, Ludtke K. Effect of treatment of temporomandibular disorders (TMD) in patients with cervicogenic headache: a single-blind, randomized controlled study. *Cranio.* 2011;29:43–56.
17. Villafane JH, Silva GB, Diaz-Parreno SA, Fernandez-Carnero J. Hypoalgesic and motor effects of kaltenborn mobilization on elderly patients with secondary thumb carpometacarpal osteoarthritis: a randomized controlled trial. *J Manipulative Physiol Ther.* 2011;34:547–556.

## Appendix 1. References for the systematic review

1. Gemmell H, Miller P. Relative effectiveness and adverse effects of cervical manipulation, mobilisation and the activator instrument in patients with sub-acute non-specific neck pain: results from a stopped randomised trial. *Chiropr Osteopat.* 2010;18:20.
2. Cuccia AM, Caradonna C, Annunziata V, Caradonna D. Osteopathic manual therapy versus conventional conservative therapy in the treatment of temporomandibular disorders: a randomized controlled trial. *J Bodyw Mov Ther.* 2010;14:179–184.
3. Sillevs R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. *J Man Manip Ther.* 2010;18:181–190.

18. de Camargo VM, Albuquerque-Sendin F, Berzin F, Stefanelli VC, de Souza DP, Fernandez-de-las-Penas C. Immediate effects on electromyographic activity and pressure pain thresholds after a cervical manipulation in mechanical neck pain: a randomized controlled trial. *J Manipulative Physiol Ther.* 2011;34:211-220.
19. Voigt K, Liebnitzky J, Burmeister U, Sihvonien-Riemenschneider H, Beck M, Voigt R, et al. Efficacy of osteopathic manipulative treatment of female patients with migraine: results of a randomized controlled trial. *J Altern Complement Med.* 2011;17:225-230.
20. Lau HM, Wing Chiu TT, Lam TH. The effectiveness of thoracic manipulation on patients with chronic mechanical neck pain - a randomized controlled trial. *Man Ther.* 2011;16:141-147.
21. Szlezak AM, Georgilopoulos P, Bullock-Saxton JE, Steele MC. The immediate effect of unilateral lumbar Z-joint mobilisation on posterior chain neurodynamics: a randomised controlled study. *Man Ther.* 2011;16:609-613.
22. Yeo HK, Wright A. Hypoalgesic effect of a passive accessory mobilisation technique in patients with lateral ankle pain. *Man Ther.* 2011;16:373-377.
23. Martel J, Dugas C, Dubois JD, Descarreaux M. A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain. *BMC Musculoskelet Disord.* 2011;12:41.
24. Puentedura EJ, Landers MR, Cleland JA, Mintken PE, Huijbregts P, Fernandez-de-Las-Penas C. Thoracic spine thrust manipulation versus cervical spine thrust manipulation in patients with acute neck pain: a randomized clinical trial. *J Orthop Sports Phys Ther.* 2011;41:208-220.
25. Petersen T, Larsen K, Nordsteen J, Olsen S, Fournier G, Jacobsen S. The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: a randomized controlled trial. *Spine (Phila Pa 1976).* 2011;36:1999-2010.
26. Bronfort G, Maiers MJ, Evans RL, Schulz CA, Bracha Y, Svendsen KH, et al. Supervised exercise, spinal manipulation, and home exercise for chronic low back pain: a randomized clinical trial. *Spine J.* 2011;11:585-598.
27. Castien RF, van der Windt DA, Grooten A, Dekker J. Effectiveness of manual therapy for chronic tension-type headache: a pragmatic, randomised, clinical trial. *Cephalalgia.* 2011;31:133-143.
28. Solanki D, Kage V. A Comparative Study on Immediate effect of Adductor Stretch MWM versus MET in Subjects with Hip Adductor Tightness - Randomized Clinical Trial. *Indian Journal of Physiotherapy and Occupational Therapy.* 2012;6:44-47.
29. Cruser d A, Maurer D, Hensel K, Brown SK, White K, Stoll ST. A randomized, controlled trial of osteopathic manipulative treatment for acute low back pain in active duty military personnel. *J Man Manip Ther.* 2012;20:5-15.
30. Nagrale AV, Patil SP, Gandhi RA, Learman K. Effect of slump stretching versus lumbar mobilization with exercise in subjects with non-radicular low back pain: a randomized clinical trial. *J Man Manip Ther.* 2012;20:35-42.
31. Schenk R, Dionne C, Simon C, Johnson R. Effectiveness of mechanical diagnosis and therapy in patients with back pain who meet a clinical prediction rule for spinal manipulation. *J Man Manip Ther.* 2012;20:43-49.
32. Pace do Amaral MT, Freire de Oliveira MM, Ferreira Nde O, Guimaraes RV, Sarian LO, Gurgel MS. Manual therapy associated with upper limb exercises vs. exercises alone for shoulder rehabilitation in postoperative breast cancer. *Physiother Theory Pract.* 2012;28:299-306.
33. Viswas R, Ramachandran R, Korde Anantkumar P. Comparison of effectiveness of supervised exercise program and Cyriax physiotherapy in patients with tennis elbow (lateral epicondylitis): a randomized clinical trial. *ScientificWorldJournal.* 2012;2012:939645.
34. Miller JE, Newell D, Bolton JE. Efficacy of chiropractic manual therapy on infant colic: a pragmatic single-blind, randomized controlled trial. *J Manipulative Physiol Ther.* 2012;35:600-607.
35. Puhl AA, Injeyan HS. Short-term effects of manipulation to the upper thoracic spine of asymptomatic subjects on plasma concentrations of epinephrine and norepinephrine—a randomized and controlled observational study. *J Manipulative Physiol Ther.* 2012;35:209-215.
36. Pentelka L, Hebron C, Shapleski R, Goldshtein I. The effect of increasing sets (within one treatment session) and different set durations (between treatment sessions) of lumbar spine posteroanterior mobilisations on pressure pain thresholds. *Man Ther.* 2012;17:526-530.
37. Yang JL, Jan MH, Chang CW, Lin JJ. Effectiveness of the end-range mobilization and scapular mobilization approach in a subgroup of subjects with frozen shoulder syndrome: a randomized control trial. *Man Ther.* 2012;17:47-52.
38. Balthazard P, de Goumoens P, Rivier G, Demeulenaere P, Ballabeni P, Deriaz O. Manual therapy followed by specific active exercises versus a placebo followed by specific active exercises on the improvement of functional disability in patients with chronic non specific low back pain: a randomized controlled trial. *BMC Musculoskelet Disord.* 2012;13:162.
39. Martinez-Segura R, De-la-Llave-Rincon AI, Ortega-Santiago R, Cleland JA, Fernandez-de-Las-Penas C. Immediate changes in widespread pressure pain sensitivity, neck pain, and cervical range of motion after cervical or thoracic thrust manipulation in patients with bilateral chronic mechanical neck pain: a randomized clinical trial. *J Orthop Sports Phys Ther.* 2012;42:806-814.
40. Dunning JR, Cleland JA, Waldrop MA, Arnot CF, Young IA, Turner M, et al. Upper cervical and upper thoracic thrust manipulation versus nonthrust mobilization in patients with mechanical neck pain: a multicenter randomized clinical trial. *J Orthop Sports Phys Ther.* 2012;42:5-18.
41. Saavedra-Hernandez M, Castro-Sanchez AM, Arroyo-Morales M, Cleland JA, Lara-Palomo IC, Fernandez-de-Las-Penas C. Short-term effects of kinesio taping versus cervical thrust manipulation in patients with mechanical neck pain: a randomized clinical trial. *J Orthop Sports Phys Ther.* 2012;42:724-730.
42. Brantingham JW, Parkin-Smith G, Cassa TK, Globe GA, Globe D, Pollard H, et al. Full kinetic chain manual and manipulative therapy plus exercise compared with targeted manual and manipulative therapy plus exercise for symptomatic osteoarthritis of the hip: a randomized controlled trial. *Arch Phys Med Rehabil.* 2012;93:259-267.
43. Nee RJ, Vicenzino B, Jull GA, Cleland JA, Coppieters MW. Neural tissue management provides immediate clinically relevant benefits without harmful effects for patients with nerve-related neck and arm pain: a randomised trial. *J Physiother.* 2012;58:23-31.
44. Bronfort G, Evans R, Anderson AV, Svendsen KH, Bracha Y, Grimm RH. Spinal manipulation, medication, or home exercise with advice for acute and subacute neck pain: a randomized trial. *Ann Intern Med.* 2012;156:1-10.

45. Sharma SS. Randomized Comparison of effectiveness of Clinical Exercises and Manual Therapy Procedures versus Clinical Exercises alone in the Treatment of Osteoarthritis of Knee. *Indian Journal of Physiotherapy & Occupational Therapy*. 2013;7:188-204.
46. Tuncer AB, Ergun N, Tuncer AH, Karahan S. Effectiveness of manual therapy and home physical therapy in patients with temporomandibular disorders: A randomized controlled trial. *J Bodyw Mov Ther*. 2013;17:302-308.
47. Lee J, Lee Y, Kim H, Lee J. The Effects of Cervical Mobilization Combined with Thoracic Mobilization on Forward Head Posture of Neck Pain Patients. *J Phys Ther Sci*. 2013;25:7-9.
48. Cook C, Learman K, Showalter C, Kabbaz V, O'Halloran B. Early use of thrust manipulation versus non-thrust manipulation: a randomized clinical trial. *Man Ther*. 2013;18:191-198.
49. Lin JH, Shen T, Chung RC, Chiu TT. The effectiveness of Long's manipulation on patients with chronic mechanical neck pain: a randomized controlled trial. *Man Ther*. 2013;18:308-315.
50. Saavedra-Hernandez M, Arroyo-Morales M, Cantarero-Villanueva I, Fernandez-Lao C, Castro-Sanchez AM, Puentedura EJ, et al. Short-term effects of spinal thrust joint manipulation in patients with chronic neck pain: a randomized clinical trial. *Clin Rehabil*. 2013;27:504-512.
51. Cleland JA, Mintken PE, McDevitt A, Bieniek ML, Carpenter KJ, Kulp K, et al. Manual physical therapy and exercise versus supervised home exercise in the management of patients with inversion ankle sprain: a multicenter randomized clinical trial. *J Orthop Sports Phys Ther*. 2013;43:443-455.
52. Villafane JH, Cleland JA, Fernandez-de-Las-Penas C. The effectiveness of a manual therapy and exercise protocol in patients with thumb carpometacarpal osteoarthritis: a randomized controlled trial. *J Orthop Sports Phys Ther*. 2013;43:204-213.
53. Masaracchio M, Cleland JA, Hellman M, Hagins M. Short-term combined effects of thoracic spine thrust manipulation and cervical spine nonthrust manipulation in individuals with mechanical neck pain: a randomized clinical trial. *J Orthop Sports Phys Ther*. 2013;43:118-127.
54. French HP, Cusack T, Brennan A, Caffrey A, Conroy R, Cuddy V, et al. Exercise and manual physiotherapy arthritis research trial (EMPART) for osteoarthritis of the hip: a multicenter randomized controlled trial. *Arch Phys Med Rehabil*. 2013;94:302-314.
55. von Heymann WJ, Schloemer P, Timm J, Muehlbauer B. Spinal high-velocity low amplitude manipulation in acute nonspecific low back pain: a double-blinded randomized controlled trial in comparison with diclofenac and placebo. *Spine (Phila Pa 1976)*. 2013;38:540-548.
56. Visser LH, Woudenberg NP, de Bont J, van Eijs F, Verwer K, Jenniskens H, et al. Treatment of the sacroiliac joint in patients with leg pain: a randomized-controlled trial. *Eur Spine J*. 2013;22:2310-2317.
57. Castro-Sanchez AM, Aguilar-Ferrandiz ME, Mataran-Penarrocha GA, Sanchez-Joya Mdel M, Arroyo-Morales M, Fernandez-de-las-Penas C. Short-term effects of a manual therapy protocol on pain, physical function, quality of sleep, depressive symptoms, and pressure sensitivity in women and men with fibromyalgia syndrome: a randomized controlled trial. *Clin J Pain*. 2014;30:589-597.
58. La Touche R, Paris-Alemayn A, Mannheimer JS, Angulo-Diaz-Parreno S, Bishop MD, Lopez-Valverde-Centeno A, et al. Does mobilization of the upper cervical spine affect pain sensitivity and autonomic nervous system function in patients with cervico-craniofacial pain?: A randomized-controlled trial. *Clin J Pain*. 2013;29:205-215.
59. de Oliveira RF, Liebano RE, Costa Lda C, Rissato LL, Costa LO. Immediate effects of region-specific and non-region-specific spinal manipulative therapy in patients with chronic low back pain: a randomized controlled trial. *Phys Ther*. 2013;93:748-756.
60. Licciardone JC, Minotti DE, Gatchel RJ, Kearns CM, Singh KP. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomized controlled trial. *Ann Fam Med*. 2013;11:122-129.
61. Poulsen E, Hartvigsen J, Christensen HW, Roos EM, Vach W, Overgaard S. Patient education with or without manual therapy compared to a control group in patients with osteoarthritis of the hip. A proof-of-principle three-arm parallel group randomized clinical trial. *Osteoarthritis Cartilage*. 2013;21:1494-1503.
62. Abbott JH, Robertson MC, Chapple C, Pinto D, Wright AA, Leon de la Barra S, et al. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee: a randomized controlled trial. 1: clinical effectiveness. *Osteoarthritis Cartilage*. 2013;21:525-534.
63. Espi-Lopez GV, Gomez-Conesa A. Efficacy of manual and manipulative therapy in the perception of pain and cervical motion in patients with tension-type headache: a randomized, controlled clinical trial. *J Chiropr Med*. 2014;13:4-13.
64. Marron-Gomez D, Rodriguez-Fernandez AL, Martin-Urrialde JA. The effect of two mobilization techniques on dorsiflexion in people with chronic ankle instability. *Phys Ther Sport*. 2015;16:10-15.
65. Casanova-Mendez A, Oliva-Pascual-Vaca A, Rodriguez-Blanco C, Heredia-Rizo AM, Gogorza-Arroitaonandia K, Almazan-Campos G. Comparative short-term effects of two thoracic spinal manipulation techniques in subjects with chronic mechanical neck pain: a randomized controlled trial. *Man Ther*. 2014;19:331-337.
66. Izquierdo Perez H, Alonso Perez JL, Gil Martinez A, La Touche R, Lerma-Lara S, Commeaux Gonzalez N, et al. Is one better than another?: A randomized clinical trial of manual therapy for patients with chronic neck pain. *Man Ther*. 2014;19:215-221.
67. Paanalahti K, Holm LW, Nordin M, Asker M, Lyander J, Skillgate E. Adverse events after manual therapy among patients seeking care for neck and/or back pain: a randomized controlled trial. *BMC Musculoskelet Disord*. 2014;15:77.
68. Espi-Lopez GV, Rodriguez-Blanco C, Oliva-Pascual-Vaca A, Benitez-Martinez JC, Lluch E, Falla D. Effect of manual therapy techniques on headache disability in patients with tension-type headache. Randomized controlled trial. *Eur J Phys Rehabil Med*. 2014;50:641-647.
69. Ofner M, Kastner A, Wallenboeck E, Pehn R, Schneider F, Groell R, et al. Manual khalifa therapy improves functional and morphological outcome of patients with anterior cruciate ligament rupture in the knee: a randomized controlled trial. *Evid Based Complement Alternat Med*. 2014;2014:462840.
70. Rabin A, Shashua A, Pizem K, Dickstein R, Dar G. A clinical prediction rule to identify patients with low back pain who are likely to experience short-term success following lumbar stabilization exercises: a randomized controlled validation study. *J Orthop Sports Phys Ther*. 2014;44:6-B13.
71. Packer AC, Pires PF, Dibai-Filho AV, Rodrigues-Bigaton D. Effects of upper thoracic manipulation on pressure pain sensitivity in women with temporomandibular disorder: a randomized, double-blind, clinical trial. *Am J Phys Med Rehabil*. 2014;93:160-168.

72. Reid SA, Rivett DA, Katekar MG, Callister R. Comparison of mulligan sustained natural apophyseal glides and maitland mobilizations for treatment of cervicogenic dizziness: a randomized controlled trial. *Phys Ther.* 2014;94:466-476.