Assessing the effect of sample size, methodological quality and statistical rigour on outcomes of randomised controlled trials on mobilisation, manipulation and massage for low back pain of at least 6 weeks duration

Hettinga DM, Hurley DA, Jackson A, May S, Mercer C, Roberts L

CRD summary
The review concluded that many trials of manual therapy for lower back pain had shortcomings, but there remained evidence from higher quality trials to support the use of a manual therapy package compared with General Practitioner care for non-specific lower back pain of at least six weeks duration. The authors’ conclusions are suitably cautious and consistent with the available evidence.

Authors’ objectives
To assess the effect of sample size, methodological quality, and statistical rigour on outcomes of randomised controlled trials (RCTs) on manual therapy for non-specific lower back pain of at least six weeks in duration, and to report the results from RCTs with adequate sample size, quality and statistical rigour.

Searching
MEDLINE, EMBASE, CINAHL, AMED, PEDro, the Cochrane Library and the library collection of the Chartered Society of Physiotherapy were searched to June 2005. Reference lists of identified systematic reviews were also searched.

Study selection
Randomised controlled trials (RCTs) that compared massage, mobilisation and/or manipulation as a single intervention versus control or an alternative intervention, in adult patients (aged over 18 years) with non-specific lower back pain (of at least six weeks duration) were eligible for inclusion. Trials were also included if they combined manual therapies with other non-physiotherapy interventions.

Relevant outcomes included pain, function, physiological status, or return to work/sick leave.

The included trials evaluated the following interventions: manual therapy; massage; mobilisation in combination with manipulation; General Practitioner (GP) care plus/minus exercise; or manipulation. These interventions were compared with Chinese acupuncture, self-care, untreated controls, relaxation therapy, education programme, medication, diathermy, placebo, or sham manipulation. Some trials compared different durations of manual therapy. The reported outcomes included pain, function, disability outcomes, psychological status, and depression index.

The authors did not state how many reviewers performed study selection.

Assessment of study quality
Two reviewers independently assessed trial quality using an adapted version of the van Tulder criteria, assessing 10 methodological quality factors, including allocation concealment, blinding, similarity of groups at baseline, co-interventions, compliance rates, withdrawals and drop-outs, timing of outcome assessments, and analysis by intention-to-treat. Sample size was also considered, in terms of the number of participants in the manual therapy arm of the trial, as was statistical rigour. For trials to be deemed high quality they had to have more than 40 participants in the manual therapy arm, score above 5 out of 10 on the modified van Tulder scale, and report statistical tests that compared the change in the intervention group with the change in the control group.

Data extraction
Data was extracted on the change in pain and function after any follow-up period and used to calculate the percent improvement (or worsening).

The authors did not state how many reviewers performed data extraction.
Methods of synthesis
The percentage change in pain and function were plotted on L’Abbe plots to identify outliers, and the effects of quality on results. The difference in change for both pain and function outcomes was subsequently plotted against quality scores and manual therapy sample size to determine the effects of quality and sample size on outcomes. A narrative synthesis was also presented looking at the results of the high quality RCTs only.

Results of the review
Ten RCTs (n=2,499 participants) were included in the review. The trial sample size ranged from 24 to 1,334 participants. The methodological quality assessment ranged from 1 to 7 out of 10 points, with only three trials scoring more than 5. Half of the trials did not have statistical tests to compare changes; half of the trials had a sample size in the manual therapy group less than 40 participants. Only two trials fulfilled all of the authors’ criteria and were deemed high quality.

Quantitative analysis (all trials that compared intervention with control): Three trials showed statistically significant improvements in function, and two trials showed statistically significant changes in pain. Low quality RCTs and RCTs with small manual therapy sample sizes showed a trend for a greater variation in the difference in change than larger and higher quality RCTs.

Qualitative analysis (higher quality trials only): One trial that fulfilled the quality, sample size and statistical rigour criteria demonstrated that compared with GP care, manipulation in combination with mobilisation was more effective for pain relief and functional improvements, and equally effective for fear-avoidance behaviour. Another trial that also fulfilled the quality criteria showed that mobilisation was as effective as sham manipulation in terms of pain, function and psychological status.

Authors’ conclusions
Many RCTs in the area of manual therapy for lower back pain had shortcomings in sample size, methodological quality, and/or statistical rigour. However, there remained evidence from higher quality RCTs to support the use of a manual therapy package, compared with GP care, for non-specific lower back pain of at least six weeks duration.

CRD commentary
Inclusion criteria for the review were clearly defined. Several relevant databases were searched; search terms were not provided. Publication bias was not assessed and could not be ruled out. The potential for language bias was unclear. Suitable methods were used to minimise the risk of reviewer bias during quality assessment, but the authors did not report on whether such methods were used for study selection and data extraction.

Trial quality assessment was undertaken according to a modified version of Van Tulder criteria; an assessment of the trial size and statistical rigour was also conducted, which gave a good indication of trial quality. However, there was a lack of details on the trial participants which may affect generalisability. The trial outcomes were assessed using L’Abbe plots. Other graphical techniques were used to determine the effects of quality and trial size on outcomes, which appeared appropriate.

Overall, the authors’ conclusions were suitably cautious and consistent with the review aims and available evidence.

Implications of the review for practice and research
Practice: The authors stated that at this point in time, manipulation used in isolation cannot be recommended. Likewise no recommendations can be made for massage interventions due to the lack of RCTs fulfilling all criteria for best available evidence.

Research: The authors stated that future RCTs to inform clinical practice should have an adequate sample size, methodological quality and statistical rigour. Future trials should use relevant and valid outcome measures and identify subgroups where specific treatments may be most effective.
Funding
Chartered Society of Physiotherapists (CSP); CSP Charitable Trust, UK.

Bibliographic details

DOI
10.1016/j.physio.2007.10.008

Original Paper URL
http://www.physiotherapyjournal.com/article/S0031-9406(07)00177-0/abstract

Indexing Status
Subject indexing assigned by CRD

MeSH
Humans; Low Back Pain; Odds Ratio; Pain Measurement; Randomized Controlled Trials as Topic

AccessionNumber
12008107181

Date bibliographic record published
07/04/2009

Date abstract record published
22/12/2010

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.