A systematic review and meta-analysis of clinical trials on physical interventions for lateral epicondylalgia

L Bisset, A Paungmali, B Vicenzino, E Beller

A systematic review of the literature on the effectiveness of physical interventions for lateral epicondylalgia (tennis elbow) was carried out. Seventy-six randomised controlled trials were identified, 28 of which satisfied the minimum criteria for meta-analysis. The evidence suggests that extracorporeal shock wave therapy is not beneficial in the treatment of tennis elbow. There is a lack of evidence for the long-term benefit of physical interventions in general. However, further research with long term follow up into manipulation and exercise as treatments is indicated.

Aetiological factors of tennis elbow are predominantly repetitive hand tasks, especially those in the food processing, retail, and transportation industries, where the workplace often involves lifting, handling heavy loads, or carrying loads on the shoulders. An estimated 10% of the population experience some type of elbow pain, and 75–80% of these elbow complaints are attributed to repetitive hand tasks.10 Up to 50% of all tennis players experience some type of elbow pain, and 75–80% of these elbow complaints are attributed to tennis elbow.11 Many treatments have been advocated in the management of LE, possibly implying that much is unknown about its aetiology and how best it should be treated.

Labelle et al11 attempted to perform a quantitative meta-analysis of various treatments for LE, using a search of the Embase and Medline databases from 1966 to 1990 and including French and English studies. They concluded that there was insufficient evidence to support the efficacy or lack thereof of any one type of intervention. The most current systematic review of physical treatments for LE by Smidt et al10 covered Medline, Embase, Cinahl, Cochrane Controlled Trial Register, and Current Contents databases and performed citation tracking up to January 1999. They concluded that there was still insufficient evidence to draw conclusions on the effects of all therapeutic approaches, except ultrasound (US), for which they found weak evidence to support its effectiveness over a placebo.

The Cochrane Library contains a series of systematic reviews for the treatment of LE, covering such interventions as extracorporeal shock wave therapy (ESWT; review last amended 2001), acupuncture (databases searched to June 2002) within the Cochrane Library has looked at the effectiveness of transverse friction massage on general tendonopathy conditions.12 The prevailing conclusion drawn by many of these reviews was that of insufficient evidence to support any one treatment over another, predominantly because of the low statistical power, inadequate internal validity, and insufficient data reported in most of the trials.

Before the systematic review of Smidt et al,10 there was no review of the literature pertaining to the effectiveness of manual therapy and exercises for LE, commonly used treatments for musculoskeletal pain in clinical practice. Within the field of manual therapy for LE, there are several randomised controlled trials (RCTs) that specifically evaluated Cyriax manipulation, transverse friction massage, and cervical mobilisations.13–16 Following on from the review by Smidt et al,10 there are several recently published studies on the effect of manipulation techniques applied to the elbow, wrist, and cervical spine, on outcomes for LE.17–21 Similarly, studies on the use of ESWT as a treatment for LE have increased in the literature.22–26

An updated systematic review is justified on the basis of further evidence coming to light since the review of Smidt et al,10 who reviewed data before 2000. The aim of this review was to look at the effectiveness of physical interventions on clinically relevant outcomes for LE to discern if there are any advances on current best practice guidelines.

METHODS

Search strategy

One reviewer searched the full set of the following databases: Medline, Cinahl, Embase, Web of Science, Allied and Complimentary Medicine, SPORTDiscus, and Physiotherapy Evidence Database (PEDro), without language restrictions, using the recommended Cochrane Library search strategy.26 The Cochrane Controlled Trial Register was searched for RCTs on LE/tennis elbow, and references from retrieved articles and systematic reviews were also screened.

Abbreviations: CI, confidence interval; ESWT, extracorporeal shock wave therapy; LE, lateral epicondylalgia; MGS, maximum grip strength; MWM, mobilisations with movement; NSAID, non-steroidal anti-inflammatory drug; PF, pain-free grip strength; PVAS, pain visual analogue scale; RCT, randomised controlled trial; RR, relative risk; SMD, standardised mean difference; US, ultrasound
Selection
For this systematic review we included all studies that met the following conditions: participants with a diagnosis of tennis elbow or LE, as confirmed by lateral elbow pain that increased on palpation and/or during resisted wrist dorsiflexion, and at least one of the interventions included a relevant physical intervention. A physical intervention was defined as any intervention that was physical in nature—that is, not solely pharmaceutical or surgical. By definition, the studies also required randomisation of participants, comparison between at least two groups with one intervention being the physical intervention, and the use of at least one clinically relevant outcome measure, such as pain, grip strength, or global improvement. All databases were searched during September 2003, and those comparing a physical intervention with corticosteroid injections or NSAIDs were included, but those involving surgery were not. Nor were studies that used the same physical intervention in both groups.

Abstracts of papers identified by the search strategy were screened for inclusion by one reviewer (LB). If there was any indecision, the full text was retrieved and reviewed by three reviewers (LB, AP, BV). A full copy of all identified articles was then retrieved. Articles were not restricted by language, and a translator was used when necessary.

Quality assessment
The modified PEDro rating scale (appendix 1) used to assess methodological quality was derived from an 11 point scale initially developed to rate the quality of RCTs on the Physiotherapy Evidence Database (www.pedro.fhs.usyd.edu.au). This was adapted from the list developed by Verhagen et al using the Delphi consensus technique. An additional four criteria, deemed important in other rating scales, were also included. Two qualified PEDro raters (LB and AP), independent and blind to each other, rated all papers. If there was an initial disagreement on any criterion, it was reassessed by each reviewer independently. Unresolved disagreements were identified and discussed in a consensus meeting (LB and AP). Any disagreement that was unresolved by general consensus was taken to a third reviewer (BV), who was independent to the initial deliberations, and a final consensus was reached.

Each criterion was accompanied by a strict descriptive list and was rated either yes or no to minimise ambiguity. Each criterion was then given a score: yes = 1 point, no = 0 point. A total score for the methodological quality of each study was calculated by summing up the number of positive criteria (maximum score of 15), and papers rating at least 50% (greater than or equal to 8/15) were then deemed of a suitable methodological quality to be included in the review. It has been previously shown that the inclusion of poorer quality studies in a meta-analysis is associated with an increased estimate of benefit and therefore may alter the interpretation of the effect of an intervention.

Inter-rater reliability
The inter-rater reliability was evaluated by reporting the percentage of initial agreement and the κ statistics, both for overall agreement and each criterion.

Data management and statistical analysis
Data synthesis was facilitated using RevMan 4.2. The summary statistics were the standardised mean difference (SMD; 95% confidence interval (CI)) and relative risk (RR; 95% CI) on a random effects model. A positive SMD represents an effect in favour of the intervention, with values greater than 0.8 considered a large clinical effect, 0.5 a moderate effect, and 0.2 a weak effect. To rate clinical significance for the RR, we followed Smidt et al and set it at 0.7 as favouring the placebo/control group and 1.5 as favouring the intervention. Where possible, the mean change scores and standard deviations of the change scores were used to calculate the SMD. If this was not possible, and provided that the baseline scores were not significantly different, the SMD and CI were calculated from the post-intervention mean (SD) scores. If insufficient data were available from the article itself, a formal written request was made to the author in an attempt to obtain the relevant data. All data entry and conversion were performed by one investigator (LB) and then checked by another (BV).

Inter-rater reliability
Out of a total of 1140 criteria rated, there were 99 (8.7%) initial disagreements between raters, with a κ statistic for inter-rater agreement of 0.824. After a consensus meeting between the two reviewers (LB and AP), only nine (<1%) decisions could not be resolved, and the third reviewer (BV) was called upon to make the final decision. Inter-rater agreements were moderate to good (0.598–0.936) for all criteria except for criterion 6 (similarity of baseline measures), which had a weak to moderate agreement with a κ of 0.485.

RESULTS
Figure 1 illustrates the systematic review process of identifying, refining, and final selection of RCTs for the study. After rating by reviewers, only 28 studies (table 1) scored above the a priori minimum quality score of 50%—that is, eight out of 15 criteria. This section reports on these 28 papers.

Outcome measures
Pain scores were reported in most studies using either a continuous visual analogue scale (PVAS) or an ordinal points system. Grip strength was reported in most trials as either maximum grip strength (MGS; 11 studies, 38%) or pain-free grip strength (PFGS; 14 studies, 50%), and 15 (52%) studies reported the dichotomous rating of success through a global improvement or patient satisfaction scale. Table 2 lists the range of outcome measures across the included studies.

Twentyfour of the 28 studies—that is, all except those in refs—performed a short term outcome assessment (less than six weeks), and only eight studies included a long term follow up (more than six months). Five studies only looked at the effects immediately after the intervention.

NON-ELECTROThERAPEUTIC INTERVENTIONS
Twelve studies assessed a range of non-electrotherapeutic interventions: four examined acupuncture, four manipulated mobilisation treatments, and there was one tape study, one exercise programme, and two brace studies.

Exercise
Only one study that satisfied the quality criteria specifically evaluated an exercise programme. Pienimaki et al compared a six to eight week exercise programme of stretches and exercises (isometric and isotonic) with a treatment of pulsed US across the same time span, and showed that the SMD for PVAS at rest was 0.97 (95% CI 0.30 to 1.63) and 0.66 (95% CI 0.01 to 1.31) for PVAS under strain. MGS was not significantly different between groups (table 3). This suggests a favourable effect in that exercise may improve pain in LE but not MGS.

All other studies that included exercise in their treatment group did so as a co-intervention applied concomitantly with other modalities. Their results cannot be attributed solely to exercise and cannot be compared.
Manipulation techniques

Three studies of adequate quality (rated eight, 11, and 13 out of 15 criteria respectively)\(^{19} \text{ to } 23\) looked at the immediate effects of a cervical or elbow manipulation on outcome measures. One study looked at a six week treatment programme of wrist manipulation compared with a placebo group. Data abstraction results show that at the end of the six week treatment period, PVAS during the day, PFGS, MGS, and pressure pain threshold were not significantly different between groups (table 3). In summary, there are no long term studies of adequate methodological quality on manual therapy. However, there appears to be some evidence of positive initial effects in favour of elbow manipulative therapy techniques that suggests the need for further research.

Orthotics and taping

Of the nine studies assessed for methodological quality that included an orthotic as an intervention or co-intervention, only two papers were of adequate quality for this review.\(^{47} \text{ to } 51\)

Jensen \textit{et al}\(^{51}\) compared an off the shelf orthotic (Rehband) with corticosteroid injections over an initial period of six weeks. No significant differences were found between the two groups for MGS, PVAS, dumbbell test, function VAS, or global improvement (table 3). Both groups had significantly improved from baseline. Jensen \textit{et al}\(^{47}\) concluded that the use of an orthotic was as effective as a steroid injection in the early management of LE, with respect to PVAS and patients’ self rating of their condition and suggested that orthotics should be the treatment of choice, as it is the simpler treatment with the fewest adverse side effects.

Wuori \textit{et al}\(^{51}\) compared an off the shelf orthotic (Count’R-Force Tennis Elbow brace; Medical Sports, Inc, Arlington, Virginia, USA) with two different types of placebo braces (a neoprene elbow sleeve (The Body Glove Airprene Elbow Support; Body Glove International, Redondo Beach, California, USA) and a modified DePuy patella strap (DePuy Inc, Warsaw, Indiana, USA)), as well as a control group. This was a repeated measures design study, with subjects acting as their own controls, and it looked at the immediate effects on PFGS and PVAS. No significant differences were found for PFGS and PVAS between the Count’R-Force tennis elbow brace and the placebo DePuy brace nor between the Count’R-Force brace and the Airprene Elbow Support (table 3).

One other study which also rated well on methodological quality assessed the immediate effects of a specific taping technique on outcome measures of PFGS and pressure pain threshold, over both a placebo and control group. Vicenzino \textit{et al}\(^{52}\) looked at the immediate effects of a diamond shaped taping technique and found a significant improvement in the intervention group for both PFGS (24% improvement from baseline) and pressure pain threshold immediately and 30 minutes after the intervention, when compared with a control. However, when compared with a placebo, the only significant effect seen was for pressure pain threshold immediately after the intervention (table 3). No firm conclusions on orthotics or tape can be confidently drawn from the outcomes of only three studies that have different timelines for measurements and different comparison groups. Further research is required before any firm conclusions can be drawn.

Acupuncture

Four out of 10 papers rated on acupuncture were of adequate methodological quality to be included in this review.\(^{68} \text{ to } 75\) Three studies compared acupuncture with placebo.\(^{68} \text{ to } 75\) The remaining paper compared acupuncture with US.\(^{74}\) Owing to
differences in the applied interventions and timing of outcome assessment, no pooling of studies was possible.

All four acupuncture studies used a blinded assessor, and Molsberger and Hille\(^5\) and Fink et al\(^4\) also blinded their subjects. None of the three studies used concealed allocation, nor was intention to treat analysis reported. Molsberger and Hille\(^5\) found that acupuncture resulted in relief of pain for a statistically significant longer time than placebo (SMD 1.20; 95% CI 0.58 to 1.82) and was significantly more likely to result in overall success (RR 3.17; 95% CI 1.54 to 6.52) after 10 treatments. Fink et al\(^5\) found significant improvement in PVAS early after treatment in favour of acupuncture, but no clinically significant difference at two months (table 3). Similarly, the disabilities of the arm, shoulder, and hand questionnaire (DASH) was significantly different in favour of acupuncture over placebo at two weeks (SMD 0.77; 95% CI 0.14 to 1.40), but not so by two months (SMD 0.48; 95% CI −0.15 to 1.11). Fink et al\(^5\) also showed a significant difference in favour of the acupuncture group after treatment (PVAS SMD 0.86; 95% CI 0.23 to 1.50) that was no longer significant by two months (PVAS SMD 0.31; 95% CI −0.31 to 0.93).

Davidson et al\(^4\) found less of a difference when comparing acupuncture with US. Their study showed a significant improvement in all outcome measures over time in both groups, but PFGS and PVAS were not significantly different between the two groups (table 3). Their measure of the DASH was not different between the two groups at the end of treatment (SMD 0.42; 95% CI −0.57 to 1.42), nor at one month (SMD −0.07; 95% CI −1.05 to 0.91).

---

Each criterion met is marked in grey. See appendix 1 for definitions of criteria.
There appears to be some evidence to support the efficacy of acupuncture over a placebo as a treatment for LE in short term outcomes. However, this benefit appears to be short-lived—that is, two to eight weeks. There also appears to be no difference between US and acupuncture in outcomes for LE, although no firm inferences can be made from only one study.

**ELECTROTHERAPEUTIC INTERVENTIONS**

Forty-one studies that were initially rated investigated an electrotherapeutic modality, and of those 17 were found of be of sufficient methodological quality to be included in this review. Of those included, five studies examined the effectiveness of US treatment, six examined laser therapy, two ESWT, and four looked at ionisation or electromagnetic field. Despite the presence of some contradiction in the short term effects of laser treatment, the pooled data presented show no evidence of effect over that of a placebo intervention, in either the short or long term.

**ESWT**

There has been growth in the number of studies of ESWT over recent years, but with very few showing adequate methodological quality. Of a total of eight studies, only two met the level of quality required to be included in this analysis.

One such well controlled study was by Haake et al. This trial adhered to good research design by using a placebo, blinding patients and assessors to outcomes, reporting side effects, concealing group allocation, and analysing data on an intention to treat basis. Fourteen (10.4%) patients in the active group and 16 (11.7%) in the placebo group were lost to follow up at the primary end point of one month after the end of treatment. Haake et al gave three treatments at weekly intervals and showed no differences between treatment groups for any of the measured end points. Significantly more side effects were documented in the ESWT group (99 participants with side effects/134 participants) than in the placebo group (38/136) (RR 2.64; 95% CI 1.09; 95% CI 0.77 to 1.53).

**Table 2** Outcome measures in 28 randomised control trials of physical treatments for lateral epicondylalgia

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain VAS Total</td>
<td>25</td>
</tr>
<tr>
<td>No mention of specifics</td>
<td>11</td>
</tr>
<tr>
<td>Wrist extension</td>
<td>8</td>
</tr>
<tr>
<td>24 hours</td>
<td>6</td>
</tr>
<tr>
<td>Palpation</td>
<td>6</td>
</tr>
<tr>
<td>Rest</td>
<td>5</td>
</tr>
<tr>
<td>Finger extension</td>
<td>4</td>
</tr>
<tr>
<td>Grip</td>
<td>2</td>
</tr>
<tr>
<td>Isometric supination/pronation</td>
<td>3</td>
</tr>
<tr>
<td>Stretch</td>
<td>3</td>
</tr>
<tr>
<td>Movement</td>
<td>2</td>
</tr>
<tr>
<td>Pinch</td>
<td>1</td>
</tr>
<tr>
<td>ADLs</td>
<td>1</td>
</tr>
</tbody>
</table>

**Pain and function rating scales**

- Patient assessment of subjective improvement* 13
- Functional VAS 4
- Patient self assessment pain scale 3
- Assessor’s assessment of functional impairment 4
- Pain/disability questionnaire* 2
- DASH 2
- Patient assessment of duration of symptoms 2
- Pain-free index 2
- 8 Item ADL score 1
- Roles and Maudley scale* 1
- Assessor’s assessment of pain VAS 1

**Physical evaluation**

- Pain-free grip strength 14
- Maximum grip strength 11
- Weight test 8
- Pressure pain threshold 6
- Range of movement in wrist 2
- Pinch strength 1
- Isometric wrist extension 1
- Isokinetic wrist flexion/extension 1
- Manual muscle test (0–5) 1
- Sympathetic nervous system indicators 1
- Thermal pain threshold 1
- Thermography 1
- Upper limb tension test 2b 1

Note that several papers used more than one measure of pain.
* Measures used to calculate global improvement (relative risk) in tables 3–5.

VAS, Visual analogue scale; ADL, activities of daily living; DASH, disabilities of the arm, shoulder, and hand questionnaire.

**Figure 2** Pooled measures of global improvement for laser treatment, extracorporeal shock wave therapy (ESWT), iontophoresis, and ultrasound (US) (black diamonds) and grip strength for mobilisations with movement (MWM) and laser treatment (white squares). Timing of measures: 1 = 2–6 weeks; 2 = 6–12 weeks; 3 = 26 weeks after the intervention.
1.98 to 3.53). The most reported side effect in the ESWT treated group was transitory reddening of the skin (42/134; 31.3%), pain (15/134; 11.2%), and small haematomas (10/134; 0.7%). For the placebo group, the most reported side effect was a placebo effect.

Another recent double blind, placebo controlled study that also used intention to treat analysis, looked at ESWT of 1500 pulses without local anaesthetic, with treatment applied three times at monthly intervals. At the primary end point of three months from baseline (one month after completion of treatment), there was no significant difference between the two groups for changes in pain scores. There was a worsening of symptoms reported as a side effect in two of the three studies looking at ionisation (table 3).

In summary, only one paper reported long term outcomes of ESWT over that of placebo effect.

Within the ionisation studies, data presented by Nirschl et al illustrated a reduction in pain in the short term for the intervention group (corticosteroid solution) that was not significantly different from the placebo group (change in PVAS at two days follow up, table 4). Similarly there was no difference between groups at one month (change in PVAS SMD 0.12; 95% CI −0.18 to 0.41) or in the patients’ self rated global improvement score between groups (table 4).

In comparing ionisation with two dose levels of a NSAID (pirprofen) with sham ionisation, Grossi et al found by the end of the treatment period (10 treatment sessions over two weeks) that PVAS was significantly lower in both the high and low dose NSAID groups than in the sham group. They also reported that significantly more patients rated their functional impairment as “absent” or “mild” on a functional impairment rating scale at the end of the treatment period in the NSAID ionisation groups (high dose: RR1.55; 95% CI 1.09 to 2.19; low dose: RR1.03; 95% CI 0.83 to 1.28). In table 4.

Electromagnetic field and ionisation

Four of nine studies evaluating the effectiveness of other electrotherapeutic modalities (electromagnetic field therapy and ionisation) were of adequate methodological quality, with one study looking at electromagnetic field versus a placebo, and the other three studies looking at ionisation. The study of Devereaux et al provided insufficient data on pain or grip strength to calculate an effect size. Eight patients in each group (53%) showed full recovery after eight weeks of treatment, the authors concluding that there was no advantage of pulsed electromagnetic field therapy over that of placebo effect.
refute the use of this treatment in the long term management of LE. Data abstraction from two papers showed no evidence of effect of corticosteroid iontophoresis over saline iontophoresis in the short term. There is contradictory evidence that NSAID ionisation may be more beneficial than either saline or sham ionisation immediately after treatment. Moreover, this contradiction in results and the heterogeneity of interventions makes it difficult to draw conclusions from the literature for this modality.

**US and phonophoresis**

Five out of 13 studies rated were of adequate methodological quality to be included in this review. Two studies compared US with a placebo,

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Duration of LE (months)</th>
<th>Sample size (N)</th>
<th>INT (n)</th>
<th>Week</th>
<th>Global improvement (RR) Pain (SMD) MGS (SMD) PFGS (SMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;57&lt;/sup&gt; (Nd:YAG) 204 mW/cm&lt;sup&gt;2&lt;/sup&gt;</td>
<td>6.4</td>
<td>47</td>
<td>12</td>
<td>4</td>
<td>0.37 (0.21 to 0.94)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;55&lt;/sup&gt; (GaAs 30 mW/830 nm)</td>
<td>36</td>
<td>8</td>
<td>4</td>
<td>1.10 (0.63 to 1.91)</td>
<td>0.08 (0.58 to 0.73)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;56&lt;/sup&gt; (HeNe 632.8 nm, 1.56 mW)</td>
<td>57</td>
<td>8</td>
<td>6</td>
<td>1.08 (0.49 to 2.40)</td>
<td>0.08 (0.62 to 0.69)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;57&lt;/sup&gt; (GaAs 904 nm, 0.07 mW)</td>
<td>3.5</td>
<td>30</td>
<td>8</td>
<td>3</td>
<td>7.00 (0.39 to 124.83)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;56&lt;/sup&gt; (GaAs, 904 nm, 1.5 mW, 3.5/3 cm&lt;sup&gt;2&lt;/sup &gt;)</td>
<td>5.5</td>
<td>58</td>
<td>10</td>
<td>3</td>
<td>1.60 (0.59 to 4.31)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;56&lt;/sup&gt; (HeNe 632.8 nm, 5 mW)</td>
<td>12</td>
<td>0.91 (0.51 to 2.91)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;57&lt;/sup&gt; (GaAs 904 nm, 4 mW)</td>
<td>26</td>
<td>1.63 (0.98 to 2.78)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laser v placebo Haker (1991a) (GaAs 904 nm, 12 mW)</td>
<td>9.5</td>
<td>49</td>
<td>10</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>ESWT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESWT v placebo&lt;sup&gt;54&lt;/sup&gt; (2000 pulses)</td>
<td>25</td>
<td>246</td>
<td>3</td>
<td>6</td>
<td>1.15 (0.75 to 1.77)</td>
</tr>
<tr>
<td>ESWT v placebo&lt;sup&gt;55&lt;/sup&gt; (1500 pulses)</td>
<td>14</td>
<td>71</td>
<td>3</td>
<td>4</td>
<td>1.02 (0.66 to 1.56)</td>
</tr>
<tr>
<td>Iontophoresis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iontophoresis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;56&lt;/sup&gt; (CUS 1 MHz, &gt;1)</td>
<td>99</td>
<td>10</td>
<td>6</td>
<td>1.29 (0.78 to 2.13)</td>
<td>0.32 (0.09 to 0.73)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;57&lt;/sup&gt; (PUS 1.4, 1.0 MHz, 1.0 W/cm&lt;sup&gt;2&lt;/sup&gt; 1.0 W/cm&lt;sup&gt;2&lt;/sup&gt; 10 min)</td>
<td>43</td>
<td>10</td>
<td>4</td>
<td>1.20 (0.60 to 2.38)</td>
<td>0.32 (0.78 to 1.85)</td>
</tr>
<tr>
<td>Laser v placebo&lt;sup&gt;58&lt;/sup&gt; (PUS 1.4, 1.0 W/cm&lt;sup&gt;2&lt;/sup&gt; 10 min)</td>
<td>40</td>
<td>9</td>
<td>5</td>
<td>2.70 (0.34 to 21.53)</td>
<td></td>
</tr>
<tr>
<td>Laser v US&lt;sup&gt;47&lt;/sup&gt; (PUS 1.4, 0.5–1.3 w/cm&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>4</td>
<td>40</td>
<td>9</td>
<td>5</td>
<td>3.27 (0.44 to 24.34)</td>
</tr>
</tbody>
</table>

Summary of validity score, sample size, and effect size (95% confidence interval) for global improvement, pain, maximum grip strength (MGS), and pain-free grip strength (PFGS) for study time intervals on included interventions. Effect size measured as standardised mean difference (SMD) for continuous outcome measures and relative risk (RR) for dichotomous outcome measures.

**INT** Number of treatment sessions; **Week** timing of outcome measures from baseline, N, no data available; **ESWT**, extracorporeal shock wave therapy; **cortic**, corticosteroid; **NSAID**, non-steroidal anti-inflammatory drug; **US**, ultrasound; **CUS**, continuous ultrasound; **PUS**, pulsed US; **DF**, deep friction massage.

US and acupuncture and found that, at four weeks follow up,

Of the two studies that looked at US versus placebo detuned US,<sup>55, 57</sup> the latter found no significant difference between groups for outcome measures of global improvement and PFGS (table 4). Although Lundberg et al<sup>57</sup> reported that PVAS was significantly improved in the US group at three months compared with the placebo group, there was no advantage of US over placebo on MGS or global improvement at three months (table 4). The pooled effect of global improvement at three months showed no difference between groups (RR 1.01; 95% CI 0.62 to 1.65). That is, pain, but not MGS or global improvement, was significantly better after US treatment compared with detuned US placebo at three months.

Davidson et al<sup>54</sup> compared eight treatment sessions of US and acupuncture and found that, at four weeks follow up,
there were significant improvements in all outcome measures for both groups, but that there was no difference between groups. They did not perform intention to treat analysis, nor allocation concealment with this study.

Stratford et al.\(^4\) used a 2 × 2 factorial design with phonophoresis (no drug versus drug) as one factor and transverse friction massage (no frictions versus frictions) as the other. Outcome measures after nine treatment sessions did not support the use of transverse friction massage or hydrocortisone coupling gel, finding no added benefits over that of US alone (table 4).

Based on the best evidence synthesis, there is insufficient evidence to either support or refute the use of US as a unimodal treatment for LE when based on pooled data and studies that compared it with other active treatments or a placebo.

**COMBINED PHYSICAL INTERVENTIONS**

Two studies evaluated a programme that combined a number of physical interventions—for example, deep friction massage, US, stretches, and exercises. Stujs et al.\(^2\) used this combined physical intervention as a comparator for wrist manipulation in the treatment of LE. Their data were presented in the manipulation section.

**Combined physical intervention of deep friction massage, US, and exercise**

Smidt et al.\(^6\) evaluated a six week combined programme of massage, US, and exercises, and compared this group with a corticosteroid injection group and a wait and see (control) group. At six weeks, there were significant differences between all three groups. The success rate at six weeks between the combined physical intervention and control group was weakly in favour of the intervention and this was maintained at 12 months, although these differences were not significant. There were no significant differences in the pain threshold between the physical intervention and the control group across all timed measures except for pressure pain threshold at six weeks and PFGS and patient satisfaction at 12 months (table 5).

At six weeks the corticosteroid injection group was significantly better than the physical intervention group in PVAS during the day, global improvement, MGS, pressure pain threshold, and PFGS. However, these differences were no longer evident by three months, and then reversed, with a significant difference favouring the physical intervention group at six and 12 months (table 5). A negative aspect of the apparent early success for corticosteroid injection in the first six weeks was a significant level of recurrence of the injury with this intervention at three and six months. From six months onwards, the combined physical intervention performed significantly better than the corticosteroid group, indicating that this was superior to corticosteroid injection in the long term but not significantly different from the control group.

In summary, there was evidence of a marginal advantage over the long term (more than six months) in using a combined physical intervention approach of deep friction massage, US, and exercise in the treatment of LE when compared with a corticosteroid injection, but not compared with a wait and see (no treatment) condition.

**DISCUSSION**

In this systematic review, the methodological quality of 76 studies was assessed, and the quality was found acceptable in 28 studies. Data from these 28 studies were extracted and reported as effect sizes for continuous data and relative risk for categorical data. It is important to understand that conclusions drawn are dependent on a few decision rules. These include the clinical relevance of results, the quality assessment schema used (modified PEDro, appendix 1) and the 50% acceptable quality level for inclusion of studies into this systematic review. To test the robustness of the rating scale and the cut-off point used, a sensitivity analysis found that the inclusion of studies rating greater than 40% did not change the conclusions drawn in this review. Another potential limitation to this review was that only one reviewer (LB) identified and screened all articles for inclusion, although this process was checked at several stages by a second reviewer (BV).

Although every effort was made to accurately translate the foreign language papers, one must consider the possibility that the accuracy of the quality assessment and data abstraction may be affected. Because of this, all foreign language papers are therefore identified as such in table 1 and the reference list.

Of the 15 criteria rated, two were conspicuously absent in a majority of papers (table 1). Criterion five, which evaluated the concealment of subject allocation, and criterion 12, which assessed the data analysis on an intention to treat basis, were not present in 86% and 93% of all papers respectively. In an observational study of 250 controlled trials from 33 meta-analyses, Schultz et al.\(^4\) reported that the estimated treatment effects were larger in studies that had inadequate allocation concealment. This potentially skews the findings of a systematic qualitative review in favour of the treatment under consideration. In addition to this potential bias, the lack of concealment of subject allocation and the dearth of analyses on an intention to treat basis reflects the difficulty in conducting clinical trials of a reasonable quality. The number of outcome measures used (table 2), the often inadequate description of the treatment parameters, and the small number of subjects in individual studies—that is, low-power—also contribute to the difficulty encountered in comparing studies and interpreting results. Nonetheless, there is a need for greater diligence in the design, implementation, and reporting of RCTs.

The lack of long term follow up studies for non-electro-therapeutic interventions limits any firm inferences to clinical practice. Not withstanding this, we found preliminary evidence of positive moderate to large effects (see Data management and statistical analysis) for joint manipulations such as MWM on measures of PFGS and pressure pain threshold in LE. The case was similar for exercise and tape. However, the MWM and tape studies were primarily mechanism studies which did not include long term follow up, reporting only initial effects. There is currently insufficient evidence for the use of manual therapy, tape, orthotics, and exercise in the management of LE, but there are preliminary studies that appear to warrant further evaluation.

A group of interventions have shown evidence of short term benefit (up to three months). This includes US, ionisation, and acupuncture. This effect appears to disappear, and there is evidence of no difference from placebo in the long term. To determine the applicability of these modalities in the management of LE therefore, the cost of the treatment needs to be weighed against the temporary relief in symptoms gained.

One RCT\(^9\) has shown the lack of benefit of a six week programme of a combined physical intervention compared with a control group, in both the short and long term. Although this is the result of only one study, the size and quality of this particular study in itself supports reasonable evidence of no effect for a combined physical intervention in the management of LE. The combined intervention in this RCT involved US, transverse friction massage, and exercise.
There is a growing body of evidence to show the lack of effect of laser and ESWT as treatments for LE, in both the short and long term. Pooled results for ESWT and laser treatment showed nil effect on pain, MGS, and global improvement, indicating that these treatments may not be beneficial in treating LE. This review differs from previous ones that have looked at the effect of ESWT11 65 on LE, as they concluded that there was insufficient evidence to either support or refute this modality in the management of LE. Moreover, the reported risk of adverse side effects from the use of high dose ESWT suggests that it should not be recommended as a course of treatment in this condition.

Despite the prevalence of LE and the substantial loss of work associated with this condition, there is surprisingly little consensus on its management. In summary, the fact remains that the literature to date has assessed the effect of a range of physical interventions on various outcomes for LE and failed to elucidate any long term beneficial effects over that of a placebo group. Evidence is accruing that does not support the effect of laser and ESWT as treatments for LE, in both the short and long term. Pooled results for ESWT and laser treatment showed nil effect on pain, MGS, and global improvement, indicating that these treatments may not be beneficial in treating LE. This review differs from previous ones that have looked at the effect of ESWT11 65 on LE, as they concluded that there was insufficient evidence to either support or refute this modality in the management of LE. Moreover, the reported risk of adverse side effects from the use of high dose ESWT suggests that it should not be recommended as a course of treatment in this condition.

ACKNOWLEDGEMENTS

We thank Hildegard Reul-Hirche, Tina Souvlis, Henry Tsao, and Tanja Sjelonek for help with translation of foreign language papers, and Ross Darnell for help with data abstraction.

What is already known on this topic

Lateral epicondylalgia (LE) is a chronic condition that has proven to be challenging to treat in a primary care setting. To date, systematic reviews looking at a range of outcome measures on treatments for LE have failed to draw any firm conclusions as to what treatment is most effective in managing this condition.

What this study adds

This systematic review has assessed literature not included in previous reviews. Pooling of some data was possible and added support to the findings of this review. There is evidence that extracorporeal shockwave therapy is not beneficial in the treatment of LE.

### Table 5 Combined physical interventions for lateral epicondylalgia (LE)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Duration of LE (months)</th>
<th>Sample size (N)</th>
<th>INT</th>
<th>Week</th>
<th>Global improvement (RR)</th>
<th>MGS (SMD)</th>
<th>PGFS (SMD)</th>
<th>Pressure pain threshold (SMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-friction massage + exercise</td>
<td>11</td>
<td>183</td>
<td>9</td>
<td>6</td>
<td>0.51 (0.39 to 0.67)</td>
<td>1.20</td>
<td>-0.61</td>
<td>-1.07</td>
</tr>
<tr>
<td>versus corticosteroid injection</td>
<td>12</td>
<td>26</td>
<td>52</td>
<td>1.31 (1.09 to 1.57)</td>
<td>0.40 (0.05 to 0.76)</td>
<td>0.15</td>
<td>-0.20 to 0.50</td>
<td>0.71 (0.35 to 1.07)</td>
</tr>
</tbody>
</table>

Summary of validity score, sample size, and effect size (95% confidence interval) for global improvement, pain, maximum grip strength (MGS), and pain-free grip strength (PGFS) and pressure pain threshold for study time intervals on included interventions. Effect size measured as standardised mean difference (SMD) for continuous outcome measures and relative risk (RR) for dichotomous outcome measures.

INT, Number of treatment sessions; Week, timing of outcome measures from baseline; US, ultrasound.

---

APPENDIX 1

Rating scale and description of criteria developed from the PEDro rating scale (additional criteria 2, 3, 10, and 15). PEDro criteria reprinted with permission.

1. Eligibility criteria were specified. This criterion is satisfied if the report describes the source of subjects and a list of criteria used to determine who was eligible to participate in the study.

2. The sample size was justified. Was statistical evidence provided regarding the power of the study and its effect size?

3. Ethical clearance and consent were obtained for the study. Was there mention of the study being approved by a research ethics body, and was written consent obtained from the subjects for their participation in the study?

4. Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received). A study is considered to have used random allocation if the report states that allocation was random. The precise method of randomisation need not be specified. Procedures such as coin tossing and dice rolling should be considered random. Quasi-randomisation allocation procedures such as allocation by hospital record number or birth date, or alternation, do not satisfy this criterion.

5. Allocation was concealed. Concealed allocation means that the person who determined if a subject was eligible for inclusion in the trial was unaware, when this decision was made, of which group the subject would be allocated to. A point is awarded for this criterion, even if it is not stated that allocation was concealed, when the report states that allocation was by sealed opaque envelopes or that allocation involved contacting the holder of the allocation schedule who was “off site”.

---

Authors’ affiliations

L Bisset, A Paungmali, Division of Physiotherapy, University of Queensland, St Lucia, Queensland, Australia

B Vicenzino, Musculoskeletal Pain and Injury Research Unit, Division of Physiotherapy, University of Queensland

E Beller, Queensland Clinical Trials Centre, University of Queensland

Financial support: The University of Queensland Research Development Scheme; National Health and Medical Research Council Project Grant No 252710.

Competing interests: none declared
(6) The groups were similar at baseline with regard to the most important prognostic indicators. At a minimum, in studies of therapeutic interventions, the report must describe at least one measure of the severity of the condition being treated and at least one (different) key outcome measure at baseline. The rater must be satisfied that the groups’ outcomes would not be expected to differ, on the basis of baseline differences in prognostic variables alone, by a clinically significant amount.

Criteria 6, 10–14. Key outcomes are those outcomes that provide the primary measure of the effectiveness (or lack of effectiveness) of the therapy. In most studies, more than one variable is used as an outcome measure.

(7) There was blinding of all subjects.

(8) There was blinding of all therapists who administered the treatment.

(9) There was blinding of all assessors who measured at least one key outcome.

Criteria 7–9. Blinding means that the person in question (subject, therapist, or assessor) did not know which group the subject had been allocated to. In addition, subjects and therapists are only considered to be “blind” if it could be expected that they would have been unable to distinguish between the treatments applied to different groups. In trials in which key outcomes are self-reported—for example, visual analogue scale, pain diary—the assessor is considered to be blind if the subject was blind.

(10) The measurements are known to be reliable. Were the outcome measures used referenced for their reliability? If more than one assessor was used for the outcome measures, were inter-tester reliability studies performed?

(11) Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups. This criterion is only satisfied if the report explicitly states both the number of subjects initially allocated to groups and the number of subjects from whom key outcome measures were obtained. In trials in which outcomes are measured at several points in time, a key outcome must have been measured in more than 85% of subjects at the time of primary interest.

(12) All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention to treat”. An intention to treat analysis means that, where subjects did not receive treatment (or the control condition) as allocated, and where measures of outcomes were available, the analysis was performed as if subjects received the treatment (or control condition) they were allocated to. This criterion is satisfied, even if there is no mention of analysis by intention to treat, if the report explicitly states that all subjects received treatment or control conditions as allocated.

(13) The results of between group statistical comparisons are reported for at least one key outcome. A between group statistical comparison involves statistical comparison of one group with another. Depending on the design of the study, this may involve comparison of two or more treatments, or comparison of treatment with a control condition. The analysis may be a simple comparison of outcomes measured after the treatment was administered, or a comparison of the change in one group with the change in another (when a factorial analysis of variance has been used to analyse the data, the latter is often reported as a group/time interaction). The comparison may be in the form hypothesis testing (which provides a “p” value, describing the probability that the groups differed only by chance) or in the form of an estimate (for example, the mean or median difference, or a difference in proportions, or number needed to treat, or a relative risk or hazard ratio) and its confidence interval.

(14) The study provides both point measures and measures of variability for at least one key outcome. A point measure is a measure of the size of the treatment effect. The treatment effect may be described as a difference in group outcomes, or as the outcome in (each of) all groups. Measures of variability include standard deviations, standard errors, confidence intervals, interquartile ranges (or other quantile ranges), and ranges. Point measure and/or measures of variability may be provided graphically (for example, SDs may be given as error bars in a figure) as long as it is clear what is being plotted—for example, as long as it is clear whether error bars represent SDs or SEs. Where outcomes are categorical, this criterion is considered to have been met if the number of subjects in each category is given for each group.

(15) Side effects were reported. Does the beneficial effect of the intervention outweigh the adverse effects? Were side effects reported?

REFERENCES

1 Allander E. Prevalence, incidence and remission rates of some common rheumatic diseases or syndromes. J Rheumatol 1974;1:145–53.
Physical interventions for lateral epicondylalgia

Molsberger A
Rompe J
Clark M
39
Review Manager (RevMan) [computer programme]
Krasheninnikoff M
46
51
Rompe J
52
35
Fink M
Jensen B
44
21
Vicenzino B
44
Haker E
44

A quick search of PubMed using the relevant MeSH headings ("Tennis Elbow/rehabilitation" OR "Tennis Elbow/therapy") returns over 500 hits. This systematic review by Bisset and colleagues tells us that this body of literature amounts to little of substance.

The authors found 28 randomised trials of physical interventions for lateral epicondyliitis that satisfied at least eight of 15 quality criteria. These 28 trials constitute most of what we know about the effects of physical interventions for lateral epicondyliitis.

The trials themselves provide little guidance for clinical practice. As the authors point out, the trials tend to be small (the median sample size is 21 per group), and only eight trials provided long term follow up. Several trials just examined the immediate effect of a single treatment.

The clearest finding is that extracorporeal shock wave therapy does not produce clinically meaningful reductions in pain, but occasionally produces mild adverse effects. This is consistent with the findings from recent high quality randomised trials, which show there is no real clinical benefit from the use of extracorporeal shock wave therapy for plantar fascitis and rotator cuff tendinitis.3,4

The data suggest that a single session of manipulation of the elbow or neck can produce immediate short term reductions in pain, but it is not known how long the effects last. In addition, one study found that exercise produces short term reductions in pain. However, estimates of the size of the effects of manipulation and exercise are associated with considerable uncertainty. These studies cannot tell us if the effects of exercise and manipulation are large enough and sustained enough to be clinically worthwhile. The best single trial showed that corticosteroid injections are much more effective than physiotherapy (ultrasound, friction massage, and exercise) in the short term, but physiotherapy is more effective than corticosteroid injections in the long term.5

What are the implications for clinical practice? Extracorporeal shock wave therapy should not be recommended. Current evidence suggests that corticosteroid injections offer the best prospects for short term relief of symptoms, and physiotherapy (perhaps the exercise component of physiotherapy) offers the best prospect for good long term outcomes.

R D Herbert
School of Physiotherapy and Centre for Evidence-Based Physiotherapy, University of Sydney, Sydney, NSW, Australia; R.Herbert@fhs.usyd.edu.au

REFERENCES