LACK OF SCIENTIFIC EVIDENCE FOR THE TREATMENT OF LATERAL EPICONDYLITIS OF THE ELBOW

AN ATTEMPTED META-ANALYSIS

H. LABELLE, R. GUIBERT, J. JONCAS, N. NEWMAN, M. FALLAHA, C-H. RIVARD

From the University of Montreal and Sainte-Justine Hospital

We have reviewed 185 articles published since 1966 to assess the scientific evidence for methods of treatment for lateral epicondylitis of the elbow. Of the 185 articles, 78 discussed treatment, but since the natural history of the syndrome is uncertain we considered only those series with concurrent control groups. Only 18 of these were randomised and controlled studies. We then graded these papers for scientific validity, using the methods of Chalmers et al (1981).

The mean score of the 18 articles was only 33%, with a range from 6% to 73%. A minimum of 70% is required for a valid clinical trial, and we therefore concluded that there was insufficient scientific evidence to support any of the current methods of treatment. There were too many methodological differences to allow a quantitative meta-analysis, but our qualitative review established the importance of the natural evolution of the syndrome and of the placebo effect of all treatments. Properly designed, controlled trials are needed.

Lateral epicondylitis (LE) or 'tennis elbow' is common and results in considerable morbidity and financial cost. In Sweden the prevalence is 1% to 3%, but increases to 19% in men between 40 and 50 years of age (Allander 1974). The syndrome is seen mainly in those whose work involves repetitive movements of the forearm and hand: it is reported that 7.4% of industrial workers and 40% to 50% of tennis players in the USA are at some time affected by it (Dimberg 1987). In 1981, over 1500 workers

in Quebec made claims to the workers' compensation board for this condition, generating a cost of \$4442 (Canadian) per case and an average of 62 days of work lost (1982 Province of Quebec, Workman's Compensation Board statistics).

Many papers have been published on LE since the first clinical description (Runge 1873), and review articles have reported over 40 different treatments used either alone or in combination. These include anti-inflammatory drugs, steroids, many physiotherapy techniques, cast immobilisation, orthoses, surgical operations and less conventional methods such as radiotherapy, acupuncture and vitamins.

The choice of treatment for an individual case remains controversial because it is empirical and based on the personal experience of the physician treating the patient. Our aim was to review the current literature to determine if there was any significant scientific evidence to favour a particular method of treatment for acute lateral epicondylitis.

MATERIAL

We used the Medline and EMBASE (Excerpta Medica) databases to make a thorough review of the English and French medical and paramedical literature for the years 1966 to 1990 (inclusive), locating and studying 185 articles. These included 78 on medical treatment or surgery, 35 current concepts or review articles, 27 on the

H. Labelle, MD, FRCS C, Associate Professor of Surgery, University of Montreal

J. Joncas, RN, Research Assistant

C-H. Rivard, MD, FRCS C, Professor of Surgery, University of Montreal
Orthopaedic Service, Sainte-Justine Hospital, 3175 Côte Sainte-Catherine, Montreal, Quebec, Canada H3T 1C5.

R. Guibert, MD, MSc, Assistant Professor, Department of Family Medicine and Research Groupe de Recherche Interdisciplinaire en Santé
University of Montreal, PO Box 6128-Succ. A, Montreal, Quebec, Canada H3C 3J7.

N. Newman, MD, FRCS C, Assistant Professor of Surgery, University of Montreal
Hôtel-Dieu Hospital, 3840 Sainte-Urbain Street, Montreal, Quebec, Canada H2W 1T8.

M. Fallaha, MD, FRCS C, Assistant Professor of Surgery, University of Montreal
Maisonnette-Rosemont Hospital, 5415 L'Assomption Boulevard, Montreal, Quebec, Canada H1T 2M4.

Correspondence should be sent to Dr H. Labelle.

© 1992 British Editorial Society of Bone and Joint Surgery

0301-620X/92/5377 \$2.00

J Bone Joint Surg [Br] 1992; 74-B:646-51.

Table I. The study design of 78 published articles on the treatment of lateral epicondylitis of the elbow

Study	Number
Case reports	7
Case series without controls	51
Case series with historical controls	2
Controlled non-randomised trial	1
Controlled randomised trial	3
Controlled randomised single-blind trial	1
Controlled randomised double-blind trial	13

Table II. Evaluation form A, adapted from Chalmers et al (1981), showing the 15 items scored to evaluate the study design of a clinical trial

Items	Possible points	Items	Possible points
1 Selection description	0 to 3	9 Observer blinded to results	0 to 4
2 Reject log*	0 to 3	10 Prior estimate of numbers	0 to 3
3 Therapeutic regimen definition	0 to 3	11 Stopping rules	0 to 3
4 Control appearance	0 to 1.5	12 Testing randomisation	0 to 3
5 Control taste†	0 to 1.5	13 Testing blinding	0 to 3
6 Randomisation blinding	0 to 10	14 Testing compliance	0 to 3
7 Patient blindability	0 to 8	15 Biological equivalence	0 to 3
8 Observer blindability	0 to 8	<i>Total possible score 60</i>	

*a full record of eligible population *not* accepted for the trial
 †confirmation that the taste (of oral medication), or the sensation of a treatment does not differ between treated and control groups

Table III. Evaluation form B adapted from Chalmers et al (1981) showing the 13 items scored to evaluate the data analysis of a clinical trial

Items	Possible points	Items	Possible points
1 Dates of study description	0 to 2	8 Regression/correlation	0 to 2
2 Results of randomisation	0 to 2	9 Statistical analysis	0 to 4
3 Major end-points	0 to 4	10 Withdrawals	0 to 4
4 Post type 2 estimate	0 to 3	11 Handling withdrawals	0 to 4
5 Confidence limits	0 to 3	12 Side-effects	0 to 3
6 Time series analysis	0 to 2	13 Retrospective evaluation	0 to 3
7 Timing of events	0 to 4	<i>Total possible score 40</i>	

diagnosis and pathology of the disease, 24 letters to editors, 15 articles on measuring instruments and six on the biomechanical aspects of the syndrome.

We found no information in these articles on the natural course of the disease which could explain the variations in treatment. The definitions of acute and chronic epicondylitis were unclear and varied enormously.

There are two important consequences. First, ignorance of the natural history of LE means that treatment seems to be based on either the personal experience of the different authors, or on a few even earlier reports rather than on scientific evidence. Secondly, the results of uncontrolled therapeutic trials must remain doubtful since the rate of recovery or improvement of the syndrome in the absence of treatment is unknown. The differing selection of cases from one study to another must lead to potentially important differences in the stage of evolution of the disease, and make such uncontrolled studies inconclusive.

Table I classifies the 78 reviewed articles on the treatment of LE according to the design of the study. Most of them did not include control groups, and were therefore excluded from further analysis. We made a more detailed study of the 18 controlled trials in which the study design had a minimum of internal validity.

METHOD

We initially intended to make a quantitative meta-analysis but found it impossible to group the studies because of the considerable variation in treatment, selection criteria and measures of efficacy. We therefore made a qualitative study of the 18 selected papers using the system proposed by Chalmers et al (1981) which evaluates the design, conduct and analysis of papers by scoring the quality of the research carried out. It is therefore possible to interpret the results appropriately, especially when there are conflicting results between two studies, or studies with small sample sizes and inconclusive results.

We used Chalmers' technique with great care. Two evaluation forms, with 28 individually scored items, allowed for a maximum score of 100. These forms were completed independently by four trained health professionals (HL, JJ, NN, MF), with adequate experience in clinical trials. To reduce observer bias, only the material and method and results sections of each paper were photocopied, so that the authors' names, the journal that published the paper, the discussion and the conclusions were unknown to the researcher. Differences in scores among the evaluators were resolved at group discussions.

The contents of the evaluation forms are summarised in Tables II and III. The first form examined the study design by giving particular importance to blinding of patients, physicians and measuring observers, and the presence and method of randomisation of the patients

and of the physicians when applicable. The second form evaluated the quality of the data analysis, the statistical analysis and the presentation of results. We used a revision of Chalmers' original methods of 1981 updated in 1984 (personal communication), and we believe that this is the first use of this evaluation technique in the orthopaedic literature.

RESULTS

Tables IV to VIII give our evaluations for the 18 clinical trials, expressing results as percentages of the maximum possible score, and allowing for items which were not applicable to every study and which were therefore excluded from the calculations.

The average score for the 18 trials was 33%, with a minimum of 6% for the weakest study design (Kivi 1983) and a maximum of 73% for the strongest one (Stratford et al 1989). This system has been applied to several hundred therapeutic trials, and an arbitrary score of 70% is considered to be the minimum required for a good quality design for controlled therapeutic trials. The average for our 18 studies was very low, only one exceeding 70%.

DISCUSSION

It must be kept in mind that the 18 trials chosen were the best studies *with a control group*, and interpretation of our results and conclusions must be considered in this light. Chalmers' evaluation is a quality index; it puts the results, discussion and conclusion sections of a study into perspective. This is particularly useful in assessing conflicting results reported in different studies.

We further classified the papers into groups according to the therapy used.

Ultrasound (Table IV). In this group the trial of Binder et al (1985) is the only one that showed a significant therapeutic effect of ultrasound as against placebo

ultrasound ($p < 0.01$, $n = 76$). The study scored 44% as compared with that of Lundeburg, Abrahamsson and Haker (1988) which scored 38% and showed no significant difference between the ultrasound and the placebo, although it did demonstrate superiority of ultrasound over rest only ($p < 0.01$, $n = 99$). These results suggest that there is a significant placebo effect with ultrasound treatment. Unfortunately, neither of the studies includes a power analysis which could estimate the possibility of a type 2 error, i.e., missing an existing difference when comparing ultrasound and placebo.

One excellent study, by Stratford et al (1989), with a score of 73%, showed no significant difference between ultrasound with phonophoresis and placebo ultrasound, or between phonophoresis with or without friction massage in 40 subjects. Finally, a study by Halle, Franklin and Karalja (1986), with a score of 30%, confirmed the previous conclusions, finding no significant difference between ultrasound with phonophoresis, placebo ultrasound, transcutaneous electrical nerve stimulation (TENS), and steroid infiltrations with hydrocortisone in 48 subjects. In the last two studies samples were too small to avoid a type 2 error.

Ultrasound may therefore have a therapeutic effect: a significant improvement from baseline values was detected in all four studies, but whether this effect is superior to that of a placebo remains unclear. The addition of friction massage or cortisone ointment did not appear to have any significant advantages. More controlled trials with larger sample sizes are needed to establish the therapeutic value of this treatment.

Ionisation (Table V). Three trials examined the therapeutic effect of ionisation. Famaey et al (1982), with a score of 27%, reported a significant difference between ionisation with a diclofenac solution compared with a placebo solution ($p < 0.01$). This was only partially confirmed by two Italian studies. Vecchini and Grossi (1984), with a score of 45%, showed only a moderate therapeutic effect in 11 patients treated by ionisation and diclofenac

Table IV. Results for four papers on treatment by ultrasound

Authors	Sample size	Method of treatment	Conclusions	Quality score (%)
Binder et al 1985	76	Ultrasound vs placebo	Ultrasound better ($p < 0.01$) Improvement in both groups	44
Lundeburg et al 1988	99	Ultrasound vs placebo ultrasound vs rest	Ultrasound = placebo but better than rest ($p < 0.01$)	38
Stratford et al 1989	40	Ultrasound + hydrocortisone ointment without friction vs ultrasound + placebo + friction vs ultrasound + placebo without friction	No difference Improvement in every group	73
Halle et al 1986	48	Ultrasound vs phonophoresis vs TENS vs steroid injection	No difference Improvement in every group	30

Table V. Results for three papers on treatment by ionisation

Authors	Sample size	Method of treatment	Conclusions	Quality score (%)
Famaey et al 1982	97	Diclofenac ionisation vs placebo ionisation	Diclofenac better ($p < 0.01$) Improvement in both groups	27
Vecchini and Grossi 1984	24	Diclofenac ionisation vs placebo ionisation	Improvement in both groups ($p = 0.005$) Pharmacological ionisation better ($p < 0.001$)	45
Grossi et al 1986	73	Sham ionisation vs placebo ionisation vs pirprofen ionisation (two-dose levels)	Improvement ($p < 0.01$) for all treatments Pharmacological ionisation better ($p < 0.001$)	56

Table VI. Results for three papers on treatment with non-steroid anti-inflammatory drugs

Authors	Sample size	Method of treatment	Conclusions	Quality score (%)
Rosenthal 1984	50	Flurbiprofen vs piroxicam	No difference Improvement in both groups*	44
Saartok and Eriksson 1986	21	Naproxen vs steroid injection	No difference No improvement	37
Adelaar et al 1987	18	Diflunisal vs naproxen	No difference No improvement	11

*($p < 0.001$)

solution as compared with 13 treated by ionisation and a placebo, and these authors could not exclude the possibility that the therapeutic effect in diclofenac-treated patients was due to systemic absorption. Grossi et al (1986), scoring 56%, compared four different types of treatment: sham ionisation, ionisation with a placebo and ionisation with two different dosages of pirprofen. All four treatments gave a significant improvement ($p < 0.01$), being greatest for the two groups with pirprofen ($p < 0.001$). There was no difference between sham ionisation and placebo ionisation, which suggests that there is a placebo effect with simple ionisation.

Table VII. Results for five papers on treatment by steroid injection

Authors	Sample size	Method of treatment	Conclusions	Quality score (%)
Day et al 1978	95	Methylprednisolone vs xylocaine vs saline	Steroid better ($p < 0.001$) Xylocaine = saline	26
Brattberg 1983	60	Acupuncture vs steroid injection	Acupuncture better ($p < 0.005$) No improvement with steroids	17
Clarke and Woodland 1975	50	Methylprednisolone vs hydrocortisone	No difference Improvement in both groups	13
Hughes and Currey 1969	50	Steroid hypospray vs steroid injection	No difference Improvement in both groups	7
Kivi 1983	80	Oral indomethacin vs steroid injection	No difference Improvement in both groups	6

Table VIII. Results for three papers on other treatments

Authors	Sample size	Method of treatment	Conclusions	Quality score (%)
Burton 1988	33	Manipulation vs manipulation with forearm strap vs manipulation with topical NSAID*	No difference Improvement in all groups	19
Percy and Carson 1981	80	Dimethyl sulphoxide 40% vs placebo	No difference Improvement in both groups	41
Devereaux et al 1985	30	Electromagnetic field vs placebo	No difference Improvement in both groups	63

*non-steroid anti-inflammatory drugs

These three studies appear to show that ionisation, of itself, has little therapeutic effect and that ionisation with non-steroidal anti-inflammatory drugs (NSAID) may work systemically rather than through a local effect. **Oral NSAID** (Table VI). None of the three studies with NSAID (scores of 11%, 37% and 44% respectively) showed any significant differences between the various anti-inflammatory drugs. Adelaar, Maddy and Emroch (1987) compared naproxen with diflunisal in 18 patients; Rosenthal (1984) compared piroxicam with flurbiprofen in 50 patients; and Saartok and Eriksson (1986) compared naproxen with steroid injections in 21 patients. Only the study by Rosenthal showed a significant difference ($p < 0.001$) before and after the anti-inflammatory treatment. There was no placebo-controlled trial in this group, and it therefore appears that the therapeutic value of oral NSAID is at present unknown.

Steroid injections (Table VII). Five therapeutic trials, all with low scores, examined treatment with various local steroid injections. Only that by Day, Govindasamy and Patnaik (1978), scoring 26%, showed a significant therapeutic effect of methylprednisolone solution compared with xylocaine and saline solutions in 95 subjects ($p < 0.001$). Brattberg (1983) scoring 17%, contradicted these results: he found significantly better ($p < 0.005$, $n = 60$) results with acupuncture than with steroid injection: the latter group showed no improvement. Neither Hughes and Currey (1969) nor Clarke and Woodland (1975), scoring 7% and 13%, found any differences between various preparations of steroids. All the treatments appeared to show some effect, but there was a 66% recurrence rate six months after treatment. Kivi (1983), scoring only 6%, reported similar results when comparing various oral steroids with indomethacin in 80 patients.

Most studies in this group suggest a positive therapeutic effect for steroid injections, but the very low scores and some conflicting results make further investigations mandatory.

Other treatments (Table VIII). Three studies examined methods of treatment infrequently used. Burton (1988), scoring 19%, found no difference between physical manipulation, manipulation with forearm straps, and/or anti-inflammatory topical creams; all gave similar improvement. Percy and Carson (1981), scoring 41%, compared topical dimethyl sulphoxide and a placebo cream in 80 patients, finding a significant improvement in 53% of the placebo group, and no difference between the two groups. Devereaux, Hazleman and Thomas (1985), scoring 63%, compared electromagnetic fields with a placebo. Both groups improved significantly and again there was no difference between them.

To summarise, critical analysis of the literature has revealed four trends:

1) Most of the randomised and controlled trials on the treatment of LE have a poor methodological design, as reflected by their low average score. This means that the

results and conclusions drawn from the five studies which reported positive effects for ultrasound, steroid injections, ionisation or acupuncture over other methods of treatment must be interpreted with caution: many potential biases may have altered the results.

2) A definite placebo effect was shown in six studies, which was equivalent to that of the treatment in three of them. This suggests that there could be significant improvements with any treatment, due either to a placebo effect or to the natural history of LE. It underlines the importance of comparing all methods of treatment with a control group, and supports our decision not to consider the published results of case series without controls.

3) The sizes of the sample groups were generally small (average 54, range 18 to 99). This may help to explain why 13 of the 18 trials found no significant differences between various treatments, because of possible type 2 errors.

4) Significant improvements from baseline values were found in all but two studies, confirming either that all treatments have a therapeutic effect or that LE improves spontaneously with time. Since the natural history of the syndrome is poorly documented, we must assume that the latter may be correct.

Conclusions. The poor quality and contradictory results of the randomised and controlled trials reported so far in the literature means that there is not enough *scientific* evidence to favour any particular type of treatment for acute lateral epicondylitis. The importance of a placebo effect and the natural evolution of the syndrome emphasises the need to evaluate all types of therapy with properly controlled, randomised trials, using double-blind methods where possible.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

REFERENCES

- Adelaar RS, Maddy L, Emroch KS. Diflunisal vs Naproxen in the management of mild to moderate pain associated with epicondylitis. *Adv in Therapy* 1987; 4:317-27.
- Allander J. Prevalence incidence and remission rates of some common rheumatic diseases and syndromes. *Scand J Rheumatol* 1974; 3:145-53.
- Binder A, Hodge G, Greenwood AM, Hazleman BL, Page Thomas DP. Is therapeutic ultrasound effective in treating soft tissue lesions? *Br Med J* 1985; 290:512-4.
- Brattberg G. Acupuncture therapy for tennis elbow. *Pain* 1983; 16: 285-8.
- Burton AK. A comparative trial of forearm strap and topical anti-inflammatory as adjuncts to manipulative therapy in tennis elbow. *Manual Medicine* 1988; 3:141-3.
- Chalmers TC, Smith H Jr, Blackburn B, et al. A method for assessing the quality of a randomized control trial. *Controlled Clin Trials* 1981; 2:31-49.
- Clarke AK, Woodland J. Comparison of two steroid preparations used to treat tennis elbow using the hypospray. *Rheumatol Rehabil* 1975; 14:47-9.

- Day BH, Govindasamy N, Patnaik R.** Corticosteroid injections in the treatment of tennis elbow. *Practitioner* 1978; 220:459-62.
- Devereaux MD, Hazleman BL, Thomas PP.** Chronic lateral humeral epicondylitis: a double-blind controlled assessment of pulsed electromagnetic field therapy. *Clin Exp Rheumatol* 1985; 3:333-6.
- Dimberg L.** The prevalence and causation of tennis elbow (lateral humeral epicondylitis) in a population of workers in an engineering industry. *Ergonomics* 1987; 30:573-9.
- Famaey JP, Brounx G, Cleppe D, et al.** Ionisation with Voltaren: a multi-centre trial. *J belge Méd Phys* 1982; 5:55-60.
- Grossi E, Monza GC, Pollavini S, Bona L.** NSAID ionisation in the management of soft-tissue rheumatism: role played by the drug, electrical stimulation and suggestion. *Clin Exp Rheumatol* 1986; 4:265-7.
- Halle JS, Franklin RJ, Karalja BL.** Comparison of four treatment approaches for lateral epicondylitis of the elbow. *J Orthop and Sports Physical Therapy* 1986; 8:62-9.
- Hughes GR, Currey HL.** Hypospray treatment of tennis elbow. *Ann Rheum Dis* 1969; 28:58-62.
- Kivi P.** The etiology and conservative treatment of humeral epicondylitis. *Scand J Rehabil Med* 1983; 15:37-41.
- Lundeberg T, Abrahamsson P, Haker E.** A comparative study of continuous ultrasound, placebo ultrasound and rest in epicondylalgia. *Scan J Rehabil Med* 1988; 20:99-101.
- Percy EC, Carson JD.** The use of DMSO in tennis elbow and rotator cuff tendonitis: a double-blind study. *Med Sci Sports Exerc* 1981; 13:215-9.
- Rosenthal M.** The efficacy of flurbiprofen versus piroxicam in the treatment of acute soft tissue rheumatism. *Curr Med Res Opin* 1984; 9:304-9.
- Runge F.** Zur Genese und Behandlung des Schreibekrampfes. *Berliner Klin Wchnschr* 1873; 10:245-8.
- Saartok T, Eriksson E.** Randomized trial of oral naproxen or local injection of betamethasone in lateral epicondylitis of the humerus. *Orthopedics* 1986; 9:191-4.
- Stratford PW, Levy DR, Gauldie S, Miseferi D, Levy K.** The evaluation of phonophoresis and friction massage as treatments for extensor carpi radialis tendinitis: a randomized controlled trial. *Physiotherapy Canada* 1989; 41:93-9.
- Vecchini L, Grossi E.** Ionization with Diclofenac Sodium in rheumatic disorders: a double-blind placebo-controlled trial. *J Int Med Res* 1984; 12:346-50.