Conservative treatments for tennis elbow—do subgroups of patients respond differently?

L. Bisset¹,², N. Smidt³,⁴, D. A. Van der Wind⁵,⁶, L. M. Bouter⁵, G. Jull¹, P. Brooks⁷ and B. Vicenzino¹

Objectives. To determine if subgroups of patients with tennis elbow respond differently in treatment.

Methods. This study used individual patient data (n=383) from two randomized controlled trials that investigated a wait-and-see policy, corticosteroid injections and physiotherapy. Common outcome measures were: pain severity, global improvement, severity assessed by a blinded assessor, elbow disability and pain free grip strength. Subgroup analyses for previous history of elbow pain, baseline pain severity, duration of the current episode and employment status were performed at 6 and 52 weeks.

Results. Patients' age, previous elbow symptoms and baseline pain severity were similar between trials, but other characteristics differed between trial populations. Based on individual patient data from both trials, we found that corticosteroid injections were statistically and clinically superior at 6 weeks, but significantly worse at 52 weeks compared with both wait-and-see and physiotherapy. Subgroup effects were scarce and small. Patients with higher baseline pain score showed less benefit on pain outcomes between physiotherapy and a wait-and-see policy at 6 weeks. It also appeared that non-manual workers who had an injection were the only work subgroup to follow the general trend that injections were significantly worse than a wait-and-see policy on global improvement at 52 weeks.

Conclusion. The treatment outcomes were largely similar between trials and not different between most subgroups studied. In tennis elbow, it would appear that patient characteristics play only a small role in predicting treatment outcomes, which supports the generalizability of individual trial results.

KEY WORDS: Tennis elbow, Randomized controlled trials, Corticosteroid injections, Physiotherapy.

Introduction

Lateral epicondylalgia (tennis elbow) is a common musculoskeletal condition [1–5], which is often treated by physiotherapy [6, 7], corticosteroid injections [7–9] or by adopting a wait-and-see policy. Two recent randomized controlled trials (RCTs) have assessed these three interventions and found corticosteroid injections to have a strong initial effect, but poorer long-term outcomes [10, 11]. Smidt et al. [11] in a primary care RCT found a physiotherapy approach of deep friction massage, ultrasound and exercise gave no added benefit over a wait-and-see approach. A recent RCT by Bisset et al. [10] found superior effects for a physiotherapy intervention of ‘mobilization with movement’ of the elbow and therapeutic exercise in the short-term (6 weeks) over a wait-and-see policy. In the long-term (52 weeks), both studies found superior effects for physiotherapy compared with corticosteroid injections, but no difference between physiotherapy and a wait-and-see policy.

Whilst clinical trials provide insight into the efficacy of interventions for the average population, the effects of treatments may be different for patients with specific characteristics. Previous authors have suggested that patients with greater pain severity, longer duration, previous history of concomitant neck and shoulder pain, manual work and involvement of the dominant side, are associated with poorer outcomes [12–14]. It is unclear if these factors indeed modify treatment effects.

Valuable information on patient characteristics that modify treatment outcomes in tennis elbow may be gained through subgroup analyses. However, subgroup analyses within a single RCT are often hampered by insufficient statistical power to reveal differences in treatment effects, as RCTs are usually only powered to reveal differences between the intervention groups under evaluation [15]. Pooling of individual patient data from more than one RCT increases the statistical power, and thus the possibility of identifying relevant subgroup effects [16].

The main aim of this study was to examine the influence of certain population characteristics (e.g. duration of symptoms, baseline pain severity) on the effect of conservative treatment in patients with tennis elbow.

Methods

This study combines individual patient data from two pragmatic RCTs that studied conservative management of clinically diagnosed tennis elbow of >6 weeks duration. One of the RCTs, conducted by Smidt et al. [11] recruited 185 patients from 259 potential candidates who were identified by general medical practitioners in Amsterdam, Amstelveen, Alkmaar, Purmerend and Haarlem (The Netherlands); whereas the other RCT undertaken by Bisset et al. [10] recruited 198 patients from 497 responders to a public call for volunteer participants in the Brisbane region (Australia). Thus, the combined data set had a sample size of 383. Details regarding the selection and informed consent procedures for each trial are described elsewhere. Ethical approval for this study was obtained through the Medical Research Ethics Committee of the University of Queensland and all participants provided informed consent prior to participation.

Both RCTs used similar inclusion and exclusion criteria. The Dutch RCT accepted an age range of 18–70 yrs and the Australian RCT had an age limit of 18–65 yrs. Furthermore, the Dutch RCT excluded those who had had treatment of their elbow symptoms with physiotherapy or injections in the preceding 6 months, whereas the Australian RCT excluded patients who had had any treatment of the elbow pain by a health care
practitioner within the preceding 6 months (pain medication, braces or stretches excepted).

In both RCTs, patient characteristics were recorded prior to randomization, including age, sex, employment status (manual, non-manual, no work), duration of elbow symptoms and previous episodes of elbow symptoms. Both RCTs utilized concealed random allocation of participants to interventions and intention to treat analysis.

Treatments
In the Dutch RCT, 64 patients were allocated to physiotherapy, which consisted of nine treatments over 6 weeks of deep friction massage, pulsed ultrasound [17] and a progressive exercise programme [18], 62 patients received at least one (maximum of three) local injection of 1 ml triamcinolone acetonide with 1 ml lidocaine and 59 were treated according to a wait-and-see policy of a single session with their family doctor where ergonomic advice and medication (paracetamol or non-steroidal anti-inflammatory drugs) were given. In the Australian RCT, 66 participants were allocated to eight sessions (30 min duration) of elbow manipulation and exercise, 65 participants were assigned to the same corticosteroid injection as used in the Dutch RCT but limited to a maximum of two injections, and 67 participants were managed by a wait-and-see approach (ergonomic advice, medication was not given but allowed as required).

Outcome measures
For the combined analysis, we used the follow-up times of 6 and 52 weeks after randomization, as well as baseline prior to randomization. The following outcome measures were employed in both RCTs: global improvement, pain severity, pain-free grip strength, blinded assessor’s rating of severity and elbow disability on a pain-free function questionnaire. In both RCTs, global improvement was measured on a six-point Likert scale (completely recovered to ‘much worse’) and a successful outcome was qualified as those participants who rated themselves ‘completely recovered’ or ‘much improved’.

There were some differences between the RCTs in the manner in which the outcome measures were applied. In the Australian RCT, pain severity and the blinded assessor’s rating of severity was measured on a continuous visual analogue scale (0 mm = no pain, 100 mm = worst pain imaginable); whereas in the Dutch RCT, an 11-point categorical scale (0 = no pain, 10 = worst pain imaginable) was used. The Dutch RCT used a modified Pain Free Function Questionnaire [19] in which 10 items representing common activities that may cause elbow pain were rated on a five-point Likert scale (i.e. total possible score = 40), whereas the Australian RCT used an 8-item dichomous scale (yes/no; i.e. total possible score = 8) [19]. To standardize the score for these questionnaires, the measures were transformed to 0–100 scales, where 100 indicates maximum severity [20]. The two RCTs also used different instruments and different units of measure for the evaluation of grip strength [Jamar hand dynamometer (kg) in the Dutch and digital grip dynamometer (MIE Medical Research Ltd, UK; Newtons) in the Australian], so pain-free grip strength in the affected arm was expressed as a percent of maximum grip strength of the unaffected arm; with a higher score indicating better grip strength.

Statistical analyses
The data analysis used SPSS version 13.0 [21]. Homogeneity of the study population was evaluated prior to the pooling of individual patient data, using Student’s t-tests (continuous data) and chi-squared tests (dichotomous data). The demographical characteristics of the patient population included in this analysis were age, sex, employment status (manual, non-manual, no work), duration of elbow symptoms, dominant side affected and previous elbow symptoms. The outcome measures assessed at baseline were pain severity, assessor rating of severity, pain-free grip strength, global improvement and elbow disability. Baseline outcome measures that were not significantly different between RCTs (P > 0.05) were used to evaluate the treatment effects at 6 and 52 weeks’ follow-up for the following subgroups of patients: pain severity at presentation (median split; <60 mm vs ≥ 60 mm), duration of elbow complaints (median split; ≤ 14 weeks vs > 14 weeks), previous elbow complaints (yes vs no) and employment status (manual work vs non-manual work vs no work).

Firstly, interaction between treatment groups, trial and subgroup variables was evaluated in order to determine if treatment effects differed between RCTs and whether outcomes in subgroups of patients were determined by trial and/or treatment groups (P < 0.10). Secondly, because there was no significant three-way interaction (P > 0.10) but a significant treatment by subgroup interaction (P < 0.10), differences between subgroups of patients within each treatment group were then investigated using univariate analysis of covariance for continuous outcomes and logistic regression for dichomous outcomes, measured at 6 and 52 weeks’ follow-up. Trial and baseline values for the outcome measure were retained in the model as covariates.

Differences in treatment effects between subgroups of patients were presented as mean difference and 95% confidence intervals (CI) for continuous outcomes and relative risk (RR) and numbers needed to treat (NNT) for dichotomous outcomes [22].

Results
Table 1 shows the demographic and clinical patient characteristics and outcome measures for both RCTs at baseline. The populations were similar with respect to age and previous history of elbow symptoms. However, more men were included in the Australian (65%) than in the Dutch RCT (50%). Patients in the Australian RCT had on average a longer duration of elbow complaints than the patients in the Dutch RCT (mean difference 13.2 weeks; 95% CI 7.5–18.8), but patients in the Dutch RCT reported more elbow complaints on the dominant side than in the Australian RCT (mean difference 11%; 95% CI 2–20).

Except for pain severity at baseline, all outcome values at baseline were significantly different between RCTs. Differences between baseline outcomes could not be explained by differences in patient characteristics between the two RCTs (Table 1). Therefore, the outcome measures of global improvement, which was not measured at baseline, and pain severity were used to evaluate the treatment effects in subgroups of patients.

The population characteristics did not significantly influence outcomes over time. There were no significant differences in treatment effects between RCTs for pain severity at 6 or 52 weeks, however, there was a significant difference between RCTs for global improvement at 6 weeks (P = 0.014), but not at 52 weeks. Therefore, the individual patient data from both RCTs were used to assess the combined treatment effects on pain severity at 6 and 52 weeks and for global improvement at 52 weeks.

Despite differences in patient characteristics at baseline, the overall pattern of treatment effects were very similar for both RCTs. For the combined individual patient data at 6 weeks, the corticosteroid injections group was significantly superior to both the physiotherapy and the wait-and-see policy for pain severity and global improvement scores. The physiotherapy group was also significantly superior to the wait-and-see (Figs 1 and 2, Table 2). This relative success was not present at 52 weeks, with injection significantly worse than both physiotherapy and the
wait-and-see policy in the long-term (Table 2, Figs 1 and 2). At 52 weeks, physiotherapy remained significantly better than wait-and-see for pain severity, but not global improvement (Table 2). There were significant differences in treatment effects for the subgroup of baseline pain severity at 6 weeks, with a smaller and non-significant benefit of physiotherapy over the wait-and-see group for those patients reporting baseline pain $>60$ mm (Table 2), compared with those with lower baseline pain scores who received the physiotherapy treatment. There was also a significant interaction between treatment and the employment subgroups for global improvement at 52 weeks ($P = 0.055$). Specifically, only in non-manual workers was the success rate with corticosteroid injection significantly worse when compared with wait-and-see.

Discussion

From this analysis of individual patient data from two RCTs, it was readily apparent that corticosteroid injection treatment provides superior short-term (6 weeks) results but at the cost of significantly worse outcomes than wait-and-see or physiotherapy in the long-term (1 yr). At 52 weeks, the majority of participants in both the wait-and-see and physiotherapy groups reported a successful outcome (87% wait-and-see, 92% physiotherapy). Although physiotherapy appears to accelerate the recovery over a wait-and-see policy in the short-term and remained significantly superior to the wait-and-see policy in the long-term for the measure of pain severity, the size of this effect (VAS of 10 and 6/100 mm, respectively) is of questionable clinical relevance.

The overall effects observed in the combined individual patient data from the two RCTs were not substantially altered in the subgroup analyses. That is, our study did not find that patient characteristics modified treatment effects in the short or long-term, except for small effects of baseline pain severity on pain outcome at 6 weeks and employment status on global improvement outcome at 52 weeks (Table 2). In the short-term it would appear that those patients receiving physiotherapy who report more severe pain at baseline ($\geq 60$ mm) may not respond significantly better when compared with wait-and-see and clinicians may need to focus initially more on pain relieving techniques or modalities. Only in non-manual workers were the success rates in those receiving corticosteroid injections significantly lower when compared with wait-and-see. This is difficult to explain and given the number of analyses, may be a chance finding. Alternatively, manual workers—if they take sick leave—and those without work may give their elbow more rest, or alternatively, this may reflect a protective function of those engaging in incidental physical activity after injection. Further investigation is required to evaluate these assertions.

There are dangers inherent in performing subgroup analyses in a single RCT. In particular, there is a risk of performing multiple and possibly excessive analyses in order to uncover some effect of the target intervention, albeit in a subgroup of the trial population, i.e. data dredging [15]. Also, there may be insufficient statistical power to reveal differences within a single RCT, because the trial is usually only powered to reveal differences between intervention groups. We have endeavoured to address these issues by defining in advance the factors upon which to subgroup the study population and furthermore we pooled individual patient data from two RCTs. Nevertheless, the effects we report are small and should be interpreted with care.

Despite the clinical heterogeneity between the two RCT populations, patterns of treatment effects over time were very similar between RCTs. This suggests that the treatment effects are robust across an international population of patients with tennis elbow, which is a similar finding to that of Smidt et al. [20], who compared the Dutch RCT data with that from the one in the UK.

A potential source of heterogeneity between the two RCTs was the difference in recruitment of the patient populations. The Dutch RCT recruited patients who had visited a general medical practitioner for their elbow condition, while the Australian RCT recruited patients from the general community. Interestingly, Hudak et al. [13] recommended a wider spectrum of patients than just those referred from general medical practice be used in studies of tennis elbow so as to improve the information gained on clinical course and prognosis of outcome. The influence of source of participants on baseline population characteristics and treatment outcomes should be further assessed.

From the combined individual patient data of two RCTs comparing physiotherapy, corticosteroid injections and wait-and-see policy interventions, in the main, it appears that subgroups of patients determined by baseline patient characteristics do not differ on outcomes over time. However, baseline severity of pain appears to modify the effects of physiotherapy in the short-term and clinicians may expect weaker effects for physiotherapy treatment in the short-term in patients reporting high pain scores on initial assessment. Pooling of individual patient data was possible despite population heterogeneity between RCTs. This supports the generalisability of individual RCT results.
**Table 2. Subgroup results and between-group comparisons in the pooled data set of two randomized controlled trials on conservative treatment of tennis elbow**

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Subgroups</th>
<th>Wait and see</th>
<th>Corticosteroid injection</th>
<th>Physiotherapy</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>PVAS 6 Weeks</td>
<td>Pooled</td>
<td>44.1</td>
<td>(26.1)</td>
<td>13.3</td>
<td>(18.9)</td>
</tr>
<tr>
<td></td>
<td>&lt;60 mm</td>
<td>35.3</td>
<td>(21.6)</td>
<td>9.8</td>
<td>(12.0)</td>
</tr>
<tr>
<td></td>
<td>&gt;60 mm</td>
<td>52.0</td>
<td>(27.4)</td>
<td>16.8</td>
<td>(23.5)</td>
</tr>
<tr>
<td>Duration</td>
<td>≤14 weeks</td>
<td>40.4</td>
<td>(24.4)</td>
<td>9.7</td>
<td>(13.5)</td>
</tr>
<tr>
<td></td>
<td>&gt;14 weeks</td>
<td>47.5</td>
<td>(27.4)</td>
<td>16.6</td>
<td>(22.7)</td>
</tr>
<tr>
<td>Previous history</td>
<td>No (n = 263)</td>
<td>43.1</td>
<td>(25.2)</td>
<td>12.3</td>
<td>(16.6)</td>
</tr>
<tr>
<td></td>
<td>Yes (n = 110)</td>
<td>46.3</td>
<td>(28.2)</td>
<td>15.1</td>
<td>(22.7)</td>
</tr>
<tr>
<td>Employment status</td>
<td>Manual (n = 108)</td>
<td>47.1</td>
<td>(26.7)</td>
<td>15.3</td>
<td>(21.4)</td>
</tr>
<tr>
<td></td>
<td>Non-manual (n = 172)</td>
<td>42.4</td>
<td>(25.6)</td>
<td>11.1</td>
<td>(16.4)</td>
</tr>
<tr>
<td></td>
<td>No work (n = 93)</td>
<td>43.6</td>
<td>(26.9)</td>
<td>14.7</td>
<td>(19.9)</td>
</tr>
<tr>
<td>PVAS 52 Weeks</td>
<td>Pooled</td>
<td>13.5</td>
<td>(21.5)</td>
<td>13.2</td>
<td>(26.9)</td>
</tr>
<tr>
<td></td>
<td>&lt;60 mm</td>
<td>10.7</td>
<td>(19.4)</td>
<td>16.1</td>
<td>(22.5)</td>
</tr>
<tr>
<td>Duration</td>
<td>≤14 weeks</td>
<td>16.0</td>
<td>(23.2)</td>
<td>26.2</td>
<td>(30.1)</td>
</tr>
<tr>
<td></td>
<td>&gt;14 weeks</td>
<td>12.0</td>
<td>(18.7)</td>
<td>20.4</td>
<td>(26.7)</td>
</tr>
<tr>
<td>Previous history</td>
<td>No (n = 261)</td>
<td>12.9</td>
<td>(20.1)</td>
<td>19.1</td>
<td>(26.6)</td>
</tr>
<tr>
<td></td>
<td>Yes (n = 112)</td>
<td>14.7</td>
<td>(24.6)</td>
<td>25.2</td>
<td>(27.3)</td>
</tr>
<tr>
<td>Employment status</td>
<td>Manual (n = 108)</td>
<td>13.7</td>
<td>(23.8)</td>
<td>18.7</td>
<td>(28.2)</td>
</tr>
<tr>
<td></td>
<td>Non-manual (n = 172)</td>
<td>11.2</td>
<td>(17.5)</td>
<td>21.7</td>
<td>(25.2)</td>
</tr>
<tr>
<td></td>
<td>No work (n = 93)</td>
<td>17.6</td>
<td>(25.0)</td>
<td>22.9</td>
<td>(28.8)</td>
</tr>
</tbody>
</table>

Group means (s.d.) and between-group mean differences (95% CI) for pain (PVAS); and proportions (%) and between-group relative risks (95% CI) for global improvement (GI). PVAS adjusted for baseline pain score and trial and global improvement adjusted for trial. Pooled results: *in favour of corticosteroid injection, **in favour of physiotherapy, ***significant differences between treatments for subgroups, ****in favour of wait-and-see, P < 0.05.

PT, physiotherapy; Injection, corticosteroid injection.
Acknowledgements

We wish to acknowledge Dr Ross Darnell for his assistance with the statistical analyses. Financial support was received from The University of Queensland and the National Health and Medical Research Council, Primary Health Care Project Grant, Australia #252710 and the University of Queensland Research Travel Grant.

The authors have declared no conflicts of interest.

References


Rheumatology key messages

- Patient characteristics only minimally influence the outcomes of different treatments.
- The clinician can rely on results of individual trials to predict treatment outcome.

FIG. 1. Pain severity for physiotherapy, corticosteroid injection and wait-and-see policy in the pooled data set of two randomized controlled trials on conservative treatment of tennis elbow.

FIG. 2. Global improvement (success) for physiotherapy, corticosteroid and wait-and-see policy in the pooled data set of two randomized controlled trials on conservative treatment of tennis elbow.