Efficacy of physiotherapy for musculoskeletal disorders: what can we learn from research?

H BECKERMAN
L M BOUTER
G J M G VAN DER HEIJDEN
R A DE BIE
B W KOES

SUMMARY: In order to summarize the available clinical evidence for the efficacy of physiotherapy, 400 randomized clinical trials were identified from the literature. Studies were found by using bibliographic databases, citation tracking, and correspondence with researchers in the field. Focusing on disorders of the musculoskeletal system, a number of criterion-based meta-analyses were performed on 180 trials in order to summarize the available evidence. For each randomized clinical trial in each meta-analysis a methodological score was calculated using a set of explicit criteria and weighting factors applied by two or three independent reviewers who were blinded as to the outcomes, the journal and the authors of the publication. In each meta-analysis the randomized clinical trials were ordered hierarchically depending on their score for methodological quality. Meta-analyses were performed for spinal manipulation, exercise, traction, ultrasound, and laser therapy, and for disorders of the back, neck, shoulder and knee. In general, the methodological quality of the studies appeared to be low, and the efficacy of physiotherapy was shown to be convincing only for a few indications and treatments. On the other hand, because of the prevalence of serious methodological flaws, it cannot be concluded that physiotherapy has no effect.

Keywords: physiotherapy; musculoskeletal disorders; clinical trials; outcome.

Method

Introduction

HIGH quality health care is the goal of the medical profession and the expectation of patients. Insurance companies and policy makers increasingly require documentation of effectiveness of treatment. Critical evaluation of treatment techniques is needed for further development of professional skills. In the Netherlands, during the last decade there has been a rapid increase in the number of physiotherapists, the overall number of physiotherapy treatment sessions, and, consequently, in health care costs. This development prompted the Dutch government to think critically about the efficacy of isolated physiotherapeutic applications and physiotherapy as a whole. In order to gain more insight into these questions a literature review was carried out.

Methodological assessment criteria

To explore the possibility that methodological errors can seriously bias the results of the randomized clinical trials, criteria for a methodological assessment of the studies were established. Standards for the methodological quality of randomized clinical trials have been formulated by several authors. A list of criteria was used, covering the study population, the description of the therapeutic regimens, the measurement of effect, and the data analysis and presentation of results. As an example, Table 1 shows the criteria and weighting factors for the methodological assessment of randomized clinical trials dealing with the efficacy of physiotherapy for disorders of the knee (Beckerman H, et al. The efficacy of physical therapy for disorders of the knee: a criteria-based meta-analysis of 63 randomized clinical trials. Unpublished results). A breakdown of the scoring system for the criteria is shown in Appendix 1.
Table 1. Criteria for assessing the methodological quality of randomized clinical trials of physiotherapy for disorders of the knee.*

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study population</td>
<td></td>
</tr>
<tr>
<td>A Selection</td>
<td>4</td>
</tr>
<tr>
<td>B Randomization procedure</td>
<td>8</td>
</tr>
<tr>
<td>C Sample size</td>
<td>12</td>
</tr>
<tr>
<td>D Baseline characteristics</td>
<td>10</td>
</tr>
<tr>
<td>E Drop outs</td>
<td>3</td>
</tr>
<tr>
<td>F Number of drop outs</td>
<td>4</td>
</tr>
<tr>
<td>Description of therapeutic regimen</td>
<td></td>
</tr>
<tr>
<td>G Experimental intervention</td>
<td>5</td>
</tr>
<tr>
<td>H Control intervention(s)</td>
<td>5</td>
</tr>
<tr>
<td>I Co-interventions</td>
<td>6</td>
</tr>
<tr>
<td>Effect measurement</td>
<td></td>
</tr>
<tr>
<td>J Blinding of patient</td>
<td>5</td>
</tr>
<tr>
<td>K Blinding of therapist</td>
<td>5</td>
</tr>
<tr>
<td>L Outcome measures</td>
<td>5</td>
</tr>
<tr>
<td>M Blinded assessments of outcome</td>
<td>5</td>
</tr>
<tr>
<td>N Side effects</td>
<td>3</td>
</tr>
<tr>
<td>O Follow-up period</td>
<td>5</td>
</tr>
<tr>
<td>Data analysis and presentation of results</td>
<td></td>
</tr>
<tr>
<td>P Statistical analysis</td>
<td>8</td>
</tr>
<tr>
<td>Q Presentation of the results</td>
<td>5</td>
</tr>
<tr>
<td>R Power of the study</td>
<td>2</td>
</tr>
</tbody>
</table>

*Beckerman H, et al, unpublished results. aFurther details are given in Appendix 1.

Depending on the subject of the meta-analysis at issue, the criteria and weighting factors were adjusted to some extent. Each study could earn a maximum score of 100 points. Two or three reviewers, independent of each other, applied these methodological criteria to each trial. In total five reviewers, who were specialists in physiotherapy as well as research methodology, were involved in the nine meta-analyses. In most meta-analyses the reviewers were blinded as to author(s), journal and outcome of the studies. The outcome of the study was not considered relevant for the assessment of the methodological quality. However, as soon as methodological quality had been assessed, the results were unblinded and summarized. Differences between the reviewers were resolved by discussion during a consensus meeting and a further clarification of the criteria. The final score for each study is based on the full consensus between the reviewers.

These criterion-based meta-analyses resulted in a hierarchical list in which the higher scores indicate studies of a higher quality. Consequently, the outcomes of the studies can be discussed in relation to these scores.

**Results**

More than 400 randomized clinical trials eligible for the review were identified.11 Most studies (76%) were published after 1980. The oldest study was published in 1955.12 Of these 400 studies, 280 dealt with patients with musculoskeletal disorders. Of these studies 180 were included in the nine criterion based meta-analyses completed (the references to these randomized clinical trials are available from the first author on request). These meta-analyses covered spinal manipulations, exercise therapy, traction, ultrasound and laser therapy for disorders of the back, neck, shoulder and knee. The results of the methodological assessments are shown in Table 2. The main conclusions are summarized in the following paragraphs. For more details the reader is referred to the original review articles (Beckerman H, et al, unpublished results and van der Heijden GJMG, et al, traction for back and neck pain, a blinded review, unpublished results).13-17

Table 2. Meta-analyses of trials of the efficacy of physiotherapy for musculoskeletal disorders.

<table>
<thead>
<tr>
<th>Indication</th>
<th>Intervention</th>
<th>Number of RCTs*</th>
<th>Methodological median score (range)d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back pain13,14 c</td>
<td>Spinal manipulation</td>
<td>30</td>
<td>35 (20-56)</td>
</tr>
<tr>
<td></td>
<td>Exercise therapy</td>
<td>16</td>
<td>40 (24-61)</td>
</tr>
<tr>
<td></td>
<td>Traction</td>
<td>16</td>
<td>36 (23-66)</td>
</tr>
<tr>
<td>Neck pain13 c</td>
<td>Spinal manipulation</td>
<td>5</td>
<td>39 (26-50)</td>
</tr>
<tr>
<td></td>
<td>Traction</td>
<td>3</td>
<td>39 (36-51)</td>
</tr>
<tr>
<td>Shoulder complaints15</td>
<td>Physiotherapyd</td>
<td>18</td>
<td>49 (22-76)</td>
</tr>
<tr>
<td>Knee disorders16,17 d</td>
<td>Physiotherapyd</td>
<td>63</td>
<td>29 (6-52)</td>
</tr>
<tr>
<td>Musculoskeletal disorders16,17 d</td>
<td>Ultrasound therapy</td>
<td>16</td>
<td>41 (17-70)</td>
</tr>
<tr>
<td></td>
<td>Laser therapy</td>
<td>33</td>
<td>40 (4-72)</td>
</tr>
</tbody>
</table>

RCT = randomized clinical trial. *In total, 180 trials were used in one or more meta-analyses. dMaximum score 100 points. eBeckerman H, et al, unpublished results. fSee original publication for more details. gvan der Heijden GJMG, et al, unpublished results. hPercentage of the total score (maximum 25 points).

Table 3. Sample size of explanatory trials and pragmatic trials.

<table>
<thead>
<tr>
<th>Type of study</th>
<th>No. of studiesa</th>
<th>Median (25% and 75% percentiles)</th>
<th>No. of patients in smallest group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory</td>
<td>67</td>
<td>15 (12 and 33)</td>
<td></td>
</tr>
<tr>
<td>Pragmatic</td>
<td>107</td>
<td>20 (10 and 31)</td>
<td></td>
</tr>
</tbody>
</table>

aSix missing values.

**Back and neck pain**

Patients with back and neck complaints are often treated by physiotherapists. About 40% of the Dutch population visiting physiotherapists are patients with these complaints.18 General practitioners, who in the Netherlands are mostly responsible for the referral of patients, strongly believe in the efficacy of manipulation and physiotherapy.19 A survey among 293 general practitioners in the Netherlands showed that 80% of them believe that manual therapy is efficacious in the treatment of patients with chronic neck and back problems, compared with 71% for conventional physiotherapy.20

However, the opinions of general practitioners are not supported by the literature. Although there are some promising results from randomized clinical trials in favour of manipulation for back and neck pain, the poor methodological quality of the 35 trials severely limits the strength of the conclusions that can be drawn.13 No single trial scored more than 56 points. Among the studies with more than 50 points, one was positive, one was negative and two were positive in a subgroup with low back pain of two to four weeks duration. Eight trials attempted to compare manipulation with placebo therapy. In most studies the placebo therapy consisted of detuned short wave diathermy. The results were inconsistent in four of these trials manipulation was more effective than placebo therapy, in three other studies this was not the case, and in one trial manipulation was more effective in a subgroup of sufferers of acute pain only.

Overall, 18 studies found positive effects of manipulation for acute or chronic back or neck pain, and 11 studies had a negative
outcome. In addition, five studies reported positive results in one or more subgroups only. One author refrained from drawing a conclusion. In general, the negative studies seemed to have higher methodological scores. Only 14 studies measured the effect at least three months after randomization, four of which reported long term positive effects of manipulation.

To determine the efficacy of exercise therapy 16 trials, in which exercise therapy was given by physiotherapists to individual patients with back pain, were identified.¹⁴ The results of seven studies which contrasted exercise therapy with other conservative treatments are controversial. Two of these seven studies indicate favourable outcomes of exercise therapy compared with hot packs and rest, or one session of 45 minutes with instructions and education. In the latter study the differences between the groups were still present one year after randomization. Nevertheless, the five other studies indicated that exercise therapy was no better than manipulation and mobilization, hot packs, education, (bed)rest, short wave diathermy, and non-steroid anti-inflammatory drugs. All seven studies, however, had a relatively low methodological score of less than 50 points.

Three studies contrasted exercise therapy with no exercise therapy; all used a factorial design. Of these, the study with the highest methodological score found that exercises had a positive effect for patients with chronic low back pain. After three months follow up this difference had disappeared. The two other studies with relatively high methodological scores found no positive effect. There was one placebo controlled study, which found no differences in effect.

Of the 16 randomized clinical trials, eight studies compared different exercise programmes. The comparisons are mainly between isometric flexion exercises and extension exercises. Two studies suggested that extension exercises were better, whereas two other studies suggested that a flexion programme was superior. The other four studies reported no differences.

Manual or mechanical traction is another therapeutic possibility for patients with low back and neck pain. The mechanisms by which traction could be effective are spinal elongation by means of decreasing lumbar or cervical lordosis and increasing the intervertebral space. Existing muscle strain or spinal nerve root compression could be reduced, and a temporary laxation of a zygo-apophysical disc or joint capsule can theoretically be released. The clinical effects were studied on the basis of 16 trials among patients with low back pain and three trials among patients with neck complaints (van der Heijden GJMG, et al, unpublished results). Three of the 16 trials among patients with low back pain scored more than 50 points. All three studies reported no differences in effect between the treatment groups. One of the trials dealing with the efficacy of cervical traction scored more than 50 points. All three trials showed a negative outcome.

**Shoulder complaints**

Of all patients treated by physiotherapists working in Dutch primary health care, approximately 10% are treated because of shoulder complaints.¹⁸ The 18 randomized controlled trials identified do not allow strong conclusions about the efficacy of physiotherapy for shoulder complaints.¹⁹ Patients were, for instance, treated with ultrasound, laser therapy, short wave diathermy and exercise therapy. Control groups were usually treated with placebo therapy, steroid injections and other medications. Only nine studies scored 50 or more points. Among the better studies only one had results confirming the efficacy of a physiotherapeutic treatment (pulsed electromagnetic field therapy) compared with placebo therapy.

**Disorders of the knee**

In order to investigate the current state of knowledge about the efficacy of physiotherapy for disorders of the knee 63 trials were reviewed (Beckerman H, et al, unpublished results). The study populations consisted of patients with anterior cruciate ligament injuries, meniscectomy, gonarthrosis, total knee replacement, and rheumatoid arthritis of the knee. For each diagnostic group a variety of treatment possibilities were investigated for their efficacy. In general, the methodological quality of the studies appeared to be low. The median methodological score measured 29 (Table 2). Thirty one trials showed a positive result, whereas in the other trials no positive effect could be detected. It is difficult to draw a general overall conclusion concerning the efficacy of physiotherapy (such as exercises, continuous passive motion, electrical stimulation) for disorders of the knee based on the results of this meta-analysis. As indicated by the low median methodological score, the results of the studies may be seriously biased because of methodological shortcomings.

**Ultrasound therapy**

Ultrasound, at least in the Netherlands, is one of the most frequently used physiotherapeutic applications.† Based on empirical evidence it is proposed that ultrasound enhances recovery from musculoskeletal disorders as a result of its mechanical, thermal and chemical effects. Sixteen trials were included in the meta-analysis. The explanatory trials as well as the pragmatic trials showed no convincing evidence confirming the efficacy of ultrasound. The only exception is possibly the existence of clinical effects for elbow disorders.

**Laser therapy**

Recently, laser therapy has become a popular treatment in the physiotherapeutic management of musculoskeletal disorders. The efficacy of laser therapy has been assessed on the basis of the results of 33 trials among 1592 patients.²⁰ The positive studies were generally of a better quality than the studies with a negative outcome. No clear relationship could be demonstrated between the dose applied and the efficacy of laser therapy, or between the dose and the methodological score. Therefore, no conclusions can be drawn concerning the minimal and optimal dose of laser therapy. In general, the efficacy of laser therapy for musculoskeletal disorders seems to be greater than the efficacy of a placebo treatment. For rheumatoid arthritis, post-traumatic joint disorders, and myofascial pain, laser therapy seems to have a substantial specific therapeutic effect.

**Discussion**

The trials included in the meta-analyses can be considered to be the best studies available evaluating the efficacy of physiotherapy. The potential of randomized clinical trials to supply valid answers is much larger than that of uncontrolled or non-randomized studies. Although there is a surprisingly large amount of research, most randomized clinical trials investigating the efficacy of physiotherapy have major flaws in their methodology. The meta-analyses suggest that the predominantly disappointing results could be caused by the heterogeneity of the study population, the interventions and the effect measures. Furthermore, the considerable number of negative outcomes is probably due to small sample sizes.

The diagnoses of patients referred for physiotherapeutic treatment are often vague and refer to symptoms or to the localization of complaints, for instance lumbago or frozen shoulder. These diagnostic groups probably cover a wide range of diagnoses. To identify homogeneous groups of patients that are susceptible to the intervention under study, more adequate
diagnostic procedures are necessary. At the moment diagnostic procedures seem to contribute only in a global sense to the choice of physiotherapeutic treatment.\(^{20,21}\) In order to define homogeneous groups of patients, new diagnostic strategies are needed to discover factors which are prognostically relevant. Furthermore, in future randomized clinical trials, more attention needs to be paid to the explicit description of the inclusion and exclusion criteria of the patients and the selection of a prognostically homogeneous study population. Trials with prognostically homogeneous groups would give a better idea of the real value of a therapeutic intervention.

In the publications the interventions were often incompletely described. Therefore, it was generally impossible to identify the optimal dose of the investigated intervention.\(^{17}\) The great variety of methods of carrying out the same intervention in several studies might explain some of the inconsistent outcomes of studies with the same research question. Some methods might be ineffective, for instance owing to under- or over-dosage, whereas others are effective. Future publications ought to contain complete and exact information about the execution of the interventions.

One of the main problems in the evaluation of the effects of physiotherapy is the relative lack of outcome parameters that are valid, precise, sensitive and clinically relevant. Regarding clinical relevance both important complaints presented by the category of patients under study and sensitivity for changes that are considered to be of clinical relevance should be included.\(^{22,23}\) In many studies, there are strong doubts about the investigators used to measure the efficacy of physiotherapy. Consequently, the study results concerning the efficacy of physiotherapy could be biased substantially. More research into the methodological quality of measurement instruments to assess clinically relevant changes will benefit future evaluations in the field of physiotherapy.

Before concluding that there is probably no difference in effect, one should always take the power of the trials into account.\(^{8}\) The randomized clinical trials included in the meta-analyses described here had very small sample sizes. With small sample sizes there is usually a substantial chance of false negative conclusions occurring. In other words, one does not find a statistically significant difference between the treatment groups under contrast, while in fact there is such a difference. This problem could be avoided if sample sizes in physiotherapy trials were based on more realistic assumptions about the expected difference in effect. That is, of course, clinically relevant differences.\(^{24}\) These assumptions can be based on earlier publications, pilot studies or consensus meetings among experts.

The efficacy of physiotherapeutic treatments for musculoskeletal disorders has been investigated for many interventions and indications. Although some forms of physiotherapy for some indications seem to be promising, no definite conclusions can be drawn owing to the general low methodological quality of the studies. The most prevalent methodological shortcomings of randomized clinical trials in the field of physiotherapy consisted of problems with the eligibility criteria of the study population, the way in which the interventions were carried out, the lack of responsive effect measurements and false negative conclusions. Physiotherapy may well have a beneficial effect, but because of the methodological shortcomings of the trials studied, this conclusion cannot be reached.

Appendix 1. Scoring for criteria listed in Table 1.

A Description of inclusion and exclusion criteria: two points.

B Randomization procedure described: three points. Randomization procedure which excludes bias: five points.

C Size of smallest group immediately after randomization — more than 25 patients: four points, more than 50 patients: four points, more than 75 patients: four points.

D Comparability for age, duration of complaints, severity of complaints, baseline value of outcome measures, and surgery: two points each.

E Information about from which group subjects dropped out with reason for withdrawal: one point. Non-selective dropout: two points. No dropouts: three points.

F Loss to follow up: all randomized patients minus the number of patients assessed for the main outcome measure (according to the author), divided by all randomized patients multiplied by 100 — less than 20%: two points, less than 10%: two points.

G Physiotherapy explicitly described: method, intensity, duration of each treatment, total number of treatments, treatment frequency: one point each.

H All reference treatments (placebo or real) described — method, intensity, duration of each treatment, total number of treatments, treatment frequency: one point each.

I Other medical interventions: two points. Self medication avoided or comparable: three points.

J Attempted blinding of patients: three points. Blinding evaluated and successful: two points.

K Attempted blinding of therapist: three points. Blinding evaluated and successful: two points.

L Outcome measures assessed and reported: pain, strength, mobility, stability, and functional status (activities of daily living): one point each.

M Each blinded measurement mentioned under criterion L earns one point.

N Side effects described: one point, for each group separately: two points.

O Outcome measures assessed during or just after last treatment: two points. Outcome measures assessed after six months or longer: three points.


Q Frequency of the outcome measures mentioned under criterion L at the most important times of measurement. In the case of (semi) continuous variables presentation of the mean or median with a standard error or percentiles: one point each.

R Statement concerning the adequacy of sample size (power of the study): two points.

References


Acknowledgements
The authors thank WJJ Assendelft and RJD Wilkinson. This study was supported by a grant from the Dutch Ministry of Welfare, Public Health, and Cultural Affairs (project 89-42), and took place at the Department of Epidemiology and Biostatistics, University of Limburg, Maastricht, the Netherlands.

Address for correspondence
Ms H Beckerman, Department of Rehabilitation Medicine, Free University Hospital, PO Box 7057, 1007 MB Amsterdam, The Netherlands.

RCPG Publications
RESEARCH AND AUDIT

Epidemiology in Country Practice
William Pickles' own work — first published in 1939 — the classic example of original research in general practice. '... an inspiration for us today' New Zealand Family Physician. £5.50

The Measurement of the Quality of General Practitioner Care (Occasional Paper 15)
A detailed review of the literature on quality of care. Essential reading for those trying to measure and promote quality in general practice. £3.00

Practice Audit Plan
The aim of audit is to maintain and improve the quality of care, so the new proposals present general practitioners with a major opportunity to advance the care given to patients. £3.50

Practice Activity Analysis (Occasional Paper 41)
Describes the results of practice activity analysis undertaken by the RCPG Birmingham Research Unit over a number of years and discusses its role as a practical approach to audit and assessment of quality. £7.50

Who Killed Susan Thompson? (Video)
Set in a fictional four doctor partnership in an English market town, this video demonstrates the link between individual clinical competence and medical audit. £30.00 (Additional Course Books £5.00)

All the above can be obtained from the Sales Office, Royal College of General Practitioners, 14 Princes Gate, London SW7 1PU (Enquiries, Tel: 071-823 9698). Prices include postage. Cheques should be made payable to RCPG Enterprises Ltd. Access and Visa welcome (Tel: 071-225 3048, 24 hours).

RCPG EXAMINATION — 1993
The dates and venues of the next two examinations are as follows: May/July 1993

Written papers: Wednesday 5 May 1993 at centres in London, Manchester, Edinburgh, Newcastle, Cardiff, Belfast, Dublin, Liverpool, Ripon, Birmingham, Bristol and Sennelager.

Oral examinations: In Edinburgh from Monday 21 to Wednesday 23 June and in London from Thursday 24 June to Saturday 3 July inclusive.

The closing date for the receipt of applications is Friday 26 February 1993.

October/December 1993

Written papers: Tuesday 26 October 1993 at those centres listed above.

Oral examinations: In Edinburgh on Monday 6 and Tuesday 7 December and in London from Wednesday 8 to Monday 13 December inclusive.

The closing date for the receipt of applications is Friday 3 September 1993.

MRCGP is an additional registrable qualification and provides evidence of competence in child health surveillance for accreditation.

For further information and an application form please write to the Examination Department, Royal College of General Practitioners, 14 Princes Gate, London SW7 1PU, or telephone: 071-581 3232.