Effectiveness of exercise therapy: A best-evidence summary of systematic reviews

Nynke Smidt, Henrica CW de Vet, Lex M Bouter and Joost Dekker
for the Exercise Therapy Group

The purpose of this project was to summarise the available evidence on the effectiveness of exercise therapy for patients with disorders of the musculoskeletal, nervous, respiratory, and cardiovascular systems. Systematic reviews were identified by means of a comprehensive search strategy in 11 bibliographic databases (08/2002), in combination with reference tracking. Reviews that included (i) at least one randomised controlled trial investigating the effectiveness of exercise therapy, (ii) clinically relevant outcome measures, and (iii) full text written in English, German or Dutch, were selected by two reviewers. Thirteen independent and blinded reviewers participated in the selection, quality assessment and data-extraction of the systematic reviews. Conclusions about the effectiveness of exercise therapy were based on the results presented in reasonable or good quality systematic reviews (quality score ≥ 60 out of 100 points). A total of 104 systematic reviews were selected, 45 of which were of reasonable or good quality. Exercise therapy is effective for patients with knee osteoarthritis, sub-acute (6 to 12 weeks) and chronic (≥ 12 weeks) low back pain, cystic fibrosis, chronic obstructive pulmonary disease, and intermittent claudication. Furthermore, there are indications that exercise therapy is effective for patients with ankylosing spondylitis, hip osteoarthritis, Parkinson's disease, and for patients who have suffered a stroke. There is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with neck pain, shoulder pain, repetitive strain injury, rheumatoid arthritis, asthma, and bronchiectasis. Exercise therapy is not effective for patients with acute low back pain. It is concluded that exercise therapy is effective for a wide range of chronic disorders. [Smidt N, de Vet HCW, Bouter LM and Dekker J (2005): Effectiveness of exercise therapy: A best-evidence summary of systematic reviews. Australian Journal of Physiotherapy 51: 71–85]

Key words: Exercise Therapy; Exercise Movement Techniques; Meta-analysis; Physical Therapy Techniques

Introduction

Exercise therapy is a regular component in the management of various (chronic) disorders, such as musculoskeletal, neurological, cardiovascular, and respiratory disorders (Chartered Society of Physiotherapy 2001, ACCP/AACVPR Pulmonary Rehabilitation Guidelines Panel, American College of Chest Physicians, American Association of Cardiovascular and Pulmonary Rehabilitation 1997, Gordon et al 2004, Pina et al 2003, Wooff et al 2004). Exercise therapy involves the prescription of muscular contraction and bodily movement ultimately to improve the overall function of the individual and to help meet the demands of daily living (Tan and Horn 1998).

There is no up-to-date overview of the effectiveness of exercise therapy compared with no treatment or an alternative treatment (Beckerman et al 1993a, Beckerman et al 1993b, Bouter et al 1992, Herbert et al 2001). An overview of this will help: care providers to choose the most appropriate treatment option; policy makers in making decisions concerning health care; and research agencies in setting priorities in the field of physiotherapy. Our objective was therefore to assess and summarise the available evidence on the effects of exercise therapy in a best-evidence summary of systematic reviews.

Method


Selection We included systematic reviews that met the following criteria: (i) the full text of the systematic review is published and it is based on a transparent and reproducible protocol (at least reporting on inclusion criteria, search date(s), and database(s)); (ii) at least one randomised controlled trial is included in the review; (iii) exercise therapy is compared with no treatment, other conservative types of treatment (e.g. steroid injections), surgery, or some other type of exercise therapy (e.g. flexion versus extension exercises); (iv) at least one clinically relevant outcome measure is included (e.g. pain, activities of daily living (ADL), walking distance, return to work) is included; (v) the results and conclusions are presented separately for each diagnosis; (vi) reviews are written in English, German or Dutch; (vii) the focus is on patients with disorders of the following: the musculoskeletal system and connective tissue, the nervous system and sense organs, the respiratory system, and the cardiovascular system (excluding coronary heart diseases),
Figure 1. Selection of systematic reviews.

Computerised searches
- PubMed (n = 1544)
- Cochrane Library (n = 270)
- PEDro (n = 425)
- Doc-Online NPi (n = 56)
- EMBASE (n = 858)
- CINAHL (n = 833)
- Current Contents (n = 10)
- Biological Abstracts (n = 10)
- Elsevier Biobase (n = 4)
- PASCAL (n = 6)
- MEDLINE in PROCESS (n=1)

4017 titles with abstracts were screened by one reviewer

228 full text reviews retrieved and read by two independent and blinded reviewers

Selected for the review (n = 104, including 9 duplicates)*

1 Respiratory system (n = 18)
   - Cystic fibrosis (n = 2/3)†
   - Chronic obstructive pulmonary diseases (n = 5/11)‡
   - Asthma (n = 2/4)‡
   - Bronchiectasis (n = 1/1)

2 Nervous system and sense organs (n = 9)
   - Parkinson's disease (n = 3/6)
   - Multiple sclerosis (n = 0/1)
   - Cerebral palsy (n = 0/1)
   - Tetraplegia (n = 0/1)

3 Cardiovascular system (n = 18)
   - Cerebrovascular accident (n = 3/11)
   - Intermittent claudication (n = 4/7)

4 Musculoskeletal system and connective tissue (n = 50)
   - Low back pain (n = 10/21)§
   - Neck pain (n = 6/6)§
   - Shoulder pain (n = 3/4)
   - Hip/knee osteoarthritis (n = 3/7)
   - Ankylosing spondylitis (n = 1/2)
   - Rheumatoid arthritis (n = 1/2)
   - Fibromyalgia (n = 0/1)
   - Patellofemoral pain syndrome (n = 0/2)
   - Repetitive strain injury (n = 1/1)
   - Carpal tunnel syndrome (n = 0/1)
   - Temporomandibular joint disc displacement (n = 0/1)
   - Mix of musculoskeletal disorders (n = 1/3)

Excluded articles (n = 3789)
- no systematic review (n = 63)
- no RCTs included in SR (n = 11)
- no exercise therapy (n = 41)
- no clinically relevant outcome measures (n = 2)
- data not presented for each diagnosis (n = 1)
- systematic review not written in English, German, or Dutch (n = 1)
- disorder not topic of this review (n = 5)

Excluded (n = 124)
- no systematic review (n = 63)
- no RCTs included in SR (n = 11)
- no exercise therapy (n = 41)
- no clinically relevant outcome measures (n = 2)
- data not presented for each diagnosis (n = 1)
- systematic review not written in English, German, or Dutch (n = 1)
- disorder not topic of this review (n = 5)

†Two of the 3 systematic reviews on the effectiveness of exercise therapy for cystic fibrosis had quality scores of at least 60 points out of 100. §In one review (<60 points) exercise therapy was investigated for both asthma and COPD therefore the sum of the individual sub-categories add up to more than the total category; §§In one review (< 60 points) exercise therapy was investigated for both neck and low back pain therefore the sum of the individual sub-categories add up to more than the total category.
The quality of the systematic reviews was assessed according to the list of criteria developed by Assendelft et al (1995). This list consists of criteria for the selection of studies (30 points), assessment of the methodological quality of randomised controlled trials (20 points), description of the interventions (15 points), data presentation (20 points), and evaluation (15 points) (see Appendix I). The maximum quality score is 100 points. A total of 13 independent, blinded reviewers (see authors’ affiliations) participated in the final selection and assessment of the quality of the systematic reviews. One reviewer (NS) assessed all systematic reviews and 12 other reviewers (MEB, SMAB, AH, SHJK, GK, TL, RPSP, MR, CT, CBT, APV, DAWMW) each evaluated a selection of the included reviews. Disagreements were discussed and resolved in a consensus meeting. If consensus could not be reached, a third reviewer (RWJGO) made the final decision.

The systematic reviews were categorised according their quality score: good quality (≥ 80 points), reasonable quality (60–79 points), moderate quality (40–59 points), poor quality (20–39 points), and very poor quality (< 20 points). Our conclusions regarding the effectiveness of exercise therapy are based on the results of reasonable quality (60–79 points) or good quality (≥ 80 points) systematic reviews (De Vet et al 2001).

Data extraction An overview of each systematic review (≥ 60 points) was made, including the research question(s) and details of all the randomised controlled trials investigating exercise therapy included in the systematic review (interventions in the experimental and control group, methodological quality, sample size (statistical power), outcome measures, timing of outcome assessment, and effectiveness of the exercise therapy (statistical significance)).

The conclusions reported in each systematic review were discussed with a panel of experts in the field of physiotherapy, general practice, rehabilitation medicine, and epidemiology (JHA, RAB, JD, PJMH, RABO, ST, HCWV). For each systematic review, categorisation of the conclusions was based on the following two research questions:

A What is the effectiveness of exercise therapy, compared to no treatment, a placebo, or a wait-and-see policy?

B What is the effectiveness of exercise therapy, compared to other treatments (e.g. steroid injections)? Is one specific type of exercise therapy more effective than others?

The following are all the possible conclusions that could be drawn for Question A:

- Exercise therapy is effective, compared to no treatment, placebo, or a wait-and-see policy (positive).
- Exercise therapy is not effective, compared to no treatment, placebo, or a wait-and-see policy (negative).
- Exercise therapy is less effective than no treatment, placebo, or a wait-and-see policy (harmful).
- There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment, placebo, or a wait-and-see policy (insufficient evidence).
- There is insufficient evidence, but there are indications to support the effectiveness of exercise therapy, compared to no treatment, placebo, or a wait-and-see policy (insufficient evidence but indications).

The following are all the possible conclusions that could be drawn for Question B:

- Exercise therapy is effective, compared to other treatments (positive).
- Exercise therapy is equally effective, compared to other treatments (equal).
- Exercise therapy is less effective, compared to other treatments (negative).
- There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to other treatments (insufficient evidence).
- There is insufficient evidence, but there are indications to support the effectiveness of exercise therapy, compared to other treatments (insufficient evidence but indications).

If the panel felt that the conclusions were not sufficiently justified by the data presented in the systematic review at issue, the conclusions reported in the systematic review were not endorsed, and the panel drew its own conclusions about the effectiveness of exercise therapy. In such cases, the panel’s conclusions were based on randomised controlled trials that were of good methodological quality (≥ 50% of the quality score reported in the systematic review) with large sample sizes (smallest group n ≥ 50).

For each disorder, the panel’s final conclusions with regard to the effectiveness of exercise therapy were based on the conclusions of all available systematic reviews. If the conclusions of the systematic reviews were conflicting, the sources of discordance among the conclusions of systematic reviews were explored (Jadad et al 1997). The panel based its final conclusions on the most complete systematic review, using the decision tool described by Jadad et al (1997).

Results

Selection of studies The results of our search strategy are presented in a flow chart (Fig. 1). Out of a total of 4017 abstracts, 228 reviews were considered to be potentially eligible for our best-evidence summary. Reviewing the full text resulted in the inclusion of 104 systematic reviews, including nine duplicates. The systematic reviews have been marked with an asterisk in the reference list.

Quality assessment The overall inter-rater agreement for the quality assessment was 86% (Cohen’s Kappa 0.73). Most of the disagreements were caused by differences in
Table 1. Results and conclusions of systematic reviews (quality score ≥ 60 points) on the effectiveness of exercise therapy (n = 45)

| Systematic review \( ^{a} \) | Disease | Score \( ^{b} \) | No. RCTs | Quality \( ^{d} \) | Conclusions SR | Dissent \( ^{b} \) | Conclusions Panel | # |
|-----------------------------|---------|---------------|----------|----------------|----------------|----------------|-----------------|
| Bradley & Moran (2002)       | Cystic fibrosis | 75 3 4 0 0 | ? ? | No | ? ? |
| Bradley et al (2002)         | Bronchiectasis | 64 2 1 0 0 | ? ? | No | ? ? |
| Smith et al (1992)           | COPD, asthma and bronchitis | 65 12 3 10 2 | - ? | Yes | ? ? |
| Lacasse et al (1996)         | COPD     | 70 14 NA 14 NA | + NA | No | + NA |
| Lacasse et al (1997a, 1997b) | COPD     | 62 6 18 0 2 | + ? | No | + ? |
| Cambach et al (1999)         | COPD     | 64 12 2 2 0 | + ? | No | + ? |
| De Goede et al (2001)        | Parkinson’s dis. | 65 5 3 2 2 | + ? | Yes | ? (ind) |
| Deane et al (2002b)          | Parkinson’s dis. | 64 NA 7 NA 5 NA | ? NA | No | NA |
| van der Lee et al (2001)     | CVA      | 79 1 12 1 9 | ? (ind) | No | ? (ind) | ? (ind) |
| Robeer et al (1998)          | Inter. claudication | 76 6 6 5 6 | + + | Yes | ? (ind) |
| Girolami et al (1999)        | Inter. claudication | 66 6 NA 6 NA | + NA | No | + NA |
| Leng et al (2002)            | Inter. claudication | 74 6 6 3 2 | + ? | No | + ? |
| Philadelphia Panel (2001b)   | Neck pain | 71 3 NA 0 NA | + NA | Yes | ? ? |
| Philadelphia Panel (2001c)   | Shoulder pain | 66 1 NA 0 NA | ? NA | No | NA |
| Philadelphia Panel (2001a)   | Knee pain | 72 6 NA 3 NA | + NA | No | + NA |
| van Tulder et al (1997)      | Low back pain (≤ 6 weeks) | 69 5 8 2 0 | - ? | No | - ? |
| Low back pain (> 12 weeks)   | 69 6 16 2 4 | + ? | No | + ? |
| Hilde & Bo (1998)            | Low back pain (> 4 weeks) | 69 5 8 1 4 | ? ? | No | ? ? |
| van Tulder et al (1999, 2000a, 2002b) | Low back pain (≤ 12 weeks) | 83 3 11 2 3 | - ? | No | - ? |
interpretation when discussing the power of the randomised controlled trials (see Appendix I, item L) and the heterogeneity of randomised controlled trials and outcomes (items N1, N2, N3, N4).

The mean (standard deviation) quality score of 95 systematic reviews (excluding the duplicate reviews) was 56 (17), ranging from 17 to 88 points (see Appendix II in the addenda at the AJP website, www.physiotherapy.asn.au/AJP). The most prevalent flaws were associated with the assessment of the methodological quality of the individual randomised controlled trials in the systematic review (items D1, D2, D5, D6, F, G), the data presentation (items J1, J2, J3, J4, L) and the evaluation of the results (items N1, N3, N4). There were 45 systematic reviews with a quality score of at least 60 points. These reviews investigated the effectiveness of exercise therapy for cystic fibrosis, chronic obstructive pulmonary disease (COPD), asthma, bronchiectasis, Parkinson’s disease, cerebrovascular accident (CVA), intermittent claudication, osteoarthritis, ankylosing spondylitis, rheumatoid arthritis, repetitive strain injury (RSI), neck pain, shoulder pain, and low back pain. Systematic reviews investigating the effectiveness of exercise therapy for patients with fibromyalgia, patellofemoral pain syndrome, carpal tunnel syndrome, temporomandibular joint displacement, multiple sclerosis, and cerebral palsy had low quality scores (< 60 points) (Baker and Tickle-Degnen 2001, Crossley et al 2001, Dodd et al 2002, Feuerstein et al 1999, Kropmans et al 1999, Rossy et al 1999, Stiller and Huff 1999, Zomerdiijk et al 1998). Consequently, these disorders will not be discussed.

For each systematic review (≥ 60 points), the quality score, the total number of randomised controlled trials, the number of high quality randomised controlled trials, the conclusions reported in the review, and the final conclusions of the panel

<table>
<thead>
<tr>
<th>Disorder</th>
<th>No. reviews</th>
<th>High quality RCTs</th>
<th>Total RCTs</th>
<th>RCTs for low back pain (≥ 12 weeks)</th>
<th>RCTs for low back pain (≤ 12 weeks)</th>
<th>RCTs for low back pain (&gt; 12 weeks)</th>
<th>RCTs for low back pain (&lt; 4 weeks)</th>
<th>RCTs for low back pain (4–12 weeks)</th>
<th>RCTs for low back pain (mixed group)</th>
<th>RCTs for musculo. dis.</th>
<th>RCTs for pulmonary disease</th>
<th>RCTs for rheumatoid arthritis</th>
<th>RCTs for RSI</th>
<th>RCTs for cognitive behaviour therapy</th>
<th>RCTs for cognitive behaviour therapy</th>
<th>RCTs for multidisciplinary biopsychosocial rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low back pain</td>
<td>83</td>
<td>8</td>
<td>19</td>
<td>3</td>
<td>10</td>
<td>+</td>
<td>±</td>
<td>No</td>
<td>+</td>
<td>±</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Low back pain (mixed group)</td>
<td>83</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>van Tulder et al (2000b)</td>
<td>81</td>
<td>11</td>
<td>13</td>
<td>3</td>
<td>3</td>
<td>+</td>
<td>?</td>
<td>No</td>
<td>+</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Guzman et al (2001)</td>
<td>72</td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>van Tulder et al (2002a)</td>
<td>88</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>van Tulder et al (2002b)</td>
<td>88</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>van Tulder et al (2002)</td>
<td>88</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>?</td>
<td>?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Beckerman et al (1993a)</td>
<td>62</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Because there are nine duplicates, the number of systematic reviews presented in this table is 45. Duplicates came from van Tulder et al (1999, 2000a, 2002b), Koes et al (1991a, 1991b), Green et al (1998, 2002), van Baar et al (1998a, 1999, 2001), van den Ende et al (1998, 2002), Ram et al (2000, 2002) and Lacasse et al (1997a, 1997b). The systematic reviews are ranked in order of publication (for each disorder), equally ranked reviews are ordered alphabetically. Total quality score of the systematic review; the quality score is calculated as the sum of all items. Number of randomised controlled trials of high quality based on the methodological quality presented in the systematic review; RCTs with at least 50% of the maximum quality score were regarded as ‘high quality’. Disagreement between the conclusions in the systematic review and the conclusions of the panel. The systematic review of van der Heijden (1995) investigated exercise therapy for patients with low back pain and neck pain and is therefore presented twice in this table. Conclusions were drawn regarding the effectiveness of exercise therapy for patients with knee osteoarthritis; only one large (smallest group > 50) RCT of methodological good quality (≥ 50% quality scores) investigated the effectiveness of exercise therapy for hip osteoarthritis and found positive results on pain, observed disabilities, and patients, global assessment. RCTs investigated the effectiveness of cognitive behavioural therapy (exercise therapy is included). RCTs investigated the effectiveness of multidisciplinary biopsychosocial rehabilitation (exercise therapy is included). RCTs investigated the effectiveness of back schools (exercise therapy is included).
are presented in Table 1. In five cases the panel disagreed with the authors of the systematic review with regard to the conclusions. These disagreements were mainly caused by inadequate reporting of the results of the randomised controlled trials in the systematic review (Philadelphia Panel 2001d, Robeer et al 1998) or because the conclusions were based on both randomised controlled trials and controlled clinical trials (De Goede et al 2001, Philadelphia Panel 2001b). In one systematic review the overall conclusions were drawn for a very heterogeneous patient population, namely patients with COPD, asthma, and bronchitis (Smith et al 1992).

**Characteristics of the systematic review** Details of each systematic review (≥ 60 points), including the research question(s), information on randomised controlled trials, the conclusions of the authors, and the final conclusions of the panel are presented in the Appendix III (see addenda at the AJP website, www.physiotherapy.asn.au/AJP).

**Cystic fibrosis** Three systematic reviews investigated the effectiveness of exercise therapy for patients with cystic fibrosis (Boyd et al 1994, Bradley and Moran 2002, Thomas et al 1995). Based on the results of two reasonable quality systematic reviews, we concluded that exercise therapy in addition to percussion, vibration, and postural drainage, has beneficial effects on FEV, (Forced Expiration Volume within one second) (Bradley and Moran 2002, Thomas et al 1995). The exercise therapy consisted of aerobic exercises (e.g. swimming), strength training exercises, and inspiratory muscle training. It is unclear whether exercise therapy is also effective for outcome measures such as quality of life. There is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment (no randomised controlled trials available), or compared to treatment consisting of percussion, vibration, and postural drainage, or other treatments for patients with cystic fibrosis.

**Asthma** Four systematic reviews investigated the effectiveness of exercise therapy for patients with asthma (Ernst 2000, Gosselink and Wagenaar 1993a, Gosselink and Wagenaar 1993b, Holloway and Ram 2002, Ram et al 2002, Ram et al 2000). Based on the results of two reasonable quality systematic reviews (Holloway and Ram 2002, Ram et al 2002, Ram et al 2000), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for children and adults with asthma, compared to no treatment or other conservative treatments.

**Bronchiectasis** With regard to bronchiectasis, we found only one reasonable quality systematic review (Bradley et al 2002). Due to the strict selection criteria applied in this systematic review, only two randomised controlled trials with poor quality reporting (abstract only) were included. Based on the results of this review, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with bronchiectasis.

**Chronic obstructive pulmonary disease (COPD)** During the period 1992–2002, 11 systematic reviews on the effectiveness of exercise therapy for COPD were published (Bekkering et al 1998, Cambach et al 1999, Chavannes and Vollenberg 2002, Devine and Peacy 1996, Gosselink and Wagenaar 1993a, Gosselink and Wagenaar 1993b, Lacasse et al 1996, Lacasse et al 1997a, Lacasse et al 1997b, Lacasse et al 2002, Smith et al 1992) we concluded that exercise therapy, consisting of aerobic exercises (e.g. walking, cycling) and strengthening exercises, is effective in improving the maximum and functional exercise capacity and quality of life of patients with COPD. Exercise therapy in a supervised program is probably more effective than exercise therapy in an unsupervised program, which showed no beneficial effects, compared to no treatment (Lacasse et al 2002). However, there were no randomised controlled trials included in the systematic reviews that directly compared the effectiveness of supervised exercise therapy to unsupervised exercise therapy. There is insufficient evidence to support or refute the effectiveness of a specific type of exercise therapy. There is also insufficient evidence to draw conclusions with regard to the effectiveness of exercise therapy, compared to other conservative treatments.

**Parkinson’s disease** Six systematic reviews investigated the effectiveness of exercise therapy for patients with Parkinson’s disease (Deane et al 2002a, Deane et al 2002b, Deane et al 2002c, De Goede et al 2001, Nieuwboer et al 1994, Rubinstein et al 2002). Based on the results of three reasonable quality systematic reviews (Deane et al 2002a, Deane et al 2002b, De Goede et al 2001), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with Parkinson’s disease. There are indications that intensive exercise therapy, consisting of general mobility activities focussing on balance, posture, walking, range of motion, fine motor dexterity, and functional exercises has positive effects on the activities of daily living for patients with Parkinson’s disease. However, this is based on randomised controlled trials with poor methodological quality or randomised controlled trials with small sample sizes (Comella et al 1994, Gauthier et al 1987, Patti et al 1996).

**Cerebrovascular accident (CVA)** Eleven systematic reviews investigated the effectiveness of exercise therapy in patients who had suffered a stroke (CVA) (de Bie et al 1995, Hiraoka 2001, Kwakkel et al 1997, Langhorne et al 1996, Ottenbacher and Jannell 1993, Pedro-Cuesta et al 1992, Pomeroy and Tallis 2000, Schoppink et al 1996, Snels et al 2002, van der Lee 2001, van der Lee et al 2001). Based on the results of three reasonable quality systematic reviews (Kwakkel et al 1997, Snels et al 2002, van der Lee et al 2001), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients who had suffered a stroke or for patients with hemiplegic shoulder pain, compared to no treatment or other conservative treatments. There are indications that (time-) intensive exercise therapy has more positive effects on the activities of daily living in patients who had suffered a stroke than less intensive exercise therapy. The exercise therapy consisted of neuromuscular facilitation and functional exercises, focusing on training of toilet transfers, rising from a sitting position, and walking. However, this was based on randomised controlled trials with poor methodological quality (Peacock et al 1972, Sivenius et al 1985, Smith et al 1981, Werner and Kessler 1996). More research is needed to confirm these results.

Based on the results of four reasonable quality systematic reviews (Brandsma et al 1998, Girolami et al 1999, Leng et al 2002, Robeer et al 1998), we concluded that exercise therapy is effective for patients with intermittent claudication, compared to no treatment. Exercise therapy consisted of (treadmill) training in walking, and lower limb strengthening exercises (e.g. stair climbing). The patients were encouraged to continue with daily walking exercises at home until they felt moderate pain. There are also indications that exercise therapy is more effective in improving maximal walking time than angioplasty (Creasy et al 1990) (Weighted Mean Difference (WMD) = 3.30 minutes; 95% CI 2.21 to 4.39) or antiplatelet therapy (Mannarino et al 1991) (WMD = 1.06 minutes; 95% CI 0.15 to 1.97), and there are indications that exercise therapy is equally as effective as surgery (Lundgren et al 1989) (WMD = -1.66 minutes; 95% CI -4.55 to 1.23).

However, this was based on randomised controlled trials that either had small sample sizes or the methodological quality was not described in the review (unclear) (Creasy et al 1990, Lundgren et al 1989, Mannarino et al 1991). No conclusions can be drawn with regard to the effectiveness of a specific type of exercise therapy for patients with intermittent claudication.

**Osteoarthritis** Seven systematic reviews investigated the effectiveness of exercise therapy for patients with knee or hip osteoarthritis (Fransen et al 2002, McCarthy and Oldham 1999, Pendleton et al 2000, Petrella 2000, Philadelphia Panel 2001a, Puett and Griffin 1994, van Baar et al 1998a, Van Baar et al 1999, van Baar et al 2001). Based on the results of three reasonable or good quality systematic reviews, we concluded that exercise therapy, consisting of strengthening, stretching, and functional exercises, is effective for patients with knee osteoarthritis, compared to no treatment (Fransen et al 2002, Philadelphia Panel 2001a, van Baar et al 1998a, van Baar et al 1999, van Baar et al 2001). There are indications that exercise therapy (e.g. strengthening and stretching exercises, functional training, and ADL instruction) is effective for patients with hip osteoarthritis. However, this is based on one large randomised controlled trial with good methodological quality (van Baar et al 1998b). There is insufficient evidence to support or refute the effectiveness of a specific type of exercise therapy (individual, group therapy, or hydrotherapy) for patients with knee or hip osteoarthritis.

**Ankylosing spondylitis** Two systematic reviews investigated the effectiveness of exercise therapy for patients with ankylosing spondylitis (Ammer 1997, Dagfinrud and Hagen 2002). Based on one good quality systematic review (Dagfinrud and Hagen 2002), we concluded that there are indications to support the effectiveness of exercise therapy, compared to no treatment for patients with ankylosing spondylitis. The exercise therapy consisted of functional exercises and exercises to improve mobility, strength, and endurance, using normal movement patterns and proprioceptive neuromuscular facilitation. The patients received disease education and were encouraged to continue their exercises daily at home. However, this was based on only one small good quality randomised controlled trial (Kraag et al 1990). No conclusions can be drawn with regard to the effectiveness of exercise therapy, compared to other types of exercise therapy or other treatments.

**Rheumatoid arthritis** Two systematic reviews investigated the effectiveness of exercise therapy for patients with rheumatoid arthritis (Augustinus et al 2000, van den Ende et al 1998, van den Ende et al 2002). Based on one reasonable quality systematic review, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with rheumatoid arthritis (van den Ende et al 1998, van den Ende et al 2002).

**Repetitive strain injury** With regard to repetitive strain injury, we found only one reasonable quality systematic review (Konijnenberg et al 2001). We concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with repetitive strain injury.

**Neck pain** We found six systematic reviews investigating the effectiveness of exercise therapy for patients with non-specific neck pain (Aker et al 1996, Gross et al 2002, Hurwitz et al 1996, Kjellman et al 1999, Philadelphia Panel 2001b, van der Heijden et al 1995). Based on the results of these six reasonable quality systematic reviews, we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy, compared to no treatment or other conservative treatments, for patients with (non-specific) neck pain.

**Shoulder pain** During the period 1997–2002, four systematic reviews on the effectiveness of exercise therapy for shoulder pain were published (Green et al 1998, Green et al 2002, Johansson et al 2002, Philadelphia Panel 2001c, van der Heijden et al 1997). Based on the results of three reasonable quality systematic reviews (Green et al 1998, Green et al 2002, Philadelphia Panel 2001c, van der Heijden et al 1997), we concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with shoulder pain or shoulder complaints.


For patients with acute low back pain (< 6 weeks) there is no difference in the effectiveness of exercise therapy (e.g. stretching, strengthening, extension/flexion exercises), compared to no treatment, care provided by a general practitioner, or manipulations (high velocity techniques). For patients with sub-acute (6 to 12 weeks) and chronic (> 12 weeks) low back pain, we concluded that exercise therapy is effective compared to no treatment. The exercise therapy consisted of aerobic exercises (e.g. walking, jogging), and intensive strengthening exercises for the abdomen and trunk muscles. Exercise therapy (e.g. aerobic exercises, progressive muscle relaxation) in combination with cognitive behavioural therapy is also more effective than no treatment for patients with chronic low back pain.
For patients with chronic (> 12 weeks) low back pain, exercise therapy (e.g. strengthening exercises) is more effective than continued care provided by a general practitioner, and equally as effective as conventional physiotherapy (e.g. traction, massage, ultrasound, mobilisation exercises, hot and cold packs). There is insufficient evidence to support or refute the effectiveness of a particular type of exercise therapy for patients with sub-acute or chronic low back pain. There are indications that intensive multidisciplinary bio-psychosocial rehabilitation with functional restoration (including intensive aerobic exercises, stretching exercises, and muscle relaxation therapy) is more effective than physical training plus back school for patients with chronic low back pain. However, this was based on only one good quality randomised controlled trial with a short and long-term follow-up (Bendix et al 1995).

There are indications that exercise therapy, consisting of abdominal strengthening exercises, in addition to back school, is effective for patients with chronic low back pain, compared to back school without exercise therapy. However, this was also based on only one randomised controlled trial with good methodological quality (Klaber-Moffett et al 1986). There is insufficient evidence to support or refute the effectiveness of cognitive behavioural therapy plus exercise therapy compared to other conservative treatments for patients with chronic low back pain. There is also insufficient evidence to draw conclusions with regard to the (in)effectiveness of back schools for patients with acute, sub-acute or chronic low back pain.

**Discussion**

Exercise therapy is effective for patients with knee osteoarthritis, sub-acute and chronic low back pain, cystic fibrosis, COPD, and intermittent claudication. Furthermore, there are indications that exercise therapy is effective for patients with ankylosing spondylitis, hip osteoarthritis, and Parkinson’s disease, and also for patients who have suffered a stroke. We concluded that there is insufficient evidence to support or refute the effectiveness of exercise therapy for patients with neck pain, shoulder pain, RSI, rheumatoid arthritis, asthma, and bronchiectasis. Exercise therapy is not effective for patients with acute low back pain. Based on the available literature, we found no evidence that exercise therapy is harmful or that it provoked harmful side effects. However, systematic reviews provide little information on the safety aspects of exercise therapy. This is mainly due to the inadequate reporting of adverse effects in randomised controlled trials (Ernst and Pittler 2001).

For certain diseases (fibromyalgia, patellofemoral pain syndrome, carpal tunnel syndrome, tempromandibular joint displacement, multiple sclerosis, cerebral palsy), only systematic reviews with low scores for quality (< 60 points) were available (Baker and Tickle-Degnen 2001, Crossley et al 2001, Dodd et al 2002, Feuerstein et al 1999, Kropmans et al 1999, Rossy et al 1999, Stillier and Huff 1999, Zomerdijk et al 1998). For these disorders, we recommend that systematic reviews be carried out using methods that accord to the current state of knowledge (Egger et al 2001).

Although a number of systematic reviews were of reasonable or good quality, there was still insufficient evidence to draw firm conclusions with regard to the (in)effectiveness of exercise therapy for neck pain, shoulder pain, repetitive strain injury, rheumatoid arthritis, asthma, and bronchiectasis. This was mainly due to the contradictory results, the poor methodological quality of the randomised controlled trials, inadequate reporting, small sample sizes, and the large variation in outcome measures and study populations. We recommend that searches be conducted for new published, large randomised controlled trials of good quality (since the last search date of the most recent systematic review of reasonable or good quality) on the effectiveness of exercise therapy for the following disorders: neck pain, shoulder pain, RSI, rheumatoid arthritis, asthma, and bronchiectasis. If no new randomised controlled trials have been published, or the retrieved randomised controlled trials are of poor methodological quality, we recommend that a new, large randomised controlled trial with good methodological quality be carried out.

We found indications to support the effectiveness of exercise therapy for patients with ankylosing spondylitis, hip osteoarthritis, Parkinson’s disease and patients who had suffered a stroke, but more randomised controlled trials are needed to confirm these results.

With regard to the disorders for which exercise therapy appeared to be effective, it still remains to be determined whether exercise therapy should be included in a supervised or an unsupervised program, and whether exercise at home is sufficient or referral should be made to a physiotherapist. There is also insufficient evidence to support or refute the effectiveness of specific types of exercise therapy for almost all disorders. More research is also needed to investigate how the short-term effectiveness of exercise therapy can be maintained in the long-term. Programs or methods with which care-providers could encourage the compliance of patients with home exercises and motivate them to continue their exercises in the future would be very useful.

This best-evidence summary of systematic reviews has a number of limitations. First, different weights were applied to the five quality criteria, including the selection of studies, methodological quality assessment of the randomised controlled trials, description of the intervention, data presentation, and evaluation. Total quality scores were calculated by summing up the weights of all quality items. The advantage of using an overall quality score is its simplicity, but methodologically it is debatable. If equal weights were applied to each quality item, the division of systematic reviews into good, reasonable, moderate, poor, and very poor quality would be quite similar, and the final conclusions with regard to the effectiveness of exercise therapy would still be the same.

Second, the choice of the cut-off point for reasonable or good quality was arbitrary. The quality of the reporting of the results of systematic reviews with low scores for quality (< 60 points) was often too poor to draw conclusions with regard to the effectiveness of exercise therapy. If, for example, the cut-off point was set at 50 points, another 11 reviews would have been included. However, our conclusions with regard to the effectiveness of exercise therapy for the disorders discussed in this review would remain the same (data not shown). We could only draw new conclusions with regard to the effectiveness of exercise therapy for patellofemoral pain syndrome.

Third, our conclusions were based on statistically significant differences, rather than clinically relevant differences. Unfortunately, based on the results presented in the
systematic reviews, it was not possible to calculate effect sizes. Therefore, the clinically relevant differences were not taken into account in our conclusions.

Finally, a few systematic reviews on the same topic reported conflicting conclusions. However, based on the guidelines developed by Jadad et al (1997), explaining differences in research questions, assessment of the quality of randomised controlled trials, number of randomised controlled trials, and statistical methods for data-analysis, the panel succeeded in drawing clear conclusions.

In conclusion, exercise therapy has been shown to be effective for a wide range of (chronic) disorders.

Footnotes
*Contributing authors (members of the Exercise Therapy Group):
Johan H Arendsen, Leiden University Medical Center, Leiden, The Netherlands
Rob A de Bie, University of Maastricht, The Netherlands
Sita MA Bierma-Zeinstra, Erasmus Medical Center, Rotterdam, The Netherlands
Lex M Bouter, VU University Medical Center, The Netherlands
Henrica CW de Vet, VU University Medical Center, Amsterdam, The Netherlands
Joost Dekker, VU University Medical Center, Amsterdam The Netherlands (chair)
Paul JM Helders, University Medical Center Utrecht, The Netherlands
Samyra HJ Keus, Leiden University Medical Center, Leiden, The Netherlands
Gert Kwakkel, VU University Medical Center, Amsterdam, The Netherlands
Ton Lenssen, University of Maastricht, The Netherlands
Rob AB Oostendorp, University Medical Center St. Radboud, Nijmegen, The Netherlands
Raymond WJG Ostelo, VU University Medical Center, Amsterdam, The Netherlands
Max Reijman, Erasmus Medical Center, Rotterdam, The Netherlands
Nynke Smidt, VU University Medical Center, Amsterdam, The Netherlands
Caroline B Terwee, VU University Medical Center, Amsterdam, The Netherlands
Carlo Theunissen, University of Maastricht, The Netherlands
Steph Thomas, Erasmus Medical Center, Rotterdam, The Netherlands
Margriet E van Baar, Erasmus Medical Center, Rotterdam, The Netherlands
Alex van ‘t Hul, VU University Medical Center, Amsterdam, The Netherlands
Roland PS van Peppen, VU University Medical Center, Amsterdam, The Netherlands
Arianne P Verhagen, Erasmus Medical Center, Rotterdam, The Netherlands
Daniëlle AWM van der Windt, VU University Medical Center, Amsterdam, The Netherlands

Acknowledgements This article was based on the work of the ‘Panel on Exercise Therapy’ of the Dutch Health Council. The authors would like to thank Karin Johnson for blinded the articles.

Correspondence Nynke Smidt, Institute for Research in Extramural Medicine, VU University Medical Center, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.
Email: <n.smidt@vumc.nl>

References
Systematic reviews have been marked with an asterisk.


Hunt DL and McKibbon KA (1997): Locating and appraising...


## Appendix 1. Criteria for the assessment of the quality of the systematic reviews.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Maximal points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study selection</strong></td>
<td>(30)</td>
</tr>
<tr>
<td><strong>A</strong> Description of inclusion and exclusion criteria of the systematic review</td>
<td></td>
</tr>
<tr>
<td>1 Study setting(s) included (i.e. industry, general practice, hospital)</td>
<td>2</td>
</tr>
<tr>
<td>2 Interventions type(s) included</td>
<td>2</td>
</tr>
<tr>
<td>3 Outcome type(s) included (i.e. pain, general improvement, disability questionnaire)</td>
<td>2</td>
</tr>
<tr>
<td>4 Years covered</td>
<td>2</td>
</tr>
<tr>
<td>5 Language(s) covered</td>
<td>2</td>
</tr>
<tr>
<td><strong>B</strong> Search strategy</td>
<td></td>
</tr>
<tr>
<td>1 Established bibliographic database included (Medline (or PubMed), and at least one other database)</td>
<td>5</td>
</tr>
<tr>
<td>2 Additional efforts to locate non-indexed randomised clinical trials (RCTs) (e.g. reference tracking, correspondence with experts, manual search of non-indexed journals)</td>
<td>5</td>
</tr>
<tr>
<td><strong>C</strong> Emphasis on RCTs: RCTs only, or results or RCTs discussed separately from other study designs</td>
<td>10</td>
</tr>
<tr>
<td><strong>Methodological Quality Assessment</strong></td>
<td>(20)</td>
</tr>
<tr>
<td><strong>D</strong> Assessment (of the validity) of RCTs included that is explicit (reproducible by readers of the review) regarding:</td>
<td></td>
</tr>
<tr>
<td>1 Similarity of treatment groups at baseline (prognostic factors)</td>
<td>2</td>
</tr>
<tr>
<td>2 Similarity of treatment characteristics (co-interventions)</td>
<td>2</td>
</tr>
<tr>
<td>3 Adequacy of treatment of missing values (dropouts, loss to follow-up)</td>
<td>2</td>
</tr>
<tr>
<td>4 Blinding of outcome assessment</td>
<td>2</td>
</tr>
<tr>
<td>5 Relevance of outcome measures</td>
<td>2</td>
</tr>
<tr>
<td>6 Adequacy of statistical analysis (i.e. intention-to-treat analysis)</td>
<td>2</td>
</tr>
<tr>
<td><strong>E</strong> Number of reviewers (at least two independent reviewers)</td>
<td>4</td>
</tr>
<tr>
<td><strong>F</strong> Blinding of reviewer(s): (blinded for source of article: journal, year of the trial, publication, institute)</td>
<td>2</td>
</tr>
<tr>
<td><strong>G</strong> Agreement of reviewer(s): reported (quantitatively in percentage agreement or Kappa statistics) and acceptable (cut-off Kappa statistics &gt; 0.60, where Kappa statistics is not reported look at percentage agreement, which should be at least 80%). In the event of reviewer, use of an assessment list with established reliability.</td>
<td>2</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>(15)</td>
</tr>
<tr>
<td><strong>H</strong> Description of (index) intervention(s) (exercises) per RCT</td>
<td></td>
</tr>
<tr>
<td>1 Description of therapeutic exercise (i.e. strength, endurance and cardiovascular fitness, mobility and flexibility, stability, relaxation, coordination, balance, and functional skills)</td>
<td>3</td>
</tr>
<tr>
<td>2 Profession or training of care provider</td>
<td>1</td>
</tr>
<tr>
<td>3 Treatment frequency or number of treatments</td>
<td>2</td>
</tr>
<tr>
<td>4 Duration of treatment period</td>
<td>2</td>
</tr>
<tr>
<td><strong>I</strong> Description of control intervention(s): per RCT</td>
<td></td>
</tr>
<tr>
<td>1 Type (e.g. conservative treatments, wait-and-see policy, surgery)</td>
<td>3</td>
</tr>
<tr>
<td>2 Treatment frequency or number of treatments</td>
<td>2</td>
</tr>
<tr>
<td>3 Duration of treatment period</td>
<td>2</td>
</tr>
<tr>
<td><strong>Data Presentation</strong></td>
<td>(20)</td>
</tr>
<tr>
<td><strong>J</strong> Outcome presentation (for the most important (clinical relevant) outcome measures)</td>
<td></td>
</tr>
<tr>
<td>1 The original data of the main outcome(s) are presented separately per RCT per group</td>
<td>5</td>
</tr>
<tr>
<td>2 Presentation of the mean difference (effect size, standardised mean differences, weighted mean differences) or ratio of outcome(s) (relative risk, risk difference, odds ratio) between intervention group(s) and control group(s)</td>
<td>3</td>
</tr>
<tr>
<td>3 Presence of confidence interval (i.e. 95% CI) or standard deviation (SD) per RCT</td>
<td>3</td>
</tr>
<tr>
<td>4 Graphic presentation of the most important outcome(s) (indicating outliers and distribution) per RCT (presentation of a tree plot, meta-analysis)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Adequate summary of research findings: statistical pooling of the most important outcome(s); discussion of the reason why pooling is not indicated or warranted; or pooling of the subset considered to be valid and similar enough</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>L</td>
<td>Discussion of the power of negative RCTs</td>
</tr>
<tr>
<td></td>
<td>1 Calculation (quantitative) of the power of each RCT</td>
</tr>
<tr>
<td></td>
<td>2 Narrative elaboration (qualitative) on the power of each negative RCT</td>
</tr>
<tr>
<td></td>
<td>3 Overall narrative elaboration on the power of the negative RCTs (i.e. remarks about small sample sizes)</td>
</tr>
<tr>
<td></td>
<td><strong>Evaluation</strong></td>
</tr>
<tr>
<td>M</td>
<td>Overall conclusion regarding the aggregated level of available RCTs on the effectiveness of the (index) intervention presented</td>
</tr>
<tr>
<td>N</td>
<td>Discussion of heterogeneity of RCTs and outcomes</td>
</tr>
<tr>
<td></td>
<td>1 Identification of relevant subgroups (e.g. age, study setting, disease classification) with explicit motivation</td>
</tr>
<tr>
<td></td>
<td>2 Discussion of variety of treatment modalities in the intervention groups (i.e. high dose exercises)</td>
</tr>
<tr>
<td></td>
<td>3 Discussion of variety of treatment modalities in control groups (placebo, existing modality)</td>
</tr>
<tr>
<td></td>
<td>4 Discussion of relationship between methodological quality of RCTs and outcome</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>