The impact of physical therapy on functional outcomes after stroke: what's the evidence?


CRD summary
This review assessed the effects of physical therapy interventions for improving functional outcomes after stroke. The authors concluded that there were small to large effect sizes for task-orientated exercise training, particularly intensive training applied early after stroke. An assessment of the evidence was difficult, owing to the multiplicity of outcome measures and the lack of defined control treatments.

Authors' objectives
To assess the effects of physical therapy interventions focused on improving functional outcomes after stroke.

Searching
MEDLINE, CINAHL, the Cochrane CENTRAL Register, the Cochrane Database of Systematic Reviews, DARE, PEDro, EMBASE and DocOnline were searched up to January 2004 for studies published in English, German or Dutch. The search terms were reported. Bibliographies of studies, reviews, narrative reviews and abstracts published in conference proceedings were screened.

Study selection
Study designs of evaluations included in the review
Randomised controlled trials (RCTs) and controlled clinical trials (CCTs) were eligible for inclusion. Studies were included regardless of quality.

Specific interventions included in the review
Studies of physical therapy interventions focused on improving functional outcomes were eligible for inclusion. Studies were excluded if they used acupuncture or drug treatment in combination with physical therapy, or if they used robotics. In the review, interventions were categorised into the following ten categories:

- traditional neurological approaches;
- programmes for training sensorimotor function or influencing muscle tone;
- cardiovascular fitness and aerobic programmes;
- methods for training mobility and mobility-related activities;
- exercises for the upper limb;
- biofeedback therapy for the upper and lower limb;
- functional and neuromuscular electrical stimulation for both limbs;
- orthotics and assistive devices for both limbs;
- treatments for hemiplegic shoulder pain and hand oedema;
- intensity of exercise therapy.

Categorisations were based on the International Classification of Functioning, Disability and Health of the World Health Organization and the American Physical Therapy Association guide to physical therapist practice. The included
Participants included in the review
Studies of adults (aged 18 years or older) who had been diagnosed with a stroke were eligible for inclusion. The range of time post-stroke varied widely both within and between the included studies (the range was from a few days to years).

Outcomes assessed in the review
Studies that evaluated effectiveness with the focus on functional outcomes were eligible for inclusion. The included studies assessed a wide variety of outcomes (full details were reported).

How were decisions on the relevance of primary studies made?
Two reviewers independently searched the electronic databases and selected studies.

Assessment of study quality
The RCTs were assessed and scored using the following criteria of the PEDro scale: eligibility criteria; random allocation, allocation concealment; baseline comparability of the treatment groups; blinding of patients, therapists and assessors; outcomes for more than 85% of patients; intention-to-treat analysis; between-group comparisons; and point measures and measures of variability. The maximum possible score was 10 points. Studies scoring 4 or more points were considered high quality, while those scoring 3 or less were considered low quality. The authors did not state that they assessed the validity of the CCTs.

Two reviewers independently assessed the validity of the RCTs and resolved any disagreements through discussion or with the help of a third reviewer, if required. Inter-rater agreement on individual validity items was assessed using Cohen's kappa statistic.

Data extraction
Eight physical therapists and two reviewers reached consensus about the categorisation of interventions. The authors did not state how other data were extracted for the review, or how many reviewers performed the data extraction. Data from crossover RCTs were extracted only for the first time period. For each study an effect size was reported and classified, as advised by Cohen, as small (less than 0.2), medium (0.2 to 0.5) or large (more than 0.5).

Methods of synthesis
How were the studies combined?
The studies were grouped by intervention. RCTs with similar interventions, patient characteristics and outcome measures were pooled in a meta-analysis and pooled effect sizes (Hedge's $g$; SES) with 95% confidence interval (CI) were calculated; individual effect sizes were weighted by sample size. Fixed-effect models were used in the absence of heterogeneity, otherwise random-effects models were used. Where pooling was not possible, the level of evidence for interventions was graded using a hierarchy of evidence described by Van Tulder et al. (see Other Publications of Related Interest no.1).

How were differences between studies investigated?
Statistical heterogeneity between the studies was assessed, but the details of the test used were not reported. The association between quality-unbiased effect sizes and validity scores of RCTs and between year of publication and PEDro score was investigated. Some other differences between the studies were discussed in the text.

Results of the review
One hundred and fifty-one studies were included (n approximately 9,139): 123 RCTs (n approximately 7,507) and 28 CCTs (n approximately 1,632). The number of patients was reported for each intervention, but the potential for double-counting it difficult to be certain of the total number of patients included in the review.

Inter-rater agreement (kappa) of RCT validity items was 0.81. The median quality score for RCTs was 5 out of 10.
points (range: 2 to 8 points), indicating moderate quality. Allocation concealment was reported in 39 RCTs, intention-to-treat analysis in 19 RCTs and blinding of the observer in 72 RCTs. In none of the RCTs was it possible to blind the patient or therapist. Other methodological limitations included small sample sizes.

The meta-analyses reported below used fixed-effect models unless otherwise stated.

There was strong evidence for task-orientated exercise training for the restoration of balance and gait and for strengthening the lower paretic limb. Effect sizes of functional outcomes ranged from SES (random-effects model) 0.13 (95% CI: 0.03, 0.23; 20 RCTs with 2,686 patients) for high-intensity exercise training to SES 0.92 (95% CI: 0.54, 1.29; 4 RCTs with 128 patients) for improving symmetry when moving from sitting to standing.

There was strong evidence for physical therapies targeting functional training of the upper limb: e.g. constraint-induced movement therapies (SES 0.46, 95% CI: 0.07, reported as 0.91 in the abstract and 0.85 in a table; 5 RCTs with 104 patients), treadmill training with body weight support (SES 0.70, 95% CI: 0.29, 1.10; 3 RCTs with 148 patients), treadmill training without body weight support therapies (SES 1.09, 95% CI: 0.56, 1.61; 2 RCTs with 65 patients), aerobics (SES 0.39, 95% CI: 0.05, 0.74; 2 RCTs with 135 patients), external auditory rhythms during gait (SES 0.91, 95% CI: 0.40, 1.42; 3 RCTs with 67 patients) and neuromuscular stimulation for glenohumeral subluxation (SES 1.41, 95% CI: 0.76, 2.06; 4 RCTs with 161 patients).

There was no or insufficient evidence for: traditional neurological treatment approaches; exercises for the upper limb; biofeedback; functional and neuromuscular electrical stimulation aimed at improving dexterity or gait performance; orthotics and assistive devices; and interventions to reduce hemiplegic shoulder pain and hand oedema.

**Authors’ conclusions**

There were small to large effect sizes for task-orientated exercise training, particularly intensive training applied early after stroke. In most high-quality RCTs these benefits were only seen in tasks that were targeted in the exercise programme.

**CRD commentary**

The review addressed a clear but very broad question that was defined in terms of the participants, intervention, outcomes and study design. Several relevant sources were searched and some limited attempts were made to minimise publication and language bias; the authors acknowledged the limitations of including studies reported in only three languages. Methods were used to minimise reviewer errors and bias in the selection of studies, assessment of validity and categorisation of interventions, but it was not clear whether similar steps were taken when extracting the rest of the data. Validity was assessed using an established checklist and specific methodological limitations of the included studies were described. The individual studies were not described in detail and control interventions were not reported.

The studies were appropriately grouped by interventions and the pooling of only clinically homogeneous studies appeared appropriate. However, since the control interventions compared with the specified interventions were not reported, it was not clear whether all pooled studies had a similar control and, hence, if the pooling was really appropriate. Significant statistical heterogeneity was found in some meta-analyses, and even meta-analysis using a random-effects model may not have been appropriate. Combining other studies in a best-evidence synthesis seemed the most appropriate method of summarising evidence from such a large number of diverse studies. The multiplicity of measures used to assess the outcomes has the potential for selective reporting of results. A minority of studies analysed data on an intention to-treat basis and although the authors acknowledged that this might have influenced results, there was no exploration of the effects of drop-outs on outcomes. The authors also limited their assessment of crossover studies to the first time period only. This results in the loss of data as crossover trials are powered such that data from all time periods is required. Such trials should instead use all the available data in an appropriate paired analysis.

Overall, an assessment of the evidence was difficult, owing to the multiplicity of outcome measures and the lack of defined control treatments. Given the above concerns, it is difficult to comment on the robustness of the review’s findings.
Implications of the review for practice and research

Practice: The authors stated that there is strong evidence that patients benefit from exercise programmes in which functional tasks are directly and intensively trained.

Research: The authors stated that there is need for further, large, high-quality RCTs and a need for consensus about the most appropriate core-set of outcomes measured to use in studies of stroke rehabilitation. In particular, there is a need for trials investigating: traditional neurological treatment approaches; exercises for the upper limb; biofeedback; functional and neuromuscular electrical stimulation aimed at improving dexterity or gait performance; orthotics and assistive devices; and interventions to reduce hemiplegic shoulder pain and hand oedema.

Funding
Royal Dutch Society for Physical Therapy, grant number IKPZ/KNGF/05/2000.

Bibliographic details

PubMedID
15609840

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Biofeedback, Psychology; Edema /rehabilitation; Hand; Humans; Muscle, Skeletal /physiopathology; Orthotic Devices; Oxygen Consumption /physiology; Physical Fitness /physiology; Physical Therapy Modalities; Recovery of Function /physiology; Shoulder Pain /physiopathology /rehabilitation; Stroke /physiopathology /rehabilitation

AccessionNumber
12005009146

Date bibliographic record published
31/01/2007

Date abstract record published
31/01/2007

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.