What does a structured review of the effectiveness of exercise interventions for persons with multiple sclerosis tell us about the challenges of designing trials?

Asano M, Dawes DJ, Arafah A, Moriello C, Mayo NE

CRD summary
This review assessed the effectiveness of exercise interventions for multiple sclerosis patients. The authors concluded that there was insufficient evidence to recommend regular exercise prescription for these patients and that there is a need for additional high quality RCTs focusing on stratified samples. Given the level of evidence presented, the authors' conclusions are likely to be reliable.

Authors' objectives
To evaluate the effectiveness of exercise interventions for multiple sclerosis patients.

Searching
MEDLINE, CINHAL, EMBASE and the Cochrane Library were searched from 1950 to December 2007. Search terms were not reported. Only published studies were included.

Study selection
Randomised controlled trials (RCTs) that compared participative active exercise interventions (prescribed exercise, exercise therapy or physical therapy other than training for activities of daily living) with a control group for patients with multiple sclerosis were eligible for inclusion. Eligible trials had to report quantitative results (mean, standard deviation or standard error). Trials were excluded if the primary aim was to assess the effects of medications, educational programs or other devices on patients' ability to exercise, or multidisciplinary approaches where insufficient details of the exercise component were given.

Mean age (range of 30 to 65 years) and time since diagnosis (0 to 40 years) varied across included trials. The majority (66%) of included patients were female. Most studies included patients with minimal to moderate level of disability.

Four main combinations (maximum of two) of exercise intervention (aerobic, yoga, resistance or stretching) were provided for a defined duration, ranging from three weeks to six months at varying frequency, length of session and intensity. Target outcome measures (48 in total) varied across trials with those for body function, structure and activity (walk tests and physiological tests) most frequently used.

Two reviewers independently assessed studies for inclusion.

Assessment of study quality
The methodological quality of each trial was assessed using the Physiotherapy Evidence Database (PEDro) Scale consisting of 11 criteria resulting in a composite score ranging from 0 (poor) to 10 (excellent), where a score of 6 or more was considered to be level 1 evidence (6 to 8 was good, 9 to 10 was excellent) and a score of 5 or less was considered to be level 2 evidence (4 to 5 was fair, less than 4 was poor).

A total of 14 reviewers independently assessed at least two studies each for methodological quality.

Data extraction
In addition to trial summary details, the following data items were extracted from included studies: target outcomes (classified using the World Health Organisation's International Classification of Functioning, Disability and Health (ICF)) and the effect size (ES). The effect size was calculated from the difference in mean change pre- and post-intervention between the control and intervention groups, then dividing the mean change by the initial pooled standard deviation (SD) of that variable. An adjusted effect size, taking into account sample size and variability among trials, was then calculated using a mathematical formula. A positive effect size indicated an improvement in an intervention group; a negative effect size indicated an improvement in a control group. An effect size of less than 0.20 was
considered to be trivial, 0.20 to 0.50 small, 0.50 to 0.80 moderate and over 0.80 strong.

The authors did not state how the data were extracted or how many reviewers performed the extraction.

Methods of synthesis
Due to the heterogeneity of the participants, primary goals, interventions and consequently the multiple target outcomes across trials, the adjusted effect size was used to describe possible effects of each intervention on target outcome measures. A forest plot of hedges’ g effect sizes and 95% confidence intervals (CI) was presented.

Results of the review
Eleven RCTs (n=502 patients) were included in the review. All included trials scored 5 or above using the scale, but no trials achieved a score of 9 or 10 (indicating excellent quality).

The effect size ranged from -0.36 to 3.50 with a mean of 0.55, 95% CI 0.45 to 0.66 for the target outcomes measured, with body function effect size range of -0.29 to 3.50, activity measure and participation effect size range of -0.06 to 0.48 and quality of life effect size range of -0.36 to 2.56.

All trials showed at least one positive effect size of moderate to large size. Five trials included outcomes with negative and small effect sizes. Most of the trials had wide 95% confidence intervals, which included the null value for the effect size.

Authors' conclusions
There was insufficient evidence to recommend regular exercise prescription for this population of patients with multiple sclerosis.

CRD commentary
This review had clear inclusion and exclusion criteria with regard to study type, participants and interventions, but the reported outcome measures only needed to be quantitative. Relevant databases were searched. Search terms were not reported. Unpublished studies were not included, so publication bias may have been introduced. It was unclear whether language restrictions were imposed, so language bias could not be ruled out. Efforts to minimise reviewer bias and error were carried out at the study selection and assessment stages of the review process, but the reviewers did not state the method used for data extraction.

Trial validity was formally assessed; however, the quality scale used (PEDro) may not have been the most appropriate for rehabilitation trials since, as the authors highlighted, three points focus on subject blinding, which is unfeasible for these trials. In addition, the PEDro scale does not evaluate appropriateness of statistical analysis or sample size, which would be a methodological issue, particularly as the authors stated that for most included trials the sample size (five to 52 per group) was too small to detect a difference in effect. In light of the heterogeneity across trials, the method used to compare trials was appropriate.

Given the level of evidence presented, the authors' conclusions are likely to be reliable.

Implications of the review for practice and research
Practice: The authors stated that the gathering of evidence in practice needs to focus upon the individual patient's needs and lifestyle.

Research: The authors stated that there is a need for high quality, adequately powered RCTs to be conducted in order to add to the limited evidence base. These trials could be stand alone or could be designed with a view to contributing to a meta-analysis, with an a priori establishment of inclusion and exclusion criteria, outcomes and time points for assessment. In addition, these trials should use stratified samples of more than 40 years plus 40 to 65 years in order to evaluate possible effects on younger and older adults, and also for groups with mild and severe disabilities.
Funding
Canadian Institute of Health Research Neuroinflammation Training Award; Richard and Edith Strauss Canada Foundation Fellowship in Rehabilitation Sciences.

Bibliographic details

PubMedID
19324979

DOI
10.1177/1352458508101877

Original Paper URL
http://msj.sagepub.com/cgi/content/abstract/15/4/412

Indexing Status
Subject indexing assigned by NLM

MeSH
Exercise; Humans; Multiple Sclerosis /rehabilitation /therapy; Randomized Controlled Trials as Topic

AccessionNumber
12009104728

Date bibliographic record published
23/09/2009

Date abstract record published
03/03/2010

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.