Cochrane Library

Cochrane Database of Systematic Reviews

Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis (Review)

Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM

Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM.
Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis.
Cochrane Database of Systematic Reviews 2009, Issue 4. Art. No.: CD006853.
DOI: 10.1002/14651858.CD006853.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis

Emalie Hurkmans[1], Florus J van der Giesen[2], Thea PM Vliet Vlieland[1], Jan Schoones[3], Els CHM Van den Ende[4]

[1]Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands. [2]Department of Physical Therapy, Leiden University Medical Center, Leiden, Netherlands. [3]Walaeus Library, Leiden University Medical Center, Leiden, Netherlands. [4]Department of Rheumatology, Sint Maartenskliniek, Nijmegen, Netherlands

Contact address: Emalie Hurkmans, Department of Rheumatology, Leiden University Medical Center, Albinusdreef 2, Leiden, Zuid-Holland, 2333 ZA, Netherlands. e.j.hurkmans@lumc.nl.

Editorial group: Cochrane Musculoskeletal Group.
Publication status and date: New, published in Issue 1, 2010.

Citation: Hurkmans E, van der Giesen FJ, Vliet Vlieland TPM, Schoones J, Van den Ende ECHM. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews* 2009, Issue 4. Art. No.: CD006853. DOI: 10.1002/14651858.CD006853.pub2.

ABSTRACT

Background

An up-to-date overview of the effectiveness and safety of dynamic exercise therapy (exercise therapy with a sufficient intensity, duration, and frequency to establish improvement in aerobic capacity and/or muscle strength) is lacking.

Objectives

To assess the effectiveness and safety of short-term (< three months) and long-term (> three months) dynamic exercise therapy programs (aerobic capacity and/or muscle strength training), either land or water-based, for people with RA. To do this we updated a previous Cochrane review (van den Ende 1998) and made categories for the different forms of dynamic exercise programs.

Search methods

A literature search (to December 2008) within various databases was performed in order to identify randomised controlled trials (RCTs).

Selection criteria

RCTs that included an exercise program fulfilling the following criteria were selected: a) frequency at least twice weekly for > 20 minutes; b) duration > 6 weeks; c) aerobic exercise intensity > 55% of the maximum heart rate and/or muscle strengthening exercises starting at 30% to 50% of one repetition maximum; and d) performed under supervision. Moreover, the RCT included one or more of the following outcome measures: functional ability, aerobic capacity, muscle strength, pain, disease activity or radiological damage.

Data collection and analysis

Two review authors independently selected eligible studies, rated the methodological quality, and extracted data. A qualitative analysis (best-evidence synthesis) was performed and, where appropriate, a quantitative data analysis (pooled effect sizes).

Main results

In total, eight studies were included in this updated review (two additional studies). Four of the eight studies fulfilled at least 8/10 methodological criteria. In this updated review four different dynamic exercise programs were found: (1) short-term, land-based aerobic capacity training, which results show moderate evidence for a positive effect on aerobic capacity (pooled effect size 0.99 (95% CI 0.29 to 1.68). (2) short-term, land-based aerobic capacity and muscle strength training, which results show moderate evidence for a positive effect on aerobic capacity and muscle strength (pooled effect size 0.47 (95% CI 0.01 to 0.93). (3) short-term, water-based aerobic capacity

training, which results show limited evidence for a positive effect on functional ability and aerobic capacity. (4) long-term, land-based aerobic capacity and muscle strength training, which results show moderate evidence for a positive effect on aerobic capacity and muscle strength. With respect to safety, no deleterious effects were found in any of the included studies.

Authors' conclusions

Based on the evidence, aerobic capacity training combined with muscle strength training is recommended as routine practice in patients with RA.

PLAIN LANGUAGE SUMMARY

Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis

This summary of a Cochrane review presents what we know from research about the effect of exercise on Rheumatoid Arthritis (RA).

The review shows that in people with rheumatoid arthritis:

- Aerobic exercise and muscle strength training on land probably improve pain and physical function slightly in the short term.

- There were no harmful side effects (such as increased pain or damage to your joints) of exercise found in this review. This was true for exercising on land or in the water, although most of the studies were not long enough to tell if exercise might cause joint damage.

What is dynamic exercise and what is rheumatoid arthritis?

Dynamic exercise therapy programs means activities with enough intensity, duration, and frequency to improve stamina or muscle strength. Exercise can be any activity that enhances physical fitness. Exercise which gives you more energy, endurance or stamina is often called aerobic exercise. People exercise for many different reasons including weight loss, strengthening muscles and for general fitness.

When you have rheumatoid arthritis, your immune system, which normally fights infection, attacks the lining of your joints. This makes your joints swollen, stiff and painful. The small joints of your hands and feet are usually affected first. There is no cure for RA at present, so treatments such as exercise aim to relieve pain and stiffness and improve your ability to move.

Best estimate of what happens to people with rheumatoid arthritis who take part in a short term land-based dynamic exercise program:

Pain (higher scores mean worse or more severe pain)

- People who took part rated their pain to be about half a point lower on a scale of 0 to 10 after 12 weeks (6% absolute improvement).
- People who took part in a dynamic exercise program rated their pain to be about half a point on a scale of 0 to 10.
- People who did not exercise rated their pain to be 1 on a scale of 0 to 10.

Physical Function (higher score means worse physical function)

- People who took part rated their physical function to be about half a point lower on a scale of 0 to 3 after 12 weeks (6% absolute improvement).
- People who took part in a dynamic exercise program rated their physical function to be about 1.5 on a scale of 0 to 3.
- People who did not exercise rated their physical function to be 1 on a scale of 0 to 3.

Summary of findings for the main comparison. Short-term land-based aerobic capacity and muscle strength training for patients with Rheumatoid Arthritis

short-term land-based aerobic capacity and muscle strength training for patients with Rheumatoid Arthritis

Patient or population: patients with Rheumatoid Arthritis
Settings: hospital, outpatient (rheumatology) clinics
Intervention: short-term land-based aerobic capacity and muscle strength training

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	short-term land-based aerobic capacity and muscle strength training				
Functional ability HAQ . Scale from: 0 to 3. Follow-up: mean 12 weeks	The mean functional ability in the control groups was **0.16 points**	The mean Functional ability in the intervention groups was **0.54 standard deviations lower** (1.11 lower to 0.02 higher)		50 (1 study)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: HAQ -6%, relative % change: HAQ -25%, NNT: n.a., SMD: -0.4 (-0.86 to 0.06)
Muscle strength Isometric extension Follow-up: mean 12 weeks	The mean muscle strength in the control groups was **-2.7 Nm**	The mean Muscle strength in the intervention groups was **0.47 standard deviations higher** (0.01 to 0.93 higher)		74 (2 studies)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: isometric extension 17%, relative % change: isometric extension 19%, NNT: 45, SMD: 0.47 (0.01 to 0.93)
Self-reported pain VAS. Scale from: 0 to 10. Follow-up: mean 12 weeks	The mean self-reported pain in the control groups was **0.9 cm**	The mean Self-reported pain in the intervention groups was **0.53 standard deviations lower** (1.09 lower to 0.04 higher)		50 (1 study)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: VAS 6%, relative % change: VAS -21%, NNT: n.a., SMD: -0.53 (-1.09 to 0.04)
Disease activity	See comment	See comment	Not estimable	74 (2 studies)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change:ESR -33%/ swollen joints -33%, absolute % change:ESR -51%/ swollen joints

						-36%,NNT:n.a., SMD:statistical heterogeneity,pooling data not possible
Radiological damage - not measured	See comment	See comment	Not estimable	-	See comment	Was not measured in included studies

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Small patient number

Summary of findings 2. Short-term land-based aerobic capacity training for patients with Rheumatoid Arthritis

Short-term land-based aerobic capacity training for patients with Rheumatoid Arthritis

Patient or population: patients with Rheumatoid Arthritis
Settings: hospital, outpatient (rheumatology) clinics
Intervention: Short-term land-based aerobic capacity training

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Short-term land-based aerobic capacity training				
Functional ability AIMS . Scale from: 0 to 10. Follow-up: mean 12 weeks	The mean functional ability in the control groups was **0.9 points** [1]	The mean Functional ability in the intervention groups was **0.06 standard deviations lower** (1.33 lower to 1.2 higher)		56 (1 study)	⊕⊕⊝⊝ **low** [2]	Absolute % change: AIMS -22%, relative % change: AIMS -43%, NNT: n.a., SMD: 0.03 (-0.46 to 0.51)

Muscle strength Isometric extension Follow-up: mean 12 weeks	The mean muscle strength in the control groups was **11 foot points**	The mean Muscle strength in the intervention groups was **0.38 standard deviations lower** (1.67 lower to 0.9 higher)		10 (1 study)	⊕⊕⊖⊖ **low** [2]	Absolute % change: isometric extension 22%, relative % change: isometric extension 18%, NNT: n.a., SMD: -0.38 (-1.67 to 0.9)
Self-reported pain AIMS. Scale from: 0 to 10. Follow-up: mean 12 weeks	The mean self-reported pain in the control groups was **-0.7 points**	The mean Self-reported pain in the intervention groups was **0.27 standard deviations lower** (0.79 lower to 0.26 higher)		56 (1 study)	⊕⊕⊕⊖ **moderate** [2]	Absolute % change: AIMS -23%, relative % change: AIMS -10%, NNT: n.a., SMD: -0.27 (-0.79 to 0.26)
Disease activity	See comment	See comment	Not estimable	0 (3 studies)	⊕⊕⊖⊖ **low** [2]	Absolute % change: ERS 5%/ tender joints -36%, relative % change:ESR 19%/ tender joints -29%, NNT:n.a., SMD: statistical heterogeneity, pooling data not possible
Radiological damage - not measured	See comment	See comment	Not estimable	-	See comment	Was not assessed in included studies

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Mean change from baseline in control group
[2] Small patient numbers

Summary of findings 3. Short-term water-based aerobic capacity training for patients with Rheumatoid Arthritis

Short-term water-based aerobic capacity training for patients with Rheumatoid Arthritis

Patient or population: patients with Rheumatoid Arthritis
Settings: hospital, outpatient (rheumatology) clinics
Intervention: Short-term water-based aerobic capacity training

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Short-term water-based aerobic capacity training				
Functional ability outcome was measured on different scales in different studies Follow-up: mean 11 days	See comment	See comment	Not estimable	88 (2 studies)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: HAQ -12%/AIMS -24%, relative % change: HAQ 7%/AIMS -43%, NNT: n.a., SMD: statistical heterogeneity, pooling data not possible
Muscle strenght Grip strength Follow-up: mean 11 weeks	The mean muscle strenght in the control groups was **11.3 Nm**	The mean Muscle strenght in the intervention groups was **0.38 standard deviations lower** (1.27 lower to 0.51 higher)		20 (1 study)	⊕⊕⊖⊖ **low** [1]	Absolute % change: grip strength 15%, relative % change: grip strength -24%, NNT: n.a., SMD: -0.38 (-1.27 to 0.51)
Self-reported pain AIMS. Scale from: 0 to 10. Follow-up: mean 12 weeks	The mean self-reported pain in the control groups was **-0.7 points**	The mean Self-reported pain in the intervention groups was **0.06 standard deviations higher** (0.43 lower to 0.54 higher)		68 (1 study)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: AIMS -12%, relative % change: AIMS 2%, NNT: n.a., SMD: 0.06 (-0.43 to 0.54)
Disease activity Follow-up: mean 11 weeks	See comment	See comment	Not estimable	88 (2 studies)	⊕⊕⊕⊖ **moderate** [1]	Absolute % change: active joints -44%, relative % change: active joints -40%, NNT: n.a., SMD: statistical heterogeneity, pooling of data not possible
Radiological damage - not measured	See comment	See comment	Not estimable	-	See comment	Was not measured in included studies

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Small patient number

Summary of findings 4. Long-term land-based aerobic capacity and muscle strength training for patients with Rheumatoid Arthritis

Long-term land-based aerobic capacity and muscle strength training for patients with Rheumatoid Arthritis

Patient or population: patients with patients with Rheumatoid Arthritis
Settings: hospital, outpatient (rheumatology) clinics
Intervention: Long-term land-based aerobic capacity and muscle strength training

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Long-term land-based aerobic capacity and muscle strength training				
Functional ability outcome was measured on different scales in different studies Follow-up: mean 24 months	See comment	See comment	Not estimable	305 (2 studies)	⊕⊕⊕⊕ **high**	Absolute % change: HAQ 74%/ MACTAR 7%, relative % change: HAQ 50%/MACTAR 0%, NNT: n.a., SMD: due to conflicting evidence pooling of data was not possible
Muscle strength Isometric extension Follow-up: mean 24 months	The mean muscle strength in the control groups was **15.3 Nm**	The mean Muscle strength in the intervention groups was **0.49 standard deviations higher** (0.06 lower to 1.04 higher)		305 (2 studies)	⊕⊕⊕⊕ **high**	Absolute % change: isometric extension 16%, relative % change: isometric extension 10%, NNT: n.a., SMD: 0.49 (-0.06 to 1.04)

Self-reported pain VAS. Scale from: 0 to 10. Follow-up: mean 24 months	The mean self-reported pain in the control groups was **0 cm**	The mean Self-reported pain in the intervention groups was **0.35 standard deviations higher** (0.46 lower to 1.16 higher)		24 (1 study)	⊕⊕⊖⊖ **low** [1]	Absolute % change: VAS 11%, relative % change: VAS 11%, NNT: n.a., SMD: 0.35 (-0.46 to 1.16)
Disease activity DAS Follow-up: mean 24 months	The mean disease activity in the control groups was **-0.7 score**	The mean Disease activity in the intervention groups was **0.14 standard deviations lower** (0.38 lower to 0.09 higher)		281 (1 study)	⊕⊕⊕⊕ **high**	Absolute % change: ESR -15%/DAS -17%, relative % change: ESR: -40%/DAS -6%, NNT: n.a., SMD: -0.16 (-0.39 to 0.06)
Radiological damage Joint score radiographics Follow-up: mean 12 weeks	The mean radiological damage in the control groups was **4 points**	The mean Radiological damage in the intervention groups was **0.15 standard deviations lower** (0.37 lower to 0.08 higher)		305 (2 studies)	⊕⊕⊕⊕ **high**	Absolute % change: joint score 0%, relative % change: joint score: 0%, NNT: n.a., SMD: -0.15 (-0.37 to 0.08)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Small patient number

BACKGROUND

Exercise therapy is a regular component in the physical therapy treatment of patients with rheumatoid arthritis (RA) since it has various health benefits. Apart from general effects including a reduced risk of coronary heart disease, diabetes, hypertension, and colon cancer, a number of specific health benefits have been described for RA, such as improved functional ability (De Jong 2003; Macera 2003; Pate 1995; Van den Ende 1996).

Exercise is recommended in various monodisciplinary guidelines, such as the Canadian Ottawa Panel evidence-based clinical practice guidelines for therapeutic exercises in the management of RA, and in multidisciplinary guidelines such as the EULAR recommendations and the ACR guideline for the management of (early) RA (ACR 2002; Combe 2007; Ottawa 2004).

Exercise therapy can be performed at different intensity levels. Dynamic exercise therapy can be defined as exercise therapy with a sufficient intensity, duration, and frequency to establish improvement in aerobic capacity or muscle strength, or both (Pollock 1998).

In a Cochrane review by Van den Ende et al 1998 it was concluded that, for individuals with RA, dynamic exercise therapy had a positive effect on aerobic capacity, muscle strength, and joint mobility with no detrimental effects on disease activity or pain (van den Ende 1998). Since then, a number of reviews on exercise therapy in RA have been published (Gaudin 2007; Hakkinen 2004; Stenström 2003). In these reviews no explicit inclusion criteria were used regarding the characteristics of the exercise programs, for example intensity, duration, frequency, and supervision of exercises (Hakkinen 2004; Stenström 2003). Moreover, radiological damage was not considered as a safety parameter (Hakkinen 2004; Stenström 2003). In addition, the methodological quality of the studies was not taken into account (Hakkinen 2004; Stenström 2003), no systematic quantitative or qualitative data analysis was applied (Gaudin 2007; Hakkinen 2004; Stenström 2003), and the heterogeneity of the interventions was not addressed in the analyses (Gaudin 2007; Hakkinen 2004; Stenström 2003).

Our objective was, therefore, to update the review by van den Ende et al 1998 (van den Ende 1998) by assessing and summarizing the available evidence up to 2009 on the effectiveness and safety of dynamic exercise therapy for people with RA using the Cochrane Collaboration methodologies (Higgins 2008).

To address the heterogeneity of the interventions, a distinction was made between (a) short-term and long-term exercise therapy programs, (b) land-based and water-based exercise therapy programs, and (c) exercises aimed to improve aerobic capacity, muscle strength, or both.

OBJECTIVES

The primary aim was to determine the effectiveness (regarding functional ability, aerobic capacity, and muscle strength) and the safety (regarding self-reported pain, disease activity, and radiological damage) of dynamic exercise therapy (land-based and water-based) in people with RA.

METHODS

Criteria for considering studies for this review

Types of studies

RCTs comparing exercise therapy in people with RA with another form of exercise therapy or with a non-exercising control group were selected. Studies were identified as randomised if the treatment assignment was described with words such as randomly, at random, or randomisation.

Types of reports

Only full-length articles or full-length unpublished reports were considered for inclusion in this review. Also, due to limited resources for translation, only articles in English, Dutch, French, or German were considered for inclusion.

Types of participants

Trials including participants (males and females > 18 years old) with classical or definite RA according to the 1958 ARA criteria (Ropes 1958) or the 1987 ARA criteria (Arnett 1987) were selected. Other types of arthritis were not eligible for this review.

Types of interventions

Trials with an exercise program assumed to be adequate to improve aerobic capacity, muscle strength, or both according to the following criteria.

- Exercise frequency of at least 20 minutes twice a week.
- Duration of exercise program at least six weeks (duration < three months was considered short-term; duration > three months was considered long-term).
- Exercise program performed under supervision.
- Aerobic exercise intensity at least 55% of the maximum heart rate (HR max); or intensity starting at 40% to 50% of the maximum oxygen uptake reserve (VO_2R) or HR max reserve (HRR). Furthermore, the intensity is increased up to 85% during the intervention.
- Progressively strengthening exercise loads starting at 30% to 50% and increasing to 80% of maximum (defined as the percentage of either one repetition maximum, one maximum voluntary contraction, maximum speed, or as maximal subjective exertion) (Pollock 1998).

Types of outcome measures

Outcome measures were selected based on their confirmed properties in terms of reliability, validity, sensitivity to change, or general acceptance in the literature. Any trials reporting data on one or more of the primary or secondary outcome measures were included (see below). Functional ability, aerobic capacity and muscle strength were considered as outcome measures for effectiveness. Self-reported pain, adverse events, disease activity, and radiological damage were considered as outcome measures for safety. If improved, these outcome measures (except for adverse events) were also indicators for effectiveness.

Primary outcomes

1. Functional ability assessed by validated questionnaires, preferably the McMaster Toronto Arthritis Patient Preference

Interview (MACTAR) (Verhoeven 2000), the Health Assessment Questionnaire (HAQ) (Spitz 1987), the physical function dimension of the Arthritis Impact Measurement Scales (AIMS) (Meenan 1980) or any other validated health related quality of life (QoL) measure comprising a functional ability dimension, like the Short Form-36 (SF-36), the Fries Index (Fries 1980), or the Functional Status Index (FSI) (Jette 1980).

2. Self-reported pain, preferably scored on a visual analogue scale (VAS).

3. Adverse effects (like exercise-induced injuries, substantially increased pain levels as a direct effect of treatment).

Secondary outcomes

1. Aerobic capacity (VO_2max, in ml/kg/min), determined by a maximal or submaximal ergometer test.

2. Muscle strength of the knee extensors (in Nm), preferably isokinetic muscle strength assessed by means of an isokinetic dynamometer; or isometric muscle strength measured with an isokinetic dynamometer.

3. Disease activity, preferably assessed by the Westergren erythrocyte sedimentation rate (ESR) (in mm/h), C-reactive protein (CRP), a swollen joint count, or a composite index such as the disease activity score (DAS).

4. Radiological joint damage of small or large joints, or both.

5. Ratio of the costs divided by the effectiveness, preferably assessed by utilities such as the EuroQol (EQ-5D) questionnaire, the SF-36-derived utility index (called the SF-6D), the time trade off (TTO) method, or the Standard Gamble (SG) (Hurst 1997; Marra 2005; Tijhuis 2000).

Search methods for identification of studies

Electronic searches

The Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library*), MEDLINE, EMBASE, CINAHL, PEDro, and Web of Science databases were searched (1997 to December 2008) to identify all RCTs relating to exercise therapy in RA. The search strategy used in the review by van den Ende et al (van den Ende 1998) was updated and expanded by adding the terms physical fitness, hydrotherapy, and water therapy. This new search strategy was repeated (to 1997) to ensure that no trials published before 1997 were missed.

The following strategy was used for the identification of RCTs: (a) MEDLINE (May 1997 to December 2008) was searched using an optimal, sensitive MEDLINE search strategy for identifying randomised controlled trials (Dickerson 1994; Robinson 2002) combined with the terms (rheumatoid arthritis OR arthritis[tiab]) AND (exercise therapy OR exercis*[tiab] OR motion therap*[tiab] OR "physical education and training"[MeSH Terms] OR physical education[tiab] OR training[tiab] OR gymnast*[tiab] OR physical fitness OR physical fitness[tiab] OR hydrotherap* OR hydrotherap*[tiab] OR water therap* OR water therap*[tiab]) (Appendix 1); (b) the same search strategy was translated by an experienced medical librarian to make it applicable to Cochrane Central Register of Controlled Trials (to December 2008), EMBASE (October 1996 to December 2008) (Appendix 2), the Web of Science database (1981 to December 2008) (Appendix 3), CINAHL (to December 2008), and PEDro (to December 2008).

Searching other resources

Reference lists of the identified RCTs were scanned; (d) for abstracts without published full-length articles, the authors were contacted by mail asking for a written report; (e) the Current Controlled Trials register was checked for ongoing trials (www.controlled-trials.com/). Authors of ongoing trials were contacted asking for any preliminary results. All search strategies were performed by the principal review author (EJH).

Timeline and update

The literature search was first performed in December 2007, the qualitative and quantitative analyses were performed between January and March 2007, and the review was completed and submitted to *The Cochrane Library* in June 2008.

Since December 2007, the review author has regularly executed the search strategy (to December 2008) to make sure that no new trials have been published. The review author will continue to run the search over at least the next two years and the review will be updated if there is new information.

Data collection and analysis

Selection for inclusion in the review, assessment of the methodological quality of the trials, and data extraction were performed in three separate steps. The articles were independently assessed by two review authors (EJH, FG). The review authors were blinded to the journal, the authors, and the institution by providing edited copies of the articles. Disagreement between the two review authors was resolved by a third review author (TVV).

Selection of trials

Firstly, all titles and abstracts were screened for the two following criteria: the study included participants with RA, and the intervention was a form of exercise therapy. Secondly, the full-length articles of the selected titles and abstracts were gathered and screened for the full set of inclusion criteria.

Assessment of methodological quality

The methodological quality of all studies was independently assessed by two review authors (EJH, FG). Disagreement was resolved by discussion. If no consensus was achieved, a third review author (TVV) was the adjudicator. The 8-point scale of methodological criteria used by Verhagen 1998 (Verhagen 1998) was adjusted and combined with criteria considered by the review authors to be relevant for the purpose of this review. The criteria 'patient blinded' and 'care provider blinded' were not considered suitable because of the characteristics of the intervention of interest and were removed from the list. Four criteria were added to the list (criteria 4, 6, 7, and 8, see below). The final list included the following items:

1. adequate allocation concealment;

2. groups similar at baseline;

3. specification of eligibility criteria;

4. sufficient description of intervention;

5. blinding of outcome assessor;

6. comparable co-interventions per group;

7. description of rate and reasons for dropping out;

8. description of rate of compliance;

9. presentation of point estimates and measures of variability;

10. intention-to-treat analysis.

All selected methodological criteria were scored as yes, no, or unclear. Equal weight was applied to all items resulting in a total methodological score ranging from 0 to 10; a methodological score of 8 or higher was regarded as high methodological quality. The agreement between the review authors regarding the methodological quality of the included studies was calculated using the kappa coefficient (+1: perfect agreement, 0: no agreement above that expected by chance, -1: complete disagreement) (Fleiss 1973).

Furthermore, to assess the risk of bias according to the new Cochrane method, the following six key domains were reported by two review authors: adequate sequence generation; allocation concealment; blinding (yes, when the assessor was blinded); incomplete outcome data; free of selective reporting; and free of other bias?

Data extraction and analysis

Two review authors independently assessed the content of the interventions of the included RCTs and categorized the RCT. First a distinction was made between short and long-term exercise programs and secondly the type of intervention was assessed as follows.

A. Intervention concerns aerobic capacity training (land based).
B. Intervention concerns aerobic capacity and muscle strength training (land based).
C. Intervention concerns aerobic capacity training (water based).
D. Intervention concerns aerobic capacity and muscle strength training (water based).
There were no studies of muscle strength training only, land based or water based.

Extracted data were recorded in a structured form. The following descriptive information was systematically extracted: number of RA participants, sex, age (years), disease duration (years), inclusion and exclusion criteria (for example criteria with regard to the amount of disease activity); description of the intervention in the dynamic exercise group and the control group; and type of intervention (A, B, C, or D). If multiple exercise interventions were compared in the one study, the review authors designated the exercise program(s) with an intensity level fulfilling the criteria for dynamic exercise as the dynamic group(s) whereas the exercise program with the lowest exercise intensity level was designated as the control group.

For continuous outcome measures, baseline values and measures of variability (preferably given as the mean and standard deviation (SD)) were extracted. If possible, mean changes from baseline were obtained directly after the intervention and, if reported, at follow-up. Means and SD were converted from medians and interquartile ranges, if necessary. For dichotomous outcomes, the number of patients incurring the event of interest was extracted for each group.

Analysis

For continuous variables, a standardized mean difference (effect size) was calculated (mean difference intervention group - mean difference control group)/pooled SD for each trial. For dichotomous variables, odds ratios with corresponding 95% confidence intervals (CI) were computed (odds of event in intervention group/odds of event in control group) per trial. For multiple time measurements, the measurements closest to the end of the intervention and the end of follow-up were included in the analyses.

Qualitative analysis

For this review a decision was made to always apply a qualitative approach (qualitative synthesis). For this qualitative analysis the strength of the evidence was defined as follows (van Tulder 2003).

Strong evidence: provided by generally consistent findings in multiple relevant, high-quality RCTs (methodological score eight or more).

Moderate evidence: provided by generally consistent findings in one relevant, high-quality RCT; or more relevant low-quality RCTs (methodological score < eight).

Limited evidence: provided by generally consistent findings in one or more relevant low-quality RCTs.

No or conflicting evidence: no RCTs available or the results were conflicting.

The outcomes of the studies were considered 'consistent' if at least 75% of the trials reported statistically significant results in the same direction, within the intervention and control groups.

Quantitative analysis

If appropriate (studies comparable with statistical homogeneity), a quantitative approach (for example meta-analysis) was applied. In the case of meta-analysis, the pooled standardized mean difference was computed using a random-effects model (Higgins 2008). If continuous as well as dichotomous data were presented in the included studies for the same outcome measure, the dichotomous data were first re-expressed as standardized mean differences (Chinn 2000) and then pooled. An effect size of less than 0.2 was classified as small, 0.2 to 0.8 as medium, and those higher than 0.8 as large (Cohen 1988). The quantitative approach was not considered to be appropriate if the included studies were clinically diverse or were statistically heterogeneous. Clinical diversity among studies was assessed by two review authors (EJH, FG) taking into account the classification of participants (severity of the disease) and the intervention (duration, frequency, and setting). Disagreement was resolved by discussion. Statistical heterogeneity was tested by applying the I^2 statistic. When I^2 was greater than 50% we considered this to indicate substantial heterogeneity.

The results from the qualitative and quantitative analyses were also included in summary of findings tables, according to the new Cochrane method; the included studies were graded according to the GRADE approach (Appendix 4).

Sensitivity analyses

For the sensitivity analyses we qualitatively and quantitatively analysed the results including only the high quality studies (methodological quality score > 80%).

Subgroup analyses

If possible, subgroup analyses were performed for men and women; young and adulthood to old age; and individuals with high and low disease activities or various degrees of joint damage.

RESULTS

Description of studies

Selection of included trials

The new search strategy was applied (to May 1997) to be sure that no trials were missed in the first review by van den Ende. This search strategy yielded no other trials than the ones identified previously (Baslund 1993; Hansen 1993; Harkcom 1985; Lyngberg 1994; Minor 1989; Van den Ende 1996).

The search strategy was again applied (May 1997 to February 2009). This resulted in a list of 827 citations. Based on the titles, abstracts, or both, a total of 32 citations were selected (Bearne 2002; Bell 1998; Bilberg 2005; Bostrom 1998; Buljina 2001; Callahan 2004; De Jong 2003; De Jong 2004a; De Jong 2004b; de Jong 2005; Hafstrom 2005; Hakkinen 1997; Hakkinen 1999; Hakkinen 2001; Hakkinen 2003a; Hakkinen 2003b; Hakkinen 2004a; Hakkinen 2004b; Hall 2004; Kauppi 1998; Komatireddy 1997; Marcora 2005; McMeeken 1999; Munneke 2005; Neuberger 1997; Neuberger 2000; O'Brien 2006; Suomi 1997; Suomi 2003; van den Ende 2000; van den Hout 2005; Westby 2000). In two cases only an abstract was available (Callahan 2004; Neuberger 2000). The authors of these two abstracts were contacted and asked for a written report. In one case (Callahan 2004) no full written report was available; in the second, the author (Neuberger 2000) did not reply.

The 31 remaining full-length articles were screened for the full set of inclusion criteria; one trial was described in six full-length articles (De Jong 2003; De Jong 2004a; De Jong 2004b; De Jong 2009; Munneke 2005; van den Hout 2005). The other 25 full-length articles were excluded, mainly because: the exercise program was not supervised, the intensity and duration of the exercise program did not meet the inclusion criteria, or the trial was not randomised.

After searching the references of the selected citations, one other trial that fulfilled the inclusion criteria was identified (Sanford-smith 1998).

No ongoing RCTs were found in the Current Controlled Trials register.

In total, eight RCTs were included (see table 'Characteristics of included studies') with two new trials added to the six included in the previous review (van den Ende 1998).

The 25 articles that were screened but not included in the review are listed in the table 'Characteristics of excluded studies'.

Description of included trials

The eight included trials involved 575 participants in total. Mainly female participants participated in six of the eight included trials (Baslund 1993; De Jong 2003; Hansen 1993; Harkcom 1985; Minor 1989; Sanford-smith 1998) whereas in one trial the proportions of females and males were equal (Van den Ende 1996) and in one trial more males than females were included (Lyngberg 1994). The average age of the participants was around 52 years in most trials, except for two trials where the average age was higher: 67 years (Lyngberg 1994) and 62 years (Sanford-smith 1998), respectively. The average disease duration varied and was between 5 to 14 years, except for one trial where the average disease duration of the participants in one of the treatment arms was 20 years (Sanford-smith 1998). In all trials participants with low to medium disease activity were included and patients with serious co-morbidities were excluded.

In four studies (Harkcom 1985; Lyngberg 1994; Sanford-smith 1998; Van den Ende 1996) RA patients were selected from the outpatient clinic; in one study participants were selected from the outpatient clinic and from the community (Minor 1989); and in three studies participants were selected from the hospital registry (Baslund 1993; De Jong 2003; Hansen 1993).

Two different comparisons were evaluated. In five trials one dynamic exercise program was compared to: no exercise (Baslund 1993; Lyngberg 1994); ROM exercises (Sanford-smith 1998); ROM exercises combined with isometric exercises or written instructions (Van den Ende 1996); or physical therapy, if considered necessary (De Jong 2003). In three trials multiple dynamic exercise programs were compared to no exercise (Hansen 1993; Harkcom 1985) or ROM exercises (Minor 1989).

In four studies the dynamic exercise program was supervised by physical therapists (Hansen 1993; Lyngberg 1994; Sanford-smith 1998; Van den Ende 1996); in one study by physical education graduate students (Harkcom 1985); and in three studies the profession of the supervisors was not mentioned (Baslund 1993; De Jong 2003; Minor 1989). None of the studies described whether the supervisors received special training.

Risk of bias in included studies

Four trials, all conducted in the year 1994 or later, met eight or more of the 10 methodological criteria (De Jong 2003; Lyngberg 1994; Sanford-smith 1998; Van den Ende 1996).

The following criteria were not fulfilled in four or more of the RCTs: adequate allocation concealment (n = 7), intention-to-treat analysis (n = 5), blinded outcome assessor (n = 4), and reporting rate of compliance (n = 2). Of the RCTs with high methodological quality the criteria regarding: adequate allocation concealment was not fulfilled in three studies; intention-to-treat analysis in two studies; blinded outcome assessor in one study; and groups similar at baseline in one study (see the table 'Characteristics of included studies' for detailed information on risk of bias).

Effects of interventions

See: **Summary of findings for the main comparison** Short-term land-based aerobic capacity and muscle strength training for patients with Rheumatoid Arthritis; **Summary of findings 2** Short-term land-based aerobic capacity training for patients with Rheumatoid Arthritis; **Summary of findings 3** Short-term water-based aerobic capacity training for patients with Rheumatoid Arthritis; **Summary of findings 4** Long-term land-based aerobic

capacity and muscle strength training for patients with Rheumatoid Arthritis

Short-term aerobic capacity training (land based)

(Table 1 and Summary of findings 2)

Three trials were concerned with a short-term dynamic exercise program that included exercises to improve aerobic capacity (Baslund 1993; Harkcom 1985; Minor 1989). One study (Minor 1989) included a follow-up period of nine months. None of these trials had high methodological quality.

Best-evidence synthesis

Concerning the effectiveness immediately after the intervention, no statistically significant effect on functional ability was found in any of the trials. In two of the trials (Baslund 1993; Minor 1989) a significant positive effect on aerobic capacity was seen ($P < 0.05$) immediately after the intervention. No statistically significant effect was found in any of the trials for muscle strength. In the one study including follow-up (Minor 1989), no significant effects of the intervention on functional ability, aerobic capacity, or muscle strength were found nine months after the intervention.

With respect to safety, immediately after the intervention or at follow-up, no significant effects of the intervention were found on self-reported pain or disease activity in any of the three studies. None of the studies mentioned adverse effects or measured radiological damage. Moreover, none of the three trials undertook a cost-effectiveness analysis.

Quantitative analysis

Concerning the effectiveness, data could be pooled for functional ability (Baslund 1993; Harkcom 1985; Minor 1989) and aerobic capacity (Harkcom 1985; Minor 1989). Pooling resulted in a non-significant trend for a positive effect on functional ability (pooled effect size (ES) 0.03, 95% CI -0.46 to 0.51) and a large, significant positive effect for aerobic capacity (pooled ES 0.99, 95% CI 0.29 to 1.68) immediately after the intervention (relative percent difference 27%), when compared to the control groups. Data could not be pooled for muscle strength due to a lack of studies.

Concerning safety, no outcome measure data could be pooled due to a lack of studies (self-reported pain, radiological damage) and statistical heterogeneity (disease activity).

From the available data it can be concluded that there is limited evidence that short-term, land-based aerobic capacity training has no effect on functional ability or muscle strength measured directly after the intervention, or at follow-up. There is moderate evidence that short-term, land-based aerobic capacity training has a positive effect on aerobic capacity directly after the intervention but not at follow-up. In addition, there is limited, moderate evidence that short-term, land-based aerobic capacity training has no detrimental effects on disease activity or self-reported pain. No evidence was found regarding the effects of the intervention on radiological damage or cost effectiveness.

Short-term aerobic capacity and muscle strength training (land based)

(Table 2 and Summary of findings for the main comparison)

Two trials were concerned with a short-term dynamic exercise program including exercises to improve both aerobic capacity and muscle strength (Lyngberg 1994; Van den Ende 1996). One study (Van den Ende 1996) included a follow-up period of three months. Both studies had high methodological quality.

Best-evidence synthesis

With respect to the effectiveness immediately after the intervention, no statistically significant effects were found in either of the included trials for functional ability whereas in one trial (Van den Ende 1996) a significant positive improvement of aerobic capacity ($P < 0.001$) and muscle strength ($P < 0.05$) was seen in the intervention group. In the study that included follow-up (Van den Ende 1996), no significant effects of the intervention were reported for functional ability, aerobic capacity, or muscle strength at follow-up.

With regard to safety, one of the two trials (Van den Ende 1996) reported a significant positive effect of the intervention on disease activity ($P < 0.05$) immediately after the intervention, but not at follow-up. Neither of the trials observed an effect of the intervention on self-reported pain immediately after the intervention or at follow-up. In one trial (Lyngberg 1994) it was reported that no exercise injuries and no heart-related problems were noted during submaximal tests or when exercising. In one trial (Van den Ende 1996) it was reported that no adverse effects were found. Radiological damage was not measured in either trial and no cost-effectiveness analysis was carried out.

Quantitative analysis

When measuring effectiveness, data could only be pooled for functional ability (Lyngberg 1994; Van den Ende 1996) and muscle strength (Lyngberg 1994; Van den Ende 1996). This resulted in a non-significant trend for a positive effect on functional ability (ES -0.40, 95% CI -0.86 to 0.06) and a significant, medium positive effect on muscle strength (ES 0.47, 95% CI 0.01 to 0.93) immediately after the intervention when compared to the control groups (relative percent difference 22%). Data could not be pooled for aerobic capacity due to statistical heterogeneity.

For safety, no outcome measure data could be pooled due to a lack of studies (self-reported pain, radiological damage) and statistical heterogeneity (disease activity).

From the available data it can be concluded that there is moderate evidence that short-term, land-based aerobic capacity and muscle strength training has no effect on functional ability measured immediately after the intervention and at follow-up. However, a positive effect on aerobic capacity and muscle strength was seen immediately after the intervention, but not at follow-up. There is moderate evidence that short-term, land-based aerobic capacity and muscle strength training does not have detrimental effects on self-reported pain and has a positive effect on disease activity immediately after the intervention, but not at follow-up. No conclusions can be drawn regarding radiological damage or the cost effectiveness of this intervention.

Short-term aerobic capacity training (water based)

(Table 3 and Summary of findings 3)

Two trials were concerned with a short-term, water-based dynamic exercise program that included exercises to improve aerobic

capacity (Minor 1989; Sanford-smith 1998). Neither of these trials included follow-up. One of the trials was of high methodological quality (Sanford-smith 1998).

Best evidence synthesis

For effectiveness, one (Minor 1989) of the two included trials (Minor 1989; Sanford-smith 1998) reported statistically significant positive effects of the intervention on functional ability (P < 0.05) and aerobic capacity (P < 0.05) immediately after the intervention. Neither trial reported an effect on muscle strength.

With respect to safety, no statistically significant effects of the intervention were found immediately after the intervention for self-reported pain or disease activity. Adverse effects were not mentioned in either study and radiological damage was not measured. Neither study included a cost-effectiveness analysis.

Quantitative analysis

With regard to effectiveness, data could only be pooled for aerobic capacity (Minor 1989; Sanford-smith 1998) resulting in a non-significant trend toward a positive effect (ES 0.47, 95% CI -0.04 to 0.98) when compared to the control groups. Data for functional ability and muscle strength could not be pooled due to statistical heterogeneity (functional ability) and a lack of studies (muscle strength).

For safety, no data could be pooled for any outcome measure because of a lack of studies (self-reported pain, radiological damage) and statistical heterogeneity (disease activity).

From the available data it can be concluded that there is limited evidence that short-term, water-based aerobic capacity training has a positive effect on functional ability and aerobic capacity, and no effect on muscle strength measured immediately after the intervention. There is limited, moderate evidence that short-term, water-based aerobic capacity training has no detrimental effects immediately after the intervention on self-reported pain and disease activity. No conclusions can be made for radiological damage and the cost effectiveness of this intervention.

Long-term aerobic capacity and muscle strength training (land based)

(Table 4 and Summary of findings 4)

Two trials investigated a long-term, land-based dynamic exercise program that included exercises to improve both aerobic capacity and muscle strength (De Jong 2003; Hansen 1993). One of the trials was of high methodological quality (De Jong 2003) and included a cost-effectiveness analysis (van den Hout 2005). This trial included a follow-up period and the results are now published (De Jong 2009). This observational follow-up study involved a selection of participants who had completed the intervention (n = 71). The participants from the intervention group (n = 60) who continued exercising at similar intensity levels as in the original intervention, though on average at a lower frequency, were compared to those participants who did not (n = 11).

Best-evidence synthesis

Concerning the effectiveness, one (De Jong 2003) of the two included trials (De Jong 2003; Hansen 1993) reported significant positive effects in the intervention group on functional ability (P < 0.05) and aerobic capacity (P < 0.001) directly after the intervention. In both trials a significant positive effect was found for muscle strength (P < 0.001, P < 0.05) directly after the intervention. In the one study including a 18-month follow-up (De Jong 2009) the improvement of muscle strength was maintained in those participants from the intervention group who continued exercising at similar intensity levels as in the original intervention (although on average at a lower frequency) but not in those who did not remain physically active.

With regard to safety, no significant effects were found on self-reported pain, disease activity, and radiological damage either immediately after the intervention or after follow-up. One trial (De Jong 2003) reported that there was a non-significant trend toward a greater increase in damage in the large joints in the dynamic exercise group. In the follow-up study (De Jong 2009) radiological damage was not increased in the intervention group participants who continued exercising at approximately similar intensity levels when compared to those who did not. Concerning cost effectiveness, dynamic exercise was not found to be cost effective in the one trial where dynamic exercise therapy was compared to usual care (De Jong 2003).

Quantitative analysis

With regard to effectiveness, data could only be pooled for muscle strength (De Jong 2003; Hansen 1993). The exercise intervention resulted in a non-significant trend toward a positive effect on muscle strength (ES 0.49, 95% CI -0.06, 1.04) when compared to the control groups. Data could not be pooled for functional ability or aerobic capacity due to conflicting evidence (functional ability) and a lack of studies (aerobic capacity).

Looking at safety, data could be pooled for disease activity (De Jong 2003; Hansen 1993) and radiological damage (De Jong 2003; Hansen 1993). Pooling resulted in a non-significant effect on disease activity (ES -0.16, 95% CI -0.39 to 0.06) and radiological damage (ES -0.15, 95% CI -0.37 to 0.08) when compared to the control groups. Data could not be pooled for self-reported pain due to a lack of studies.

From the available data it can be concluded that there is conflicting evidence that long-term aerobic capacity and muscle strength training has a positive effect on functional ability directly after the intervention. There is moderate and limited evidence that long-term aerobic capacity and muscle strength training has positive effects on aerobic capacity and muscle strength, respectively, measured immediately after the intervention. The evidence is limited that long-term aerobic capacity and muscle strength training has no detrimental effect on self-reported pain directly after the intervention. There is moderate evidence that, immediately after the intervention, long-term aerobic capacity and muscle strength training has no detrimental effect on disease activity and radiological damage. In the one trial that included an economic analysis long-term aerobic capacity and muscle strength training was not found to be cost effective when compared to physical therapy, if considered necessary by the treating rheumatologist (van den Hout 2005).

Sensitivity analyses

Four trials met eight or more methodological criteria (De Jong 2003; Lyngberg 1994; Sanford-smith 1998; Van den Ende 1996) and were used for the sensitivity analyses (see Table 5, conclusion 2).

In the subcategory short-term land based aerobic capacity training, none of the three trials fulfilled eight or more criteria so no analysis could be done.

In the subcategory short-term land based aerobic capacity and muscle strength training, both trials fulfilled the criteria of eight or more, so the conclusion remained the same.

In the subcategory short-term aerobic capacity training in water, one of two trials fulfilled eight or more criteria resulting in a different, more negative conclusion: there is moderate evidence that short-term aerobic capacity training in water has no effect on any of the outcome measures directly after the intervention, but it is safe with regard to disease activity.

In the subcategory long-term land based aerobic capacity and muscle strength training, one of two trials fulfilled the eight or more criteria, which resulted in a more positive conclusion: there is moderate evidence that long-term aerobic capacity and muscle strength training has a positive effect on functional ability, aerobic capacity, and muscle strength immediately after the intervention.

Subgroup analysis

In the included studies, no specific data were reported for specific subgroups, for example separate data for men and women or participants with high or low disease activity. As a number of studies were quite old, we considered that it was not feasible to ask for the original data sets.

DISCUSSION

This is the first systematic review, executed according to new Cochrane methodology, summarizing the results of randomised, controlled studies on the effect of dynamic exercise therapy in patients with RA through a qualitative and a quantitative approach.

The training programs in the eight selected studies (two additional trials) included short-term aerobic capacity training (land based as well as water based), aerobic capacity training combined with muscle strength training (land-based), and long-term aerobic capacity training combined with muscle strength training (land based). The results in the previous review indicated that dynamic exercise therapy is effective with regard to aerobic capacity and muscle strength. The results of this updated review suggest that dynamic exercise programs comprising aerobic capacity training (land based as well as water based) has a positive effect on aerobic capacity immediately after the intervention. If performed in water, a positive effect was also seen for functional ability. Dynamic exercise programs consisting of aerobic capacity training combined with muscle strength training (long term as well as short term) have a positive effect on aerobic capacity and muscle strength immediately after the intervention. With long-term programs, a positive effect on functional ability was also seen. No deleterious effects on disease activity, self-reported pain, or radiological damage were found in any of the training programs. In the one trial that included an economic analysis, long-term aerobic capacity and muscle strength training was not cost effective compared to physical therapy, if considered necessary by the treating rheumatologist (De Jong 2003).

In three previous published reviews it was likewise concluded that dynamic exercise therapy improves aerobic capacity and muscle strength and is safe for people with RA (Gaudin 2007; Hakkinen 2004; Stenström 2003). Although these conclusions are comparable to ours, there are a few differences. Dynamic exercise therapy consists of two types of exercises and in our review we made distinct conclusions for these different types. Secondly, the applied methodology of our review was different. In the present review pertinent inclusion criteria were applied regarding the intensity of the intervention, and the heterogeneity of the interventions was taken into account by stratifying the included trials. In addition, a quantitative approach was applied, with a calculation of the size of the effect (van den Ende 2006).

The question remains as to what the optimal duration of a dynamic exercise progam should be. None of the studies included in this review directly compared a short-term and a long-term program, so it remains unclear which dynamic exercise program has the largest effect on functional ability, aerobic capacity, or muscle strength either immediately after the intervention or at follow-up. In this review, three trials, two of short-term programs (Minor 1989; Van den Ende 1996) and one using a long-term program (De Jong 2009), included a follow-up period. In the two trials on short-term programs, participants in the intervention group were advised to continue their exercises after the intervention. In these trials it was found that the effects of the intervention diminished with time (Minor 1989; Van den Ende 1996). In the one study involving 18-months follow-up after a long-term dynamic exercise program (De Jong 2009), a selection of participants from the intervention group were advised to continue their exercises under supervision or engage in other activities of similar intensity. It was found that in those who continued exercising, only the effect on muscle strength remained. This could be due to a decreased exercise frequency as compared to the original intervention. The above mentioned follow-up results may indicate that, in the longer term, the maintenance of dynamic exercise and the ensuing health benefits may prove difficult. On the other hand, more favorable results of long-term adherence to dynamic exercises have been reported. Häkkinen et al (Hakkinen 1999; Hakkinen 2001; Hakkinen 2003a; Hakkinen 2004a; Hakkinen 2004b) found excellent maintenance up to five years after a home-based dynamic exercise program. These conflicting results indicate that more research is needed on the optimal duration of exercise programs, as well as the factors influencing maintenance after termination of the intervention.

With respect to the safety of dynamic exercise therapy, it has long been hypothesized that increasing the level of stress on the joints would increase pain, disease activity, and joint damage. However, in line with other reviews (Gaudin 2007; Hakkinen 2004; Stenström 2003), this review found no increase in pain or disease activity in any of the included trials. In updating the review by van den Ende (van den Ende 1998) we have included two trials on long-term exercise programs in which radiological progression was evaluated (De Jong 2003; Hansen 1993). This means that more firm conclusions could be drawn for safety with regard to radiological damage than in the previous review. In both studies no significant differences in radiological progression were observed between the intervention and the control groups. In the one trial with sufficient power (De Jong 2003) it was reported that there was a non-significant trend toward a greater increase in damage in the large joints in the dynamic exercise group immediately after the intervention.

Concerning the mode of delivery of the exercise program, it has been thought that dynamic exercise therapy in water would be safer than land-based dynamic exercise therapy with respect

to pain, disease activity, and joint damage (Baker 1953; Jivoff 1975). However, in the one study which directly compared land-based aerobic capacity training with water-based aerobic capacity training (Minor 1989) no differences were found with regard to safety. In the absence of high-quality studies indicating the superiority of either land-based or water-based dynamic exercise therapy, the decision on land or water-based exercise therapy should be based on the preferences of the patients themselves or the goals to be reached (Dagfinrud 2007).

In a best-evidence summary of systematic reviews, it was concluded that there was insufficient evidence to draw conclusions with regard to the effectiveness of exercise therapy for RA (Smidt 2005). In that umbrella review, a 100-point quality scoring system was applied for assessing the quality of reviews, with the quality of reviews being graded as 0 to 59 points: low quality, 40 to 59 points: reasonable quality, 60 to 79 points: moderate quality, and 80 or more: good quality. According to that scoring system the quality of the review by van den Ende (van den Ende 1998) was classified as reasonable (74 points). In a second overview of systematic reviews it was concluded that there is low level evidence with regard to the effectiveness of exercise therapy for RA (Christie 2007). The review of Van den Ende et al 1998 was not included in that analysis (only reviews between 2000 and 2007 were included) although in the discussion it was mentioned that by updating that review it would most likely be graded as a high quality review. Indeed, if the same scoring systems was applied to the current review, its quality is likely to be classified as good (Smidt 2005) or as having fewer limitations (Christie 2007) since we included two new trials of high methodological quality, added a qualitative and quantitative analysis, and included more information concerning the context of the intervention. Therefore, it is possible to draw firmer conclusions with regard to the effectiveness and safety of exercise therapy in people with RA.

This systematic review has some limitations. First, only eight studies were included. More studies on this topic have been published but over 50 studies were excluded due to our strict inclusion criteria pertaining to the intensity of the intervention as well as the methodological quality of the studies. As a result, the conclusions of this review are very specific with respect to a clearly defined type of exercise therapy in RA. With the usage of the criteria it is, however, possible that some relevant studies were excluded, such as studies on home-based exercise programs which may have had the same level of intensity (Hakkinen 1999; Hakkinen 2001; Hakkinen 2003a; Hakkinen 2004a; Hakkinen 2004b). Moreover, with respect to the intensity of muscle strength training, one of the inclusion criteria was a minimum duration of 20 minutes, whereas the intensity of muscle strength training is defined according to the number of (maximum) repetitions in some studies. It is possible that these programs did match our inclusion criteria, although it was not explicitly stated in the text. In addition, we excluded trials of moderate intensity exercise programs for which the effectiveness and safety may in part be comparable to dynamic exercise therapy. Secondly, six of the included studies were published more than 10 years ago and the medical and non-medical treatment of RA patients has markedly improved since then. It is, therefore, conceivable that the characteristics of participants included in the older studies may differ from those of current RA patients. Thirdly, the populations studied in the included trials consisted mainly of females with a low to medium level of disease activity and an average disease duration of 5 to 14 years. Therefore, the extent to which the conclusions of this review are also applicable to men and people with recent onset of RA remains unclear.

AUTHORS' CONCLUSIONS

Implications for practice

Short-term, land-based dynamic exercise programs have a positive effect on aerobic capacity (aerobic capacity training whether or not combined with muscle strength training) and muscle strength (aerobic capacity training combined with muscle strength training) immediately after the intervention, but not after a follow-up period. Short-term, water-based dynamic exercise programs have a positive effect on functional ability and aerobic capacity directly after the intervention but it is unknown whether these effects are maintained after follow-up. Long-term, land-based dynamic exercise programs (aerobic capacity and muscle strength training) have a positive effect on functional ability, aerobic capacity, and muscle strength immediately after the intervention but it is unknown whether these effects are maintained after follow-up. Whether short-term dynamic exercise programs are safe with regard to radiological damage and whether these programs are cost effective remains unclear. Long-term dynamic exercise programs appear to be safe with regard to radiological damage but they are not cost effective if compared to physical therapy, if considered necessary by the treating rheumatologist.

Based on the evidence, aerobic capacity training combined with muscle strength training is recommended for routine practice in patients with RA. The optimal duration of the intervention, mode of delivery, and extent of supervision need to be further investigated.

Implications for research

Whether the dynamic exercise program should be under supervision or not is still unknown and should be further investigated. Also, other commonly used dynamic exercise forms (such as flexibility training, stability training, coordination training) and the mode of delivery (land or water-based) should be investigated. Furthermore, dynamic exercise programs should be compared to less intensive programs such as physical activity programs that meet public health recommendations. New trials should compare different exercise and physical activity programs; aspire to an accurate description of the content, dose, and application of the interventions; and make use of standardized, validated outcome measures, including those for active participation, suitable for assessing effectiveness, safety, and cost effectiveness.

ACKNOWLEDGEMENTS

The review authors would like to thank the Cochrane Musculoskeletal Group editorial team for their editorial suggestions.

REFERENCES

References to studies included in this review

Baslund 1993 *{published data only}*

Baslund B, Lyngberg K, Andersen V, Halkjaer Kristensen J, Hansen M, Klokker M, et al. Effect of 8 wk of bicycle training on the immune system of patients with rheumatoid arthritis. *Journal of Applied Physiology* 1993;**75**:1691-5.

De Jong 2003 *{published data only}*

de Jong Z, Munneke M, Zwinderman AH, Kroon HM, Jansen A, Ronday KH, et al. Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis? Results of a randomized controlled trial. *Arthritis and Rheumatism* 2003;**48**(9):2415-24.

Hansen 1993 *{published data only}*

Hansen TM, Hansen G, Langgaard AM, Rasmussen JO. Longterm physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. *Scandinavian Journal of Rheumatology* 1993;**22**(3):107-12.

Harkcom 1985 *{published data only}*

Harkcom TM, Lampman RM, Banwell BF, Castor CW. Therapeutic value of graded aerobic exercise training in rheumatoid arthritis. *Arthritis and Rheumatism* 1985;**28**(1):32-9.

Lyngberg 1994 *{published data only}*

Lyngberg KK, Harreby M, Bentzen H, Frost B. Elderly rheumatoid arthritis on steroid treatment tolerate physical training without an increase in disease activity. *Archives of Physical Medicine and Rehabilitation* 1994;**75**:1189-95.

Minor 1989 *{published data only}*

Minor MA, Hewett JE, Webel RR, Anderson SK, Kay DR. Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. *Arthritis and Rheumatism* 1989;**32**(11):1396-405.

Sanford-smith 1998 *{published data only}*

Sanford-Smith S, Mackay-Lyons M, Nunes-Clement S. Therapeutic benefit of aquaerobics for individuals with rheumatoid arthritis. *Physiotherapy Canada* 1998;**Winter**:40-6.

Van den Ende 1996 *{published data only}*

Van den Ende CHM, Hazes JMW, Le Cessie S, Mulder WJ, Belfor DG, Breedveld FC, Dijkmans BA. Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomized clinical trial. *Annals of the Rheumatic Diseases* 1996;**55**(11):798-805.

References to studies excluded from this review

Bearne 2002 *{published data only}*

Bearne LM, Scott DL, Hurley MV. Exercise can reverse quadriceps sensorimotor dysfunction that is associated with rheumatoid arthritis without exacerbating disease activity. *Rheumatology* 2002;**41**:157-66.

Bell 1998 *{published data only}*

Bell MJ, Lineker SC, Wilkins AL, Goldsmith CH, Badley EM. A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis. *Journal of Rheumatology* 1998;**25**(2):231-7.

Bilberg 2005 *{published data only}*

Bilberg A, Ahlmen M, Mannerkorpi K. Moderately intensive exercise in a temperate pool for patients with rheumatoid arthritis: a randomized controlled study. *Rheumatology* 2005;**44**:502-8.

Bostrom 1998 *{published data only}*

Bostrom C, Harms-Ringdahl K, Karreskog H, Nordemar R. Effects of static and dynamic shoulder rotator exercises in women with rheumatoid arthritis: a randomised comparison of impairment, disability, handicap, and health. *Scandinavian Journal of Rheumatology* 1998;**27**:281-90.

Buljina 2001 *{published data only}*

Buljina AI, Taljanovic MS, Avdic DM, Hunter TB. Physical and exercise therapy for treatment of the rheumatoid hand. *Arthritis Care and Research* 2001;**45**:392-7.

Callahan 2004 *{published data only}*

Callahan LF, Mielenz T, Freburger J. A randomized controlled trial of an eight week intervention: People with arthritis can exercise (PACE). *Arthritis and Rheumatism* 2004;**50 Suppl**(9):451.

Daltroy 1995 *{published data only}*

Daltroy LH, Robb-Nicholson C, Iversen MD, Wright EA, Liang MH. Effectiveness of minimally supervised home aerobic training in patients with systemic rheumatic disease. *British Journal of Rheumatology* 1995;**34**:1064-9.

de Jong 2005 *{published data only}*

de Jong Z, Vliet Vlieland TPM. Safety of exercise in patients with rheumatoid arthritis. *Current Opinion in Rheumatology* 2005;**17**:177-82.

Ekblom 1975 *{published data only}*

Ekblom B, Lovgren O, Alderin M, Fridstrom M, Satterstrom G. Effect of short-term physical training on patients with rheumatoid arthritis. A six-month follow-up study. *Scandinavian Journal of Rheumatology* 1975;**4**:87-91.

Ekdahl 1990 *{published data only}*

Ekdahl C, Andersson SI, Moritz U, Svensson B. Dynamic versus static training in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology* 1990;**19**:17-26.

Ekdahl 1994 *{published data only}*

Ekdahl C, Ekman R, Petersson I, Svensson B. Dynamic training and circulating neuropeptides in patients with rheumatoid arthritis: a comparative study with healthy subjects. *International Journal of Clinical Pharmacology Research* 1994;**14**:65-74.

Hafstrom 2005 *{published data only}*

Hafstrom I, Hallengren M. Physiotherapy in subtropic climate improves functional capacity and health-related quality of life in Swedish patients with rheumatoid arthritis and spondylarthropathies still after 6 months. *Scandinavian Journal of Rheumatology* 2003;**32**:108-13.

Hakkinen 1994 *{published data only}*

Hakkinen A, Hakkinen K, Hannonen P. Effects of strength training on neuromuscular function and disease activity in patients with recent-onset inflammatory arthritis. *Scandinavian Journal of Rheumatology* 1994;**23**:237-42.

Hakkinen 1997 *{published data only}*

Hakkinen A, Malkia E, Hakkinen K, Jappinen I, Laitinen L, Hannonen P. Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. *British Journal of Rheumatology* 1997;**36**:1075-81.

Hakkinen 1999 *{published data only}*

Hakkinen A, Sokka T, Kotaniemi A, Kautiainen H, Jappinen I, Laitinen, L, Hannonen P. Dynamic strength training in patients with early rheumatoid arthritis increases muscle strength but not bone mineral density. *Journal of Rheumatology* 1999;**26**(6):1257-63.

Hakkinen 2001 *{published data only}*

Hakkinen A, Sokka T, Kotaniemi A, Hannonen P. A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. *Arthritis and Rheumatism* 2001;**44**(3):515-22.

Hakkinen 2003a *{published data only}*

Hakkinen A, Sokka T, Lietsalmi AM, Kautiainen H, Hannonen P. Effects of dynamic strength training on physical function, Valpar 9 work sample test, and working capacity in patients with recent-onset rheumatoid arthritis. *Arthritis and Rheumatism* 2003;**49**(1):71-7.

Hakkinen 2003b *{published data only}*

Häkkinen A, Hannonen P, Nyman K, Lyyski T, Häkkinen K. Effects of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. *Arthritis and Rheumatism* 2003;**49**(6):789-97.

Hakkinen 2004a *{published data only}*

Hakkinen A, Sokka T, Hannonen P. A home-based two-year strength training period in early rheumatoid arthritis led to good long-term compliance: a five-year follow up. *Arthritis and Rheumatism* 2004;**51**(1):56-62.

Hakkinen 2004b *{published data only}*

Hakkinen A, Sokka T, Kautiainen H, Kotaniemi A, Hannonen P. Sustained maintenance of exercise induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5 year follow up. *Annals of the Rheumatic Diseases* 2004;**63**:910-6.

Hall 1996 *{published data only}*

Hall J, Skevingon SM, Maddison PJ, Chapman K. A randomized and controlled trial of hydrotherapy in rheumatoid arthritis. *Arthritis Care and Research* 1996;**206**:15.

Hall 2004 *{published data only}*

Hall J, Grant J, Blake D, Taylor G, Garbutt G. Cardiorespiratory responses to aquatic treadmill walking in patients with rheumatoid arthritis. *Physiotherapy Research International* 2004;**9**(2):59-73.

Hart 1994 *{published data only}*

Hart LE, Goldsmith CH, Churchill EM, Tugwell P. A randomized controlled trial to assess hydrotherapy in management of patients with rheumatoid arthritis. *Arthritis and Rheumatism* 1994;**37 Suppl**:146.

Kauppi 1998 *{published data only}*

Kauppi M, Leppanen L, Heikkila S, Lahtinen T, Kautiainen H. Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. *British Journal of Rheumatology* 1998;**37**:417-20.

Kim 1994 *{published data only}*

Kim JI, Lee EO. An effect of aquatic exercise program with self-help group activities and strategies for promoting self-efficacy on pain, physiological parameters and quality of life in patients having rheumatoid arthritis. *Arthritis and Rheumatism* 1994;**37 Suppl**:159.

Kirsteins 1991 *{published data only}*

Kirsteins AE, Dietz F, Hwang SM. Evaluating the safety and potential use of a weight-bearing exercise, Tai-Chi Chuan, for rheumatoid arthritis patients. *American Journal of Physical Medicine & Rehabilitation* 1991;**70**:136-41.

Komatireddy 1997 *{published data only}*

Komatireddy GR, Leitch RW, Cella K, Browning G, Minor M. Efficacy of low load resistive muscle training in patients with rheumatoid arthritis functional class II and III. *Journal of Rheumatology* 1997;**24**:1531-9.

Landewe 1992 *{published data only}*

Landewe RB, Peeters R, Verreussel RL, Masek BA, Goei The HS. [No difference in effectiveness measured between treatment in a thermal bath and in an exercise bath in patients with rheumatoid arthritis]. *Nederlands Tijdschrift voor Geneeskunde* 1992;**136**:173-6.

Lyngberg 1988 *{published data only}*

Lyngberg KK, Danneskiold-Samsøe B, Halskov O. The effect of physical training on patients with rheumatoid arthritis: changes in disease activity, muscle strength and aerobic capacity. A clinically controlled minimized cross-over study. *Clinical and Experimental Rheumatology* 1988;**6**:253-60.

Lyngberg 1994b *{published data only}*

Lyngberg KK, Ramsing BU, Nawrocki A, Harreby M, Danneskiold-Samsøe B. Safe and effective isokinetic knee extension training in rheumatoid arthritis. *Arthritis and Rheumatism* 1994;**37**:623-8.

Marcora 2005 *{published data only}*

Marcora SM, Lemmey AB, Maddison PJ. Can progressive resistance training reverse cachexia in patients with rheumatoid arthritis? Results of a pilot study. *Journal of Rheumatology* 2005;**32**:1031-9.

McClure 1986 *{published data only}*

McClure J, Sahrmann S, Norton B. The effect of a home exercise program on the gait and function of patients with arthritis. *Arthritis and Rheumatism* 1986;**29 Suppl**:147.

McMeeken 1999 *{published data only}*

McMeeken J, Stillman B, Story I, Kent P, Smith J. The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. *Physiotherapy Research International* 1999;**4**(1):55-67.

Minor 1995 *{published data only}*

Minor MA, Hewett JE. Physical fitness and work capacity in women with rheumatoid arthritis. *Arthritis Care and Research* 1995;**8**:146-54.

Neuberger 1997 *{published data only}*

Neuberger GB, Press AN, Lindsley HB, Hinton R, Cagle PE, Carlson K, et al. Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis. *Research in Nursing and Health* 1997;**20**:195-204.

Neuberger 2000 *{published data only}*

Neuberger G, Aaronson L, Cagle P, Miller P, David S, Jamison, et al. The effects of exercise on measures of inflammation, disease activity and aerobic fitness among rheumatoid arthritis participants in either class exercise, home exercise or in a control group. *Arthritis and Rheumatism* 2000;**43 Suppl**(9):394.

Nordemar 1981 *{published data only}*

Nordemar R, Ekblom B, Zachrisson L, Lundqvist K. Physical training in rheumatoid arthritis: a controlled long-term study I. *Scandinavian Journal of Rheumatology* 1981;**10**:17-23.

Nordstrom 1996 *{published data only}*

Nordstrom DCE, Konttinen YT, Solovieva S, Friman C, Santavirta S. In- and out-patient rehabilitation in rheumatoid arthritis. *Scandinavian Journal of Rheumatology* 1996;**25**:200-6.

Noreau 1995 *{published data only}*

Noreau L, Martineau H, Roy L, Belzile M. Effects of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. *American Journal of Physical Medicine and Rehabilitation* 1995;**74**:19-27.

O'Brien 2006 *{published data only}*

O'Brien AV, Jones P, Mullis R, Mulherin D, Dziedzic K. Conservative hand therapy treatments in rheumatoid arthritis - a randomized controlled trial. *Rheumatology* 2006;**45**:577-83.

Perlman 1987 *{published data only}*

Perlman SG, Connell K, Alberti J, Conlon P, Clark A, Caldron P. Synergistic effects of exercise and problem solving education for RA patients. *Arthritis and Rheumatism* 1987;**30 Suppl**:13.

Stenstrom 1991 *{published data only}*

Stenstrom CH, Lindell B, Swanberg E, Swanberg P, Harms Ringdahl K, Nordemar R. Intensive dynamic training in water for rheumatoid arthritis functional class II—a long-term study of effects. *Scandinavian Journal of Rheumatology* 1991;**20**:358-65.

Stenstrom 1994 *{published data only}*

Stenstrom CH. Home exercise in rheumatoid arthritis functional class II: goal setting versus pain attention. *Journal of Rheumatology* 1994;**21**:627-34.

Stenstrom 1996 *{published data only}*

Stenstrom CH, Arge B, Sundbom A. Dynamic training versus relaxation training as home exercise for patients with inflammatory rheumatic diseases. A randomized controlled study. *Scandinavian Journal of Rheumatology* 1996;**25**:28-33.

Suomi 1996 *{published data only}*

Suomi R, Koceja DM. Effects of the Arthritis Foundation Aquatic Program (AFAP) on postural stability in women with arthritis. *Medicine and Science in Sports and Exercise* 1996;**28 Suppl**:8.

Suomi 1997 *{published data only}*

Suomi R, Lindauer S. Effectiveness of arthritis foundation aquatic program on strength and range of motion in women with arthritis. *Journal of Aging and Physical Activity* 1997;**5**:341-51.

Suomi 2003 *{published data only}*

Suomi R, Collier D. Effects of arthritis exercise programs on functional fitness and perceived activities of daily living measures in older adults with arthritis. *Archives of Physical Medicine and Rehabilitation* 2003;**84**:1589-94.

van den Ende 2000 *{published data only}*

van den Ende CH, Breedveld FC, le Cessie S, Dijkmans BA, de Mug AW, Hazes JM. Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. *Annals of the Rheumatic Diseases* 2000;**59**:615-21.

Van Deusen 1987 *{published data only}*

Van Deusen J, Harlowe D. A comparison of the ROM dance home exercise/rest program with traditional routines. *American Journal of Occupational Therapy* 1987;**7**:349-61.

Westby 2000 *{published data only}*

Westby MD, Wade JP, Rangno KK, Berkowitz J. A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. *Journal of Rheumatology* 2000;**27**(7):1674-80.

Wineland 1985 *{published data only}*

Wineland M, Zischke J, Dennis G, Klipple G. The physical and psychological effects of a recreational exercise program on persons with rheumatoid arthritis andosteoarthritis. *Arthritis and Rheumatism* 1985;**28 Suppl**:139.

Additional references

ACR 2002

American College of Rheumatology Subcommittee on rheumatoid arthritis guidelines. Guidelines for the management of rheumatoid arthritis: 2002 update. *Arthritis and Rheumatism* 2002;**46**(2):328-46.

Arnett 1987

Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism* 1988;**31**:315-23.

Baker 1953

Baker F. The rationale for physical therapy in arthritis. *Bulletin on the Rheumatic Diseases* 1953;**4**:57-8.

Chinn 2000

Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine* 2000;**19**:3127-31.

Christie 2007

Christie A, Jamtvedt G, Dahm KT, Moe RH, Haavardsholm EA, Hagen KB. Effectiveness of nonpharmacological and nonsurgical interventions for patients with rheumatoid arthritis: an overview of systematic reviews. *Physical Therapy* 2007;**87**(12):1697-715.

Cohen 1988

Cohen J. Statistical power analysis for the behavioral sciences. 2nd Edition. Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.

Combe 2007

Combe B, Landewé R, Lukas C, Bolosiu HD, Breedveld FC, Dougados M, et al. EULAR recommendations for the management of early arthritis: Report of a task force of the European Standing Committee for International Clinical Studies including Therapeutics (ESCISIT). *Annals of the Rheumatic Diseases* 2007;**66**(1):34-45.

Dagfinrud 2007

Dagfinrud H, Christie A. Patients with rheumatoid arthritis feel better after exercises in warm water than after similar exercises on land. *The Australian Journal of Physiotherapy* 2007;**53**(2):130.

De Jong 2004a

de Jong Z, Munneke M, Lems WF, Zwinderman AH, Kroon HM, Pauwels EK, et al. Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. *Arthritis and Rheumatism* 2004;**50**(4):1066-76.

De Jong 2004b

De Jong Z, Munneke M, Zwinderman AH, Kroon HM, Ronday KH, Lems WF, et al. Long term high intensity exercise and damage of small joints in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2004;**63**(11):1399-405.

De Jong 2009

de Jong Z, Munneke M, Kroon HM, van Schaardenburg D, Dijkmans BA, Hazes JM, Vliet Vlieland TP. Long-term follow-up of a high-intensity exercise program in patients with rheumatoid arthritis. *Clinical Rheumatology* 2009;**[Epub ahead of print]**.

Dickerson 1994

Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;**309**:1286-91.

Fleiss 1973

FleissJL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement* 1973;**33**:613-9.

Fries 1980

Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis and Rheumatism* 1980;**23**:137-45.

Gaudin 2007

Gaudin P, Leguen-Guegan S, Allenet B, Baillet A, Grange L, Juvin R. Is dynamic exercise beneficial in patients with rheumatoid arthritis?. *Joint, Bone, Spine: Revue du Rhumatisme* 2007;**75**(1):11-7.

Hakkinen 2004

Hakkinen A. Effectiveness and safety of strength training in rheumatoid arthritis. *Current Opinion in Rheumatology* 2004;**16**(2):132-7.

Higgins 2008

Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Hurst 1997

Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *British Journal of Rheumatology* 1997;**36**(5):551-9.

Jette 1980

Jette AM. Functional status index: reliability of a chronic disease evaluation instrument. *Archives of Physical Medicine and Rehabilitation* 1980;**61**:395-401.

Jivoff 1975

Jivoff L. Rehabilitation and rheumatoid arthritis. *Bulletin on the Rheumatic Diseases* 1975;**26**:838-41.

Macera 2003

Macera CA, Hootman JM, Sniezek JE. Major public health benefits of physical activity. *Arthritis and Rheumatism* 2003;**49**(1):122-8.

Marra 2005

Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, et al. A comparison of generic, indirect utility measures (the

HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Social Science and Medicine* 2005;**60**(7):1571-82.

Meenan 1980

Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. The arthritis impact measurement scales. *Arthritis and Rheumatism* 1980;**23**:146-52.

Munneke 2005

Munneke M, de Jong Z, Zwinderman AH, Ronday HK, van Schaardenburg D, Dijkmans BA, et al. Effect of a high-intensity weight-bearing exercise program on radiologic damage progression of the large joints in subgroups of patients with rheumatoid arthritis. *Arthritis and Rheumatism* 2005;**53**(3):410-7.

Ottawa 2004

Ottawa Panel. Ottawa Panel evidence-based clinical practice guidelines for therapeutic exercises in the management of rheumatoid arthritis management of rheumatoid arthritis in adults. *Physical Therapy* 2004;**84**(10):934-72.

Pate 1995

Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA* 1995;**273**(5):402-7.

Pollock 1998

Pollock ML, Gaesser GA, Butcher JD, Després JP, Dishman RK, Franklin BA, et al. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. Medicine and Science in Sports and Exercise 1998; Vol. 30, issue 6:975-99.

Robinson 2002

Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of controlled trials using PubMed. *International Journal of Epidemiology* 2002;**31**:150-3.

Ropes 1958

Ropes WR, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 Revision of diagnostic criteria for rheumatoid arthritis. *Bulletin on the Rheumatic Diseases* 1958;**9**:175-6.

Smidt 2005

Smidt N, de Vet HC, Bouter LM, Dekker J, Arendzen JH, de Bie RA, et al. Effectiveness of exercise therapy: a best-evidence summary of systematic reviews. *The Australian Journal of Physiotherapy* 2005;**51**(2):71-85.

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Spitz 1987

Spitz PW, Fries JF. The present and future of comprehensive outcome measures for rheumatic diseases. *Clinical Rheumatology* 1987;**6 Suppl 2**:105-11.

Stenström 2003

Stenström CH, Minor MA. Evidence for the benefit of aerobic and strengthening exercise in rheumatoid arthritis. *Arthritis and Rheumatism* 2003;**49**(3):428-34.

Tijhuis 2000

Tijhuis GJ, Jansen SJ, Stiggelbout AM, Zwinderman AH, Hazes JM, Vlieland TPM. Value of the time trade off method for measuring utilities in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2000;**59**(11):892-7.

van den Ende 1998

Van den Ende CHM, Vliet Vlieland TPM, Munneke M, Hazes JMW. Dynamic exercise therapy for treating rheumatoid arthritis (review). *Cochrane Database of Systematic Reviews* 1998, Issue 4. [DOI: 10.1002/14651858.CD000322.pub2]

van den Ende 2006

Van den Ende CHM, Steultjens EMJ, Bouter LM, Dekker J. Clinical heterogeneity was a common problem in cochrane reviews of physiotherapy and occupational therapy. *Journal of Clinical Epidemiology* 2006;**59**(9):914-9.

van den Hout 2005

Van den Hout WB, de Jong Z, Munneke M, Hazes JM, Breedveld FC, Vliet Vlieland TP. Cost-utility and cost-effectiveness analyses of a long-term, high-intensity exercise program compared with conventional physical therapy in patients with rheumatoid arthritis. *Arthritis and Rheumatism* 2005;**53**(1):39-47.

van Tulder 2003

Van Tulder MW, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 2003;**28**:1290-1299.

Verhagen 1998

Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of Clinical Epidemiology* 1998;**51**(12):1235-41.

Verhoeven 2000

Verhoeven AC, Broers M, van der Liden S. Validity of the MACTAR questionnaire as a functionel index in a rheumatoid arthritis clinical trial. The McMaster Toronto Arthritis. *Journal of Rheumatology* 2000;**27**(12):2801-9.

Baslund 1993

Methods	Prospective, randomised trial, 2 groups
Participants	2 males, 16 females; mean (SD) age: 48(9) yrs, mean (SD): 14 (11) yrs Inclusion criteria: RA according to ARA criteria, age < 65 yr; > 2 months before baseline Exclusion criteria: inability to perform bicycle training program Setting: hospital registry
Interventions	One dynamic group: 4 to 5x weekly bicycle training; one control group: training not allowed Duration: 8 weeks Supervisor: unknown Training supervisor: unknown
Outcomes	Aerobic capacity: VO_2max Disease activity: ESR, CRP
Notes	Study fulfilled the methodological criteria: 2, 3, 4, 7, 8, 9 and 10

Risk of bias

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Unclear risk	Not stated whether the assessor was blinded or not
Incomplete outcome data addressed? All outcomes	High risk	Not specifically stated in results or discussion section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

De Jong 2003

Methods	Prospective, randomised clinical trial, 2 groups
Participants	237 females, 63 males; mean (SD) age: experimental group 54 (18) yrs, control group 54 (16) yrs, mean (SD) DD: experimental group 8 (11) yrs, experimental group 5 (7) yrs Inclusion criteria: RA according to the 1987 ARA criteria, age 20 to 70 yrs, stable medication for 3 months, ability to cycle on a home trainer, functional class I, II and III, willingness to exercise biweekly on a fixed schedule, living in a predefined region

De Jong 2003 *(Continued)*

	Exclusion criteria: serious cardiac or lung disease preventing cardio-respiratory fitness training and prosthesis of weight-bearing joints Setting: hospital registry
Interventions	One dynamic group: high-intensity exercises 2x weekly Control group: physical therapy when necessary Duration: 2 years and a 18 months follow-up Supervisor: unknown Training supervisor: unknown
Outcomes	Aerobic capacity: ergo meter test (watts) Muscle strength: isokinetic dynamometer (newtons) Functional ability: MACTAR and HAQ Disease activity: DAS Radiological joint damage: Larsen score Cost-effectiveness: EQ-5D
Notes	Study fulfilled the methodological criteria: 1, 3, 4, 5, 6, 7, 8, 9 and 10

Risk of bias

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Low risk	A - Adequate
Blinding? All outcomes - assessor	Low risk	Stated in method section
Incomplete outcome data addressed? All outcomes	Low risk	Addressed in results section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Low risk	

Hansen 1993

Methods	Prospective, randomised clinical trial, 5 groups
Participants	49 females / 26 males, mean age: 53 yrs, mean DD: 7 yrs Inclusion criteria: age 20-6- yr, RA according 1958 ARA criteria. Exclusion criteria: Steinbrocker III and IV, co-morbidity, presence of contra-indications for training, already training 3x per week Setting: hospital registry and physical therapy practices
Interventions	4 dynamic groups, varying in amount of training and condition (water, bicycle) one control group (no exercise)

Hansen 1993 *(Continued)*

	Duration of intervention: 2 yrs Supervisor: physical therapists Training supervisor: unknown
Outcomes	Aerobic capacity: aerobic fitness Muscle strength: isokinetic strength knee Functional ability: HAQ, functional score, medicine costs Disease activity: ESR, Hb, swollen joint count, pain (VAS), morning stiffness
Notes	Study fulfilled the methodological criteria: 2, 3, 4, 5, 6, 7 and 8

Risk of bias

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Low risk	Stated in method section
Incomplete outcome data addressed? All outcomes	Low risk	Addressed in results section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

Harkcom 1985

Methods	Prospective, randomised clinical trial, 4 groups
Participants	20 females, ARA 1958 criteria, Steinbrocker II Mean (SD) age: 52 (12) yrs, mean (SD) DD: 9 (7) yrs Inclusion criteria: not reported Exclusion criteria: not reported Setting: outpatient clinic
Interventions	3 dynamic groups with bicycle exercise 3x weekly varying in degree, control group (no exercise) Duration: 12 weeks Supervisor: physical education graduate students Training supervisor: unknown
Outcomes	Aerobic capacity: VO_2max, heart rate, exercise test time

Harkcom 1985 *(Continued)*

	Muscle strength: isokinetic strength knee, grip strength Functional ability: FSI Disease activity: N of tender joints
Notes	Study fulfilled the methodological criteria: 2, 4, 7 and 9

Risk of bias

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Unclear risk	Not stated whether the assessor was blinded or not
Incomplete outcome data addressed? All outcomes	High risk	Not specifically stated in results or discussion section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

Lyngberg 1994

Methods	Prospective, randomised, single-blinded, 2 groups
Participants	22 males, 2 females; mean (SD) age: 67(9) yrs, mean (SD) DD: 9 (11) yrs, Steinbrocker II Inclusion criteria: ARA criteria, use of low dose of glucocorticosteroids, stable since 3 months Exclusion criteria: heart disease, inability to exercise Setting: outpatient clinic
Interventions	One dynamic group: bicycling and strengthening exercises (heel lifting, step climbing), 2x weekly; one control group: no exercise Duration: 3 months Supervisor: physical therapists Training supervisor: unknown
Outcomes	Aerobic capacity: VO_2 max Muscle strength: isokinetic strength (knee and ankle) Functional ability: Fries index Disease activity and pain: no swollen joints, no of tender joints, ESR, Hb
Notes	Study fulfilled the methodological criteria: 2, 3, 4, 5, 6, 7, 8 and 9

Risk of bias

Lyngberg 1994 *(Continued)*

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Low risk	Stated in method section
Incomplete outcome data addressed? All outcomes	High risk	Not specifically stated in results or discussion section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

Minor 1989

Methods	Prospective, randomised clinical trial, 3 groups
Participants	34 females, 6 males; mean (SD) age 54(14) yrs, mean (SD) DD: 11(8) yrs Inclusion: RA 1958 criteria, symptomatic weight-bearing joints, age: >20yr, DD: >6 months. Exclusion criteria: currently exercising, medical condition precluding increased activity Setting: outpatient rheumatology clinics
Interventions	2 dynamic groups: aerobic pool group and aerobic walk group, 3x weekly. One control group: ROM exercises 3x weekly, Duration: 12 weeks and a 3 months and 9 months follow-up Supervisor: three instructors with unknown profession Training supervisor: unknown
Outcomes	Aerobic capacity: VO_2 max, exercise endurance, resting blood pressure, exercise heart rate Functional ability: AIMS Disease activity: N clinical active joints, morning stiffness
Notes	Study fulfilled the methodological criteria: 3, 4, 6, 7 and 9

Risk of bias

Bias	Authors' judgement	Support for judgement
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Unclear risk	Not stated whether the assessor was blinded or not

Minor 1989 *(Continued)*

Incomplete outcome data addressed? All outcomes	Low risk	Addressed in results section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

Sanford-smith 1998

Methods	Prospective, randomised clinical trial, 2 groups
Participants	19 females, 5 males; mean (SD) age: experimental group 62 (12) yrs, control group 55 (15) yrs, mean (SD) DD: experimental group 20 (13) yrs, control group 12 (8) yrs Inclusion criteria: RA according 1958 ARA criteria, Steinbrocker functional class II and III and a stable drug regime for 3 months Exclusion criteria: unstable heart disease or already involved in an exercise program Setting: outpatient clinic
Interventions	One dynamic group: aquaerobics 3x weekly Control group: ROM exercises 2 to 3x weekly Duration: 10 weeks Supervisor: physical therapists Training supervisor: unknown
Outcomes	Functional ability: HAQ Muscle strength: grip strength Disease activity: AJC, ESR
Notes	Study fulfilled the methodological criteria: 2, 3, 4, 5, 6, 7, 8 and 9

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Low risk	Stated in method section
Incomplete outcome data addressed? All outcomes	Low risk	Addressed in results section

Sanford-smith 1998 *(Continued)*

Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

Van den Ende 1996

Methods	Prospective, randomised clinical trial, 4 groups
Participants	37 females, 37 males; mean (SD) age: 52(12) yrs, mean (SD) DD 10 (8) yrs Inclusion criteria: age 20 to 70 yr, stable on medication, able to bicycle. Exclusion criteria: arthroplasties of weight-bearing joints, co-morbidity Setting: outpatient clinic
Interventions	One dynamic group: bicycle and weight-bearing exercises 3x weekly; 3 control groups: ROM + isometric exercises; 2 supervised, one group written instructions Duration 12 weeks and a 12 week follow-up Supervisor: physical therapists Training supervisor: unknown
Outcomes	Aerobic capacity: VO_2 max Muscle strength: isometric and isokinetic muscle strength Functional ability: HAQ Disease activity and pain: ESR, no of swollen joints, no of tender joints
Notes	Study fulfilled the methodological criteria: 2, 3, 4, 6, 7, 8, 9 and 10

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Adequate sequence generation?	Low risk	Addressed in method section
Allocation concealment?	Unclear risk	Unclear how allocation was concealed
Blinding? All outcomes - assessor	Unclear risk	Not stated whether the assessor was blinded or not
Incomplete outcome data addressed? All outcomes	Low risk	Addressed in results section
Free of selective reporting?	Unclear risk	Unclear whether all measured data during the trial was also reported in the article
Free of other bias?	Unclear risk	Unclear what the influences are of changes in medication, intra-articular injections, and surgery during the trial on the outcomes

DD: Disease Duration

EPM-ROM: EPM-Range of Motion scale
HAQ: Health Assessment Questionnaire
FSI: Functional Status Index
AJC: Active Joint Count
ESR: Erythrocyte Sedimentation Rate
MACTAR: McMaster Toronto Arthritis patient preference questionnaire
EQ-5D: EuroQol-5 dimensional
QALY: Quality-adjusted life years

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Bearne 2002	Duration of intervention less than 6 weeks
Bell 1998	Intervention was not dynamic exercise program
Bilberg 2005	Level of exercise intensity unclear
Bostrom 1998	Not under supervision
Buljina 2001	Intervention was not dynamic exercise program
Callahan 2004	Full-length written report not available
Daltroy 1995	Pooled data of RA and SLE patients
de Jong 2005	Not a randomised trial
Ekblom 1975	Not a randomised trial
Ekdahl 1990	Level of exercise intensity unclear
Ekdahl 1994	Level of exercise intensity unclear
Hafstrom 2005	Not a randomised trial
Hakkinen 1994	Exercise duration less than 20 minutes per session
Hakkinen 1997	Exercise frequency unclear
Hakkinen 1999	Not under supervision
Hakkinen 2001	Not under supervision
Hakkinen 2003a	Not under supervision
Hakkinen 2003b	Exercise duration less than 20 minutes per session
Hakkinen 2004a	Not under supervision
Hakkinen 2004b	Not under supervision
Hall 1996	Duration of intervention less than 6 weeks
Hall 2004	Not a randomised trial

Study	Reason for exclusion
Hart 1994	Full-length written report not available
Kauppi 1998	Duration of intervention less than 6 weeks
Kim 1994	Full length written report in English not available
Kirsteins 1991	Not a randomised trial
Komatireddy 1997	Not under supervision
Landewe 1992	Not a randomised trial
Lyngberg 1988	Not a randomised trial
Lyngberg 1994b	Not a randomised trial
Marcora 2005	Not a randomised trial
McClure 1986	Full-length written report not available
McMeeken 1999	Level of intensity unclear
Minor 1995	Not a randomised trial
Neuberger 1997	Not a randomised trial
Neuberger 2000	Full-length written report not available
Nordemar 1981	Not a randomised trial
Nordstrom 1996	Not a randomised trial
Noreau 1995	Not a randomised trial
O'Brien 2006	Intervention is not dynamic exercise program
Perlman 1987	Not a randomised trial
Stenstrom 1991	Not a randomised trial
Stenstrom 1994	Exercise duration less than 20 minutes per session
Stenstrom 1996	Exercise duration less than 20 minutes per session
Suomi 1996	Full-length written report not available
Suomi 1997	Pooled data of RA and OA participants
Suomi 2003	Pooled data of RA and OA participants
van den Ende 2000	Duration of intervention unclear
Van Deusen 1987	Not a randomised trial
Westby 2000	Not under supervision

Study	Reason for exclusion
Wineland 1985	Full-length written report not available

DATA AND ANALYSES

Comparison 1. Short-term land-based aerobic capacity training

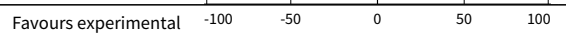
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Functional ability	2	66	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.46, 0.51]
2 Aerobic capacity	3	82	Std. Mean Difference (IV, Random, 95% CI)	0.99 [0.29, 1.68]
3 Muscle strength	1	10	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-1.67, 0.90]
4 Self-reported pain	1	56	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-0.79, 0.26]

Analysis 1.1. Comparison 1 Short-term land-based aerobic capacity training, Outcome 1 Functional ability.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Harkcom 1985	4	-0.8 (10.7)	6	0 (12)		14.63%	-0.06[-1.33,1.2]
Minor 1989	28	1 (2.5)	28	0.9 (2.4)		85.37%	0.04[-0.48,0.56]
Total ***	**32**		**34**			**100%**	**0.03[-0.46,0.51]**
Heterogeneity: Tau^2=0; Chi^2=0.02, df=1(P=0.88); I^2=0%							
Test for overall effect: Z=0.1(P=0.92)							

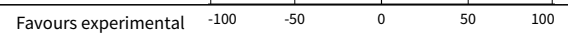
Analysis 1.2. Comparison 1 Short-term land-based aerobic capacity training, Outcome 2 Aerobic capacity.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Baslund 1993	4	7.2 (13.2)	6	0.1 (7.9)		21.1%	0.63[-0.69,1.94]
Harkcom 1985	28	3.5 (4.8)	28	-0.1 (4.8)		56.08%	0.74[0.2,1.28]
Minor 1989	8	6.1 (1.8)	8	1.3 (2.8)		22.82%	1.93[0.68,3.17]
Total ***	**40**		**42**			**100%**	**0.99[0.29,1.68]**
Heterogeneity: Tau^2=0.15; Chi^2=3.11, df=2(P=0.21); I^2=35.79%							
Test for overall effect: Z=2.78(P=0.01)							



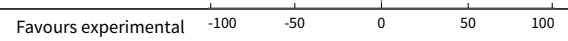
Analysis 1.3. Comparison 1 Short-term land-based aerobic capacity training, Outcome 3 Muscle strength.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Harkcom 1985	4	2 (19.3)	6	9 (14.7)		100%	-0.38[-1.67,0.9]
Total ***	**4**		**6**			**100%**	**-0.38[-1.67,0.9]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.58(P=0.56)							



Analysis 1.4. Comparison 1 Short-term land-based aerobic capacity training, Outcome 4 Self-reported pain.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Minor 1989	28	-1.2 (1.9)	28	-0.7 (1.8)		100%	-0.27[-0.79,0.26]
Total ***	**28**		**28**			**100%**	**-0.27[-0.79,0.26]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.99(P=0.32)							

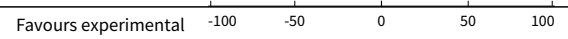


Comparison 2. Short-term land-based aerobic capacity and muscle strength training

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Functional ability	2	74	Std. Mean Difference (IV, Random, 95% CI)	-0.40 [-0.86, 0.06]
2 Muscle strength	2	74	Std. Mean Difference (IV, Random, 95% CI)	0.47 [0.01, 0.93]
3 Self-reported pain	1	50	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-1.09, 0.04]

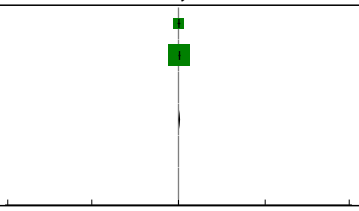
Analysis 2.1. Comparison 2 Short-term land-based aerobic capacity and muscle strength training, Outcome 1 Functional ability.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Lyngberg 1994	12	-2 (13.8)	12	0 (21.5)		33.26%	-0.11[-0.91,0.69]
Van den Ende 1996	25	-0 (0.4)	25	0.2 (0.4)		66.74%	-0.54[-1.11,0.02]
Total ***	**37**		**37**			**100%**	**-0.4[-0.86,0.06]**
Heterogeneity: Tau^2=0; Chi^2=0.76, df=1(P=0.38); I^2=0%							
Test for overall effect: Z=1.69(P=0.09)							



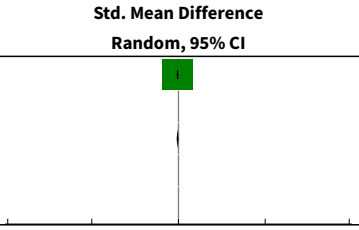
Analysis 2.2. Comparison 2 Short-term land-based aerobic capacity and muscle strength training, Outcome 2 Muscle strength.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Lyngberg 1994	12	-3 (72)	12	-21 (61.8)		33.18%	0.26[-0.55,1.06]
Van den Ende 1996	25	13.6 (21)	25	-0.7 (27.5)		66.82%	0.58[0.01,1.14]
Total ***	**37**		**37**			**100%**	**0.47[0.01,0.93]**
Heterogeneity: Tau^2=0; Chi^2=0.4, df=1(P=0.53); I^2=0%							
Test for overall effect: Z=1.99(P=0.05)							



Analysis 2.3. Comparison 2 Short-term land-based aerobic capacity and muscle strength training, Outcome 3 Self-reported pain.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Van den Ende 1996	25	0.2 (1.4)	25	0.9 (1.2)		100%	-0.53[-1.09,0.04]
Total ***	**25**		**25**			**100%**	**-0.53[-1.09,0.04]**
Heterogeneity: Not applicable							
Test for overall effect: Z=1.83(P=0.07)							

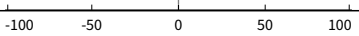


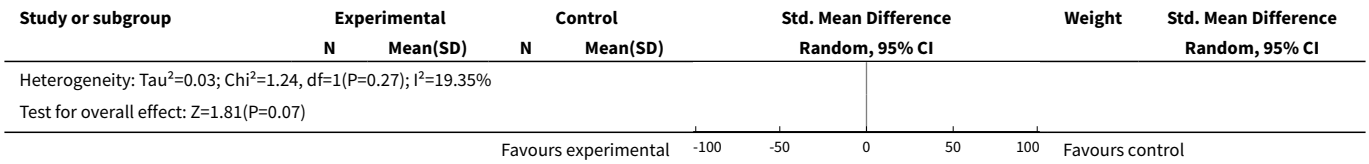
Comparison 3. Short-term water-based aerobic capacity training

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Aerobic capacity	2	88	Std. Mean Difference (IV, Random, 95% CI)	0.47 [-0.04, 0.98]
2 Muscle strenght	1	20	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-1.27, 0.51]
3 Self-reported pain	1	68	Std. Mean Difference (IV, Random, 95% CI)	0.06 [-0.43, 0.54]

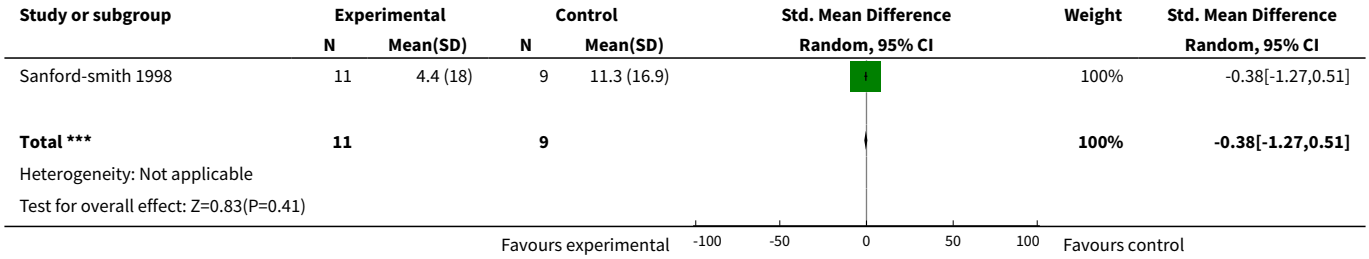
Analysis 3.1. Comparison 3 Short-term water-based aerobic capacity training, Outcome 1 Aerobic capacity.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Minor 1989	40	3.9 (7)	28	-0.1 (4.8)		70.95%	0.64[0.14,1.14]
Sanford-smith 1998	11	2.5 (6.4)	9	2 (8.3)		29.05%	0.07[-0.82,0.95]
Total ***	**51**		**37**			**100%**	**0.47[-0.04,0.98]**

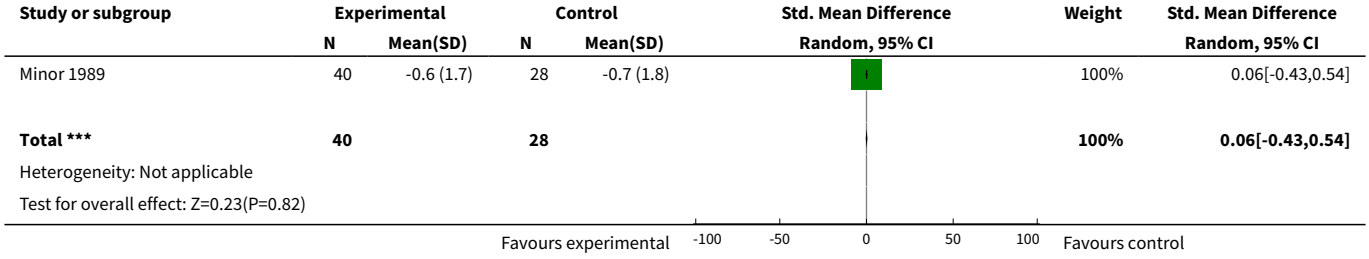




Analysis 3.2. Comparison 3 Short-term water-based aerobic capacity training, Outcome 2 Muscle strenght.



Analysis 3.3. Comparison 3 Short-term water-based aerobic capacity training, Outcome 3 Self-reported pain.

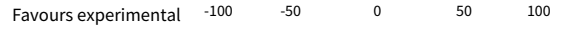


Comparison 4. Long-term land-based aerobic capacity and muscle strength training

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Aerobic capacity	1	281	Std. Mean Difference (IV, Random, 95% CI)	0.46 [0.22, 0.70]
2 Muscle strength	2	305	Std. Mean Difference (IV, Random, 95% CI)	0.49 [-0.06, 1.04]
3 Self-reported pain	1	24	Std. Mean Difference (IV, Random, 95% CI)	0.35 [-0.46, 1.16]
4 Disease activity	2	305	Std. Mean Difference (IV, Random, 95% CI)	-0.16 [-0.39, 0.06]
5 Radiological damage	2	305	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.37, 0.08]

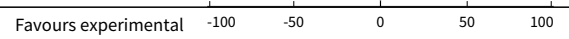
Analysis 4.1. Comparison 4 Long-term land-based aerobic capacity and muscle strength training, Outcome 1 Aerobic capacity.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
De Jong 2003	136	8.2 (37.1)	145	-6.7 (27.1)		100%	0.46[0.22,0.7]
Total ***	**136**		**145**			**100%**	**0.46[0.22,0.7]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.8(P=0)							



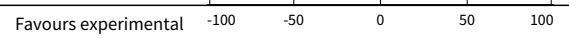
Analysis 4.2. Comparison 4 Long-term land-based aerobic capacity and muscle strength training, Outcome 2 Muscle strength.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
De Jong 2003	11	26 (10.8)	13	16 (9.8)		27.81%	0.95[0.09,1.8]
Hansen 1993	136	26.1 (50)	145	9.6 (54)		72.19%	0.32[0.08,0.55]
Total ***	**147**		**158**			**100%**	**0.49[-0.06,1.04]**
Heterogeneity: Tau^2=0.1; Chi^2=1.94, df=1(P=0.16); I^2=48.34%							
Test for overall effect: Z=1.74(P=0.08)							



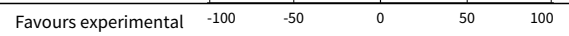
Analysis 4.3. Comparison 4 Long-term land-based aerobic capacity and muscle strength training, Outcome 3 Self-reported pain.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Hansen 1993	11	0.2 (0.6)	13	0 (0.5)		100%	0.35[-0.46,1.16]
Total ***	**11**		**13**			**100%**	**0.35[-0.46,1.16]**
Heterogeneity: Not applicable							
Test for overall effect: Z=0.85(P=0.39)							



Analysis 4.4. Comparison 4 Long-term land-based aerobic capacity and muscle strength training, Outcome 4 Disease activity.

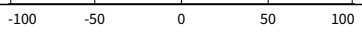
Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
De Jong 2003	11	-3 (22)	13	5 (15)		7.66%	-0.42[-1.23,0.4]
Hansen 1993	136	-0.9 (1.3)	145	-0.7 (1.5)		92.34%	-0.14[-0.38,0.09]
Total ***	**147**		**158**			**100%**	**-0.16[-0.39,0.06]**
Heterogeneity: Tau^2=0; Chi^2=0.41, df=1(P=0.52); I^2=0%							



Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Test for overall effect: Z=1.42(P=0.16)							

Analysis 4.5. Comparison 4 Long-term land-based aerobic capacity and muscle strength training, Outcome 5 Radiological damage.

Study or subgroup	Experimental N	Experimental Mean(SD)	Control N	Control Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
De Jong 2003	11	9 (29)	13	6 (15.8)		7.84%	0.13[-0.68,0.93]
Hansen 1993	136	1.5 (2.8)	145	2 (3)		92.16%	-0.17[-0.41,0.06]
Total ***	**147**		**158**			**100%**	**-0.15[-0.37,0.08]**
Heterogeneity: Tau^2=0; Chi^2=0.49, df=1(P=0.48); I^2=0%							
Test for overall effect: Z=1.29(P=0.2)							



ADDITIONAL TABLES

Table 1. Short-term land-based aerobic capacity training

Study	Functional ability	Aerobic capacity	Muscle strength	Self-reported pain	Disease activity	Radiological damage
Baslund 1993		SUBMAXIMAL BICYCLE TEST Dynamic group: Mean (SD) baseline: 27.2 (1.7) ml/kg/min After 8 weeks: 33.3 Control group: Mean (SD) baseline: 20.9 (2.9) ml/kg/min After 8 weeks: 22.2 SMD: 1.93 (0.68, 3.17) P between groups: <.05	-	-	ESR Dynamic group: Mean (SD) baseline: 21 (4) After 8 weeks: 22 Control group: Mean (SD) baseline: 28 (3) After 8 weeks: 25 SMD: 0.95 (-0.10, 2.0) P between groups: ns	-

Table 1. Short-term land-based aerobic capacity training *(Continued)*

Minor 1989	AIMS (physical activity) Dynamic group: Mean (SD) baseline: 4.6 (2.7) After 12 weeks: 3.6 Control group: Mean (SD) baseline: 4.0 (2.5) After 12 weeks: 4.9 SMD: -0.06 (-1.33, 1.20) P between groups: n.s. P between groups after 3 months: ? P between groups after 9 months: ns	MAXIMAL TREADMILL TEST Dynamic group: Mean (SD) baseline: 18.9 (4.8) After 12 weeks: 22.4 Control group: Mean (SD) baseline: 17.4 (5.9) After 12 weeks: 17.3 SMD: 0.74 (0.20, 1.28) P between groups: <.05 P between groups after 3 months:? P between groups after 9 months: ns		PAIN (AIMS) Dynamic group: Mean (SD) baseline: 5.1 (1.9) After 12 weeks: 3.9 Control group: Mean (SD) baseline: 5.5 (1.6) After 12 weeks: 4.8 SMD: -0.27 (-0.79, 0.26) P between groups: ns P between groups after 3 months:? P between groups after 9 months: ns	CLINICALLY ACTIVE JOINTS Dynamic group: Mean (SD) baseline: 12.4 (12.5) After 12 weeks: 7.9 Control group: Mean (SD) baseline: 12.4 (13.2) After 12 weeks: 12.0 SMD: -0.39 (-0.92, 0.14) P between groups: ns P between groups after 3 months:? P between groups after 9 months: ns	-
Harkcom 1985	FSI Dynamic group: Mean (SD) baseline: 16.8 (8.3) After 12 weeks: 16.0 Control group: Mean (SD) baseline: 14.8 (10.2) After 12 weeks: 14.8 SMD: 0.04 (-0.48, 0.56) P between groups: ns	MAXIMAL BICYCLE TEST Dynamic group: Mean (SD) baseline: 21.9 (9.0) ml/kg/min After 12 weeks: 29.1 Control group: Mean (SD) baseline: 18.5 (7.4) ml/kg/min After 12 weeks: 18.6 SMD: 0.63 (-0.69, 1.94) P between groups: ns	ISOMETRIC EXTENSION Dynamic group Mean (SD) baseline: 60.5 (15.6) foot pounds After 12 weeks: 62.5 Control group: Mean (SD) baseline: 63.6 (15.7) foot pounds After 12 weeks: 74.6 SMD: -0.38 (-1.67, 0.90)		N OF TENDER JOINTS (PAIN AND SWELLING) Dynamic group: Mean (SD) baseline: 32.5 (19.4) After 12 weeks: 23.0 Control group: Mean (SD) baseline: 27.5 (7.5) After 12 weeks: 23.0 SMD: -1.13 (-2.55, 0.28) P between groups: ns	-

Table 1. Short-term land-based aerobic capacity training *(Continued)*

			P between groups: ns			
Qualitative analysis	Moderate evidence	Moderate evidence	Limited evidence	Limited evidence	Moderate evidence	No evidence
Quantitative analysis	0.03 (-0.46, 0.51)*	0.99 (0.29, 1.68)*	-0.38 (-1.67, 0.90)	-0.27 (-0.79, 0.26)	na (statistical heterogeneity, I^2=70.3%)	na

* = pooled effect size

Table 2. Short-term land-based aerobic capacity and muscle strength training

Study	Functional ability	Aerobic capacity	Muscle strength	Self-reported pain	Disease activity	Radiological damage
Van den Ende 1996	HAQ Dynamic group: Mean (SD) baseline: 0.83 (0.61) After 12 weeks: 0.78 After 12 weeks follow up: 0.88 Control group: Mean (SD) baseline: 0.70 (0.61) After 12 weeks: 0.86 After 12 weeks follow up: 0.80 SMD: -0.54 (-1.11, 0.02) P between groups: After 12 weeks: ns After 12 weeks follow up: ns	SUBMAXIMAL BICYCLE TEST Dynamic group: Mean (SD) baseline: 27.6 (7.1) ml/kg/min After 12 weeks: 32.3 After 12 weeks follow up: 28.7 Control group:	ISOKINETIC EXT. 120 DEGREES/sec Dynamic group: Mean (SD) baseline: 81 (35) Nm After 12 weeks: 94.6 After 12 weeks follow up: 87 Control group: Mean (SD) baseline: 78 (48) Nm After 12 weeks: 77.3 After 12 weeks follow up: 75 SMD: 0.74 (0.16, 1.31) P between groups: After 12 weeks: P<.05 After 12 weeks follow up: ns	PAIN ON VAS Dynamic group: Mean (SD) baseline: 3.4 (2.0) cm After 12 weeks: 3.6 After 12 weeks follow up: 4.8 Control group: Mean (SD) baseline: 2.1 (1.6) cm After 12 weeks: 3.0 After 12 weeks follow up: 3.3 SMD: -0.53 (-1.09, 0.04) P between groups: After 12 weeks: ns After 12 weeks follow up: ns	N OF SWOLLEN JOINTS (range 0-26) Dynamic group: Mean (SD) baseline: 5.2 (3.2) After 12 weeks: 3.5 After 12 weeks follow up: 4.3 Control group: Mean (SD) baseline: 3.6 (3.3) After 12 weeks: 3.8 After 12 weeks follow up: 3.8 SMD: -0.76 (-1.34, -0.18) P between groups: After 12 weeks: P<.05 After 12 weeks follow up: ns	

Table 2. Short-term land-based aerobic capacity and muscle strength training *(Continued)*

			mean (SD) baseline: 25.8 (6.1) After 12 weeks: 26.1 After 12 weeks follow up: 26.3 SMD: 1.55 (0.91, 2.19) P between groups: After 12 weeks: P<.001 After 12 weeks follow up: ns		
Lyngberg 1994	FRIES INDEX Dynamic group: Mean (SD) baseline: 18 (6-30) After 12 weeks: 16 Control group: Mean (SD) baseline: 15 (0-37) After 12 weeks: 15 SMD: -0.11 (-0.91, 0.69) P between groups: ns	SUBMAXIMAL BICYCLE TEST Dynamic group: Median (range) baseline: 2.6 (1.5-4.2)	ISOKINETIC EXT. 30 degrees/sec Dynamic group: Median (range) baseline: 79 (35-181) Nm After 12 weeks: 77 Control group: Median (range) baseline: 76 (37-178) Nm After 12 weeks: 55 SMD: 0.26 (-0.55, 1.06) P between groups: ns	ERS Dynamic group: Median (range) baseline: 33 (2-97) After 12 weeks: 22 Control group: Median (range) baseline: 17 (6-48) After 12 weeks: 23 SMD: -0.04 (-0.84, 0.76) P between groups: ns	-

Table 2. Short-term land-based aerobic capacity and muscle strength training *(Continued)*

		l/min Increase after 12 weeks: 2.7 Control group: Median (range) baseline: 2.8 (1.8-3.9) l/min Increase after 12 weeks: 2.7 SMD: 0.15 (-0.65, 0.95) P between groups: ns				
Qualitative analysis	Moderate evidence	Moderate evidence	Moderate evidence	Moderate evidence	Moderate evidence	No evidence
Quantitative analysis (effect size)	-0.40 (-0.86, 0.06)*	na (statistical heterogeneity, I^2=86.2%)	0.47 (0.01 to 0.93)*	-0.53 (-1.09, 0.04)	na (statistical heterogeneity, I^2=51.5%)	na

* = pooled effect size

Table 3. Short-term water-based aerobic capacity and muscle strength training

Study	Functional ability	Aerobic capacity	Muscle strength	Self-reported pain	Disease activity	Radiological damage
Minor 1989	AIMS (physical activity) Dynamic group: Mean (SD) baseline: 4.9 (2.4) After 12 weeks: 3.7 Control group: Mean (SD) baseline: 4.0 (2.5) After 12 weeks: 4.9 SMD: -0.89 (-1.39, -0.38) P between groups: <.05	MAXIMAL TREADMILL TEST Dynamic group: Mean (SD) baseline: 19.3 (6.7) After 12 weeks: 23.2 Control group: Mean (SD) baseline: 17.4 (5.9) After 12 weeks: 17.3 SMD: 0.64 (0.14, 1.14) P between groups: <.05		PAIN (AIMS) Dynamic group: Mean (SD) baseline: 5.0 (1.6) After 12 weeks: 4.4 Control group: Mean (SD) baseline: 5.5 (1.6) After 12 weeks: 4.8 SMD: 0.06 (-0.43, 0.54) P between groups: ns	CLINICALLY ACTIVE JOINTS Dynamic group: Mean (SD) baseline: 10.5 (10.7) After 12 weeks: 5.9 Control group: Mean (SD) baseline: 12.4 (13.2) After 12 weeks: 12.0 SMD: -0.37 (-0.85, 0.12) P between groups: ns	
Sanford-Smit 1998	HAQ Dynamic group: Mean (SD) baseline: 19.2 (16.2) After 10 weeks: 16.8 Control group: Mean (SD) baseline: 12.1 (10.7) After 10 weeks: 8.4 SMD: 0.10 (-0.78, 0.98) P between groups: ns	DURATION ON TREADMILL Dynamic group: Mean (SD) baseline: 9.6 (6.1) After 10 weeks: 12.1 Control group: Mean (SD) baseline: 12.6 (7.8) After 10 weeks: 14.6 SMD: 0.07 (-0.82, 0.95) P between groups: ns	GRIP STRENGTH Dynamic group: Mean (SD) baseline: 28.5 (16.3) After 10 weeks: 32.9 Control Group: Mean (SD) baseline: 34.6 (10.0) After 10 weeks: 45.9 SMD: -0.38 (-1.27, 0.51) P between groups: ns		ACTIVE JOINT COUNT Dynamic group: Mean (SD) baseline: 8.3 (6.0) After 10 weeks: 7.5 Control group: Mean (SD) baseline: 10.6 (5.6) After 10 weeks: 7.1	

Table 3. Short-term water-based aerobic capacity and muscle strength training *(Continued)*

					SMD: 0.46 (-0.44, 1.35) P between groups: ns	
Qualitative analysis	Limited evidence	Limited evidence	Moderate evidence	Limited evidence	Moderate evidence	No evidence
Quantitative analysis (effect ize)	na (statistical heterogeneity, I^2=72.4%)	0.47 (-0.04, 0.98)*	-0.38 (-1.27, 0.51)	0.06 (-0.43, 0.54	na (statistical heterogeneity, I^2=59.9%)	na

* = pooled effect size

Table 4. Long-term land-based aerobic capacity and muscle strength training

Study	Functional ability	Aerobic capacity	Muscle strength	Self-reported pain	Disease activity	Radiological damage
Hansen 1993	HAQ Dynamic group: Median (25%-75%) baseline: 0.50 (0.50-0.75) After 24 months: 0.87 Control group: Median (25%-75%) baseline: 0.50 (0.13-0.88) After 24 months: 0.62 SMD after 24 months: 0.71 (-0.12, 1.54) P between groups: ns	-	ISOMETRIC EXTENSION Dynamic group: Median (25%-75%) baseline: 31 (26-38) kp After 24 months: 57 Control group: Median (25%-75%) baseline: 29 (19-35) kp After 24 months: 45 SMD: 0.93 (0.08, 1.79) P between groups: <.05	PAIN ON VAS Dynamic group: Median (25%-75%) baseline: 1.9 (1.2-2.6) After 24 months: 2.1 Control group: Median (25%-75%) baseline: 1.9 (1.5-2.3) After 24 months: 1.9 SMD: 0.35 (-0.46, 1.16) P between groups: ns	ESR Dynamic group: Median (25%-75%) baseline: 20 (6-42) mm/hr Increase after 24 months: 17 Control group: Median (25%-75%) baseline: 23 (12-32) Increase after 24 months: 28 SMD: -0.42 (-1.23, 0.40) P between groups: ns	X-RAY SCORE Dynamic group: Median (25%-75%) baseline: 37 (17-76) After 24 months: 46 (32-89) Control group: Median (25%-75%) baseline: 68 (58-89) After 24 months: 74 (65-97) SMD: 0.13 (-0.68, 0.93) P between groups: ns
De Jong 2003	MACTAR Dynamic group:	ERGOMETER TEST	ISOKINETIC TEST 60 DEGREES/sec		DAS4 Dynamic group:	LARSEN JOINT SCORE

Table 4. Long-term land-based aerobic capacity and muscle strength training *(Continued)*

	Mean (SD) baseline: 54.0 (4.8) After 24 months: 57.6 Control group: Mean (SD) baseline: 53.0 (5.0) After 24 months: 54.7 SMD after 24 months: 0.40 (0.16, 0.64) P between groups: at 12, 18 and 24 months: p<.05	Dynamic group: Mean (range) baseline: 162 (126-200) After 24 months: 170.2 Control group: Mean (range) baseline: 162 (162-200) After 24 months: 155.3 SMD after 24 months: 0.46 (0.22, 0.70) P between groups: <.001	Dynamic group: Mean (range) baseline: 165 (128-206) After 24 months: 191.1 Control group: Mean (range) baseline: 166 (115-227) Nw After 24 months: 175.6 SMD after 24 months: 0.32 (0.08, 0.55) P between groups: p <.001	Mean (SD) baseline: 3.3 (1.4) After 24 months: 2.4 Control group: Mean (SD) baseline: 3.4 (1.9) After 24 months: 2.7 SMD after 24 months: -0.14 (-0.38, 0.09) P between groups: ns	Dynamic group: Mean (SD) baseline: 1.5 (4.5) Increase after 12 months: 1.5 Increase after 24 months: 1.5 Control group: Mean (SD) baseline: 2.0 (5.0) Increase after 12 months: 2.0 Increase after 24 months: 2.0 SMD after 24 months: -0.17 (-0.41, 0.06) P between groups: ns
Qualitative analysis	Conflicting evidence	Moderate evidence	Moderate evidence	Limited evidence	Moderate evidence	Moderate evidence
Quantitative analysis (effect size)	na	0.46 (0.22, 0.70)	0.49 (-0.06, 1.04)*	0.35 (-0.46, 1.16)	-0.16 (-0.39, 0.06)*	-0.15 (-0.37, 0.08)*

* = pooled effect size

Table 5. Sensitivity analyses

Subcategory	Grade of evidence	Effect on/safe for:	Effect size	Grade of evidence	Effect on/ safe for	Effect size
	Conclusion 1 (based on all included studies)#			Conclusion 2 (based on the 4 studies with a methodological		

Table 5. Sensitivity analyses *(Continued)*

				quality of at least 8)##		
Short term aerobic capacity training (land-based)	Moderate evidence Moderate evidence Limited evidence	Functional ability Aerobic capacity Muscle strength	0.03 (-0.46, 0.51)* 0.99 (0.29, 1.68)* -0.38 (-1.67, 0.90)	-	-	-
	Limited evidence Moderate evidence No evidence	Self-reported pain Disease activity Radiological damage	-0.27 (-0.79, 0.26) n.a. n.a.	-	-	-
Short term aerobic capacity and muscle strength training (land-based)	Moderate evidence Moderate evidence Moderate evidence	Functional ability Aerobic capacity Muscle strength	-0.40 (-0.86, 0.06)* n.a. 0.58 (0.11, 1.04)*	Moderate evidence Moderate evidence Moderate evidence	Functional ability Aerobic capacity Muscle strength	-0.40 (-0.86, 0.06)* n.a. 0.58 (0.11, 1.04)*
	Moderate evidence Moderate evidence No evidence	Self-reported pain Disease activity Radiological damage	-0.53 (-1.09, 0.04) na na			

Conclusion 1 is based on all included studies
Conclusion 2 is based on the 4 included studies with a high methodological quality

APPENDICES

Appendix 1. MEDLINE search strategy

#1 rheumatoid arthritis

#2 arthritis[tiab]

#3 exercise therapy

#4 exercis*[tiab]

#5 motion therap*[tiab]

#6 "physical education and training"[MeSH Terms]

#7 physical education[tiab]

#8 training[tiab]

#9 gymnast*[tiab]

#10 physical fitness

#11 physical fitness[tiab]

#12 hydrotherap*

#13 hydrotherap*[tiab]

#14 water therap*

#15 water therap*[tiab])

16 randomized controlled trial

17 controlled clinical trial

18 randomized controlled trials

19 random allocation

20 double-blind method

21 single-blind method

22 clinical trial

23 clinical trials

24 "clinical trial"

#25 singl*

#26 doubl*

#27 trebl*

#28 tripl*)

#29 mask*

#30 blind*

#31 "latin square"

#32 placebos

#33 placebo*

#34 random*

#35 research design [mh:noexp]

#36 comparative study

#37 evaluation studies

#38 follow-up studies

#39 prospective studies

#40 cross-over studies

#41 control*

#42 prospective*

#43 volunteer*

#44 randomised controlled trial

#45 randomised controlled trials)

#46 #1 OR #2

#47 #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15

#48 #25 OR #26 OR #27 OR #28

#49 #16 OR #17 OR #18 #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR # 48

#50 #29 OR #30

#51 #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #50

FOUND REFERENCES UP TO DECEMBER 2008: 873

Appendix 2. EMBASE search strategy

1 Rheumatoid Arthritis/

2 arthritis.ti,ab)

3 exercise therapy/

4 exercis$.ti,ab

5 motion therap$.ti,ab

6 "Physical Education and Training"/

7 physical education.ti,ab

8 training.ti,ab

9 gymnast$.ti,ab

10 physical fitness

11 physical fitness.ti,ab

12 hydrotherap$.mp

13 hydrotherap$.ti,ab

14 water therap$.mp

15 water therap$.ti,ab)

16 randomized controlled trial/

17 clinical trial/

18 randomized.mp

19 random allocation.mp

20 randomization/

21 double-blind method.mp.

22 Double Blind Procedure/

23 single-blind method.mp

24 Single Blind Procedure/

25 clinical trial.mp

26 clinical trials.mp

27 singl$.mp

28 doubl$.mp

29 trebl$.mp

30 tripl$.mp

31 mask$.mp

32 blind$.mp

33 latin square.mp

34 placebo/

35 placebo$.mp

36 random$.mp

37 methodology/

38 comparative study/

39 comparative study.mp

40 exp evaluation/

41 exp follow up/

42 follow-up stud$.mp

43 exp prospective study/

44 prospective.mp

45 Crossover Procedure/

46 cross-over stud$.mp

47 volunteer$.mp

48 randomised.mp

49 OR/1-2

50 OR/3-15

51 OR/27-30

52 OR/16-26

53 OR/51-52

54 OR/31-32

55 OR/33-48

56 OR/54-55

57 49 AND 50 AND 53 AND 56

FOUND REFERENCES UP TO DECEMBER 2008: 783

Appendix 3. WEB of SCIENCE search strategy

1 "rheumatoid arthritis" (ti)

2 arthritis(ti)

3 "exercise therapy"

4 exercis*

5 "motion therap*"

6 "physical education"

7 training

8 gymnast*

9 "physical fitness"

10 hydrother*

11 "water therap*"

12 "randomized controlled trial*"

13 "controlled clinical trial*"

14 "random allocation"

15 "double-blind method"

16 "single-blind method"

17 "clinical trial*"

18 "latin square"

19 placebo*

20 random*

21 "research design"

22 "comparative stud*"

23 "evaluation stud*"

24 "follow-up stud*"

25 "prospective stud*"

26 "cross-over stud*"

27 control*

28 prospective*

29 volunteer*

30 "randomised controlled trial*"

31 singl*

32 doubl*

33 trebl*

34 tripl*

35 mask*

36 blind*

37 or 1-2

38 or 3-11

39 or 12-30

40 or 31-34

41 or 39-40

42 or 35-36

43 37 AND 38 AND 41 AND 42

FOUND REFERENCES UP TO DECEMBER 2008: 316

Appendix 4. GRADE approach

Underlying methodology quality rating

Randomized trials High

Downgraded randomised trials Moderate

Double-downgraded randomised trials Low

Triple-downgraded randomised trials Very low

The quality rating can be decreased or increased by the review authors by the following factors:

Decreased:

- Limitations in the design and implementation of available studies suggesting high likelihood of bias
- Indirectness of evidence (indirect population, intervention, control, outcomes)
- Unexplained heterogeneity or inconsistency of results (including problems with subgroup analyses)
- Imprecision of results (wide confidence intervals)
- High probability of publication bias

Increased:

- Large magnitude of effect
- all plausible confounding would reduce a demonstrated effect or suggest a spurious effect when results show no effect
- dose-response gradient

WHAT'S NEW

Date	Event	Description
22 April 2009	New citation required but conclusions have not changed	Change in authors
22 April 2009	New search has been performed	Update of a withdrawn review (See published notes for details on update).
26 May 2008	Amended	CMSG ID: C119-R

HISTORY

Protocol first published: Issue 4, 2007

Review first published: Issue 4, 2009

Date	Event	Description
26 May 2008	Amended	Converted to new review format.

CONTRIBUTIONS OF AUTHORS

EJ Hurkmans
Coordinator and writer of this review and independent review author for: selection based on title and abstract; check for inclusion criteria; assessment of methodological quality; data extraction.

F van der Giesen
Independent review author for: selection based on title and abstract; check for inclusion criteria; assessment of methodological quality; data extraction.

TPM Vliet Vlieland
In cases of disagreement between the two independent review authors, Vliet Vlieland made the final decision.

CHM van den Ende
Overview of the whole review process, helped with writing the protocol and review.

DECLARATIONS OF INTEREST

None known

SOURCES OF SUPPORT

Internal sources

- None, Not specified.

External sources

- None, Not specified.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Analysis: in the protocol the analysis consisted of a quantitative analysis and, if a meta-analysis was not appropriate, a qualitative analysis. Furthermore, an overall grading of the evidence according to Tugwell 2004 was to be applied. During the process of developing this review it was found that there were too many analyses, the readability of the results was not very clear. For this reason we decided to apply a best-evidence synthesis for all data and, if possible, a meta-analysis was applied to establish the effect size.

Subgroup analysis: in the protocol we indicated our intention to perform various subgroup analyses. However, in the included studies no specific data were reported for these subgroups, for example separate data for men and women or patients with a high or low disease activity. As a number of studies were quite old, it was considered not feasible to ask for the original data sets.

NOTES

This is an update of the withdrawn review "Dynamic exercise for treating rheumatoid arthritis" (van den Ende 1998). We changed the title from "Dynamic exercise therapy for treating rheumatoid arthritis" into "Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis" because we think this is more correct.

In the protocol we intended to grade the evidence according to the grading system of Tugwell 2004. During the process of analysing the data we used three different methods (meta-analysis, best-evidence synthesis, and Tugwell), which did not make the results clear. So for the readibility of this review we decided to perform best-evidence synthesis and meta-analysis.

INDEX TERMS

Medical Subject Headings (MeSH)

*Exercise Therapy; *Resistance Training; Arthritis, Rheumatoid [*rehabilitation]; Oxygen Consumption; Physical Fitness [*physiology]; Randomized Controlled Trials as Topic

MeSH check words

Humans