LITERATURE REVIEW

The Effectiveness of Hydrotherapy in the Management of Rheumatoid Arthritis: A Systematic Review

Khamis Y. Al-Qubaeissy MD, Francis A. Fatoye PhD, Peter C. Goodwin PhD & Abebaw M. Yohannes* PhD, MSc, FCCP

Department of Health Professions, Manchester Metropolitan University, Manchester, UK

Abstract

Background. Hydrotherapy is frequently indicated for the rehabilitation of patients with rheumatoid arthritis (RA); nevertheless, there has been inadequate appraisal of its effectiveness. The potential benefits of hydrotherapy for patients with RA are to improve and/or maintain functional ability and quality of life.

Objectives. The aim of this systematic review was to evaluate the effectiveness of hydrotherapy in the management of patients with RA.

Method. AMED, CINAHL, EMBASE, MEDLINE, PubMed, Science Direct and Web of Science were searched between 1988 and May 2011. Keywords used were rheumatoid arthritis, hydrotherapy, aquatic physiotherapy, aqua therapy and water therapy. Searches were supplemented with hand searches of references of selected articles. Randomized controlled trials were assessed for their methodological quality using the Physiotherapy Evidence Database (PEDro) scale. This scale ranks the methodological quality of a study scoring 7 out of 10 as 'high quality', 5–6 as 'moderate quality' and less than 4 as 'poor quality'.

Results. Initially, 197 studies were identified. Six studies met the inclusion criteria for further analysis. The average methodological quality for all studies was 6.8 using the PEDro scale. Most of the studies reported favourable outcomes for a hydrotherapy intervention compared with no treatment or other interventions for patients with RA. Improvement was particularly noted in reducing pain, joint tenderness, mood and tension symptoms, and increasing grip strength and patient satisfaction with hydrotherapy treatment in the short term.

Conclusions. There is some evidence to suggest that hydrotherapy has a positive role in reducing pain and improving the health status of patients with RA compared with no or other interventions in the short term. However, the long-term benefit is unknown. Further studies are needed. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords

Rheumatoid arthritis; hydrotherapy; aquatic physiotherapy; aqua therapy; water therapy; pain; quality of life; physical activity

*Correspondence

Abebaw Mengistu Yohannes, Reader in Physiotherapy, Department of Health Professions, Manchester Metropolitan University, Elizabeth Gaskell Building, Hathersage Road, Manchester, M13 0JA, UK. Tel: 44 (0) 161 247 2943; Fax: 44 (0) 161 247 6571.
Email: A.Yohannes@mmu.ac.uk

Published online 16 July 2012 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/msc.1028

Introduction

Rheumatoid arthritis (RA) is a chronic, systemic, inflammatory, symmetrical polyarthritis disease that can be both erosive and deforming (Arthritis Research UK, 2011; McMahon and Allard, 2002; Waldburger and Firestein, 2008). It affects many organs and tissues in the body, although the joints are usually the most severely affected (Arthritis Research UK, 2011; National Institute for Health and Clinical Excellence (NICE), 2009; Waldburger and Firestein, 2008). The disease is characterized by joint pain, swelling, tenderness and the
destruction of the synovial joints, leading to severe disability and premature mortality (NICE, 2009; Tehlirian and Bathon, 2008; Waldburger and Firestein, 2008). It results from an immune system disturbance caused by the interaction of immunological, genetic, environmental and hormonal factors (Arthritis Research UK, 2011; NICE, 2009; Waldburger and Firestein, 2008). It typically affects the small joints of the hands, especially the knuckles and second joints, such as metacarpophalangeal joint and proximal interphalangeal joint, as well as the wrists, knees, ankles, elbows, shoulders and feet (Tehlirian and Bathon, 2008). Usually, both sides of the body are equally affected in a symmetrical fashion, although any synovial joint can be affected. The lumbar spine and hips are often spared (Tehlirian and Bathon, 2008). In the USA, the average annual incidence of RA is 0.5 per 1,000 persons per year (Drosos, 2004; Tehlirian and Bathon, 2008), and in the UK it affects approximately 0.5–1% of the population (McMahone and Allard, 2002; Symmons et al., 1994, 2002). The overall prevalence of RA worldwide in the general population is 1–2%, and it affects more women than men; this prevalence is expected to rise to 5% of people by the age of 70 years in the next few decades (NICE, 2009; Symmons et al., 1994; Tehlirian and Bathon, 2008). In the UK, there are 100 new cases of inflammatory joint disease per thousand of the population per year, of whom 24 will have RA (Söderlin et al., 2002). The direct costs to the National Health Service are estimated at £560 million and to the wider economy (e.g. loss of earnings due to ill health) are estimated at £1.8 billion per annum (Comptroller and Auditor General, 2009), whereas the total costs of RA in the UK, together with the indirect costs and the effects of early mortality and lost productivity, have been approximated at between £3.8 and £4.75 billion per year (NICE, 2009).

Exercise is the cornerstone of the treatment of RA and it improves function, muscle strength and general well-being (Hurkmans et al., 2009; van den Ende et al., 2008; Vliet Vlieland and van den Ende, 2011). The term ‘hydrotherapy’ or ‘aquatic exercise’ is defined as exercise in warm water under supervision by utilizing the buoyancy, assistance and resistance of warm water to relieve pain, induce muscle relaxation and promote more effective exercise (Campion, 1997; Eversden et al., 2007; Hall et al., 2008; Schrepfer, 2002). Hydrotherapy is a safe and efficient medium treatment modality for achieving exercise-related goals and it is commonly used as part of a rehabilitation intervention for patients with rheumatic disease (Beardmore, 2008; Rintala et al., 1996).

Unblinded studies that examined the efficacy of hydrotherapy in patients with RA demonstrated a reduction in pain and an increase in quality of life (QoL), muscle strength, aerobic conditioning and physical functioning (Danneskiold-Samsøe et al., 1987; Hart et al., 1994; Minor et al., 1989). However, the generizability of the findings were limited because of small sample sizes and a lack of controlled intervention.

To our knowledge, there has been no recent exclusive systematic review to examine the efficacy of hydrotherapy for patients with RA. We hypothesized that hydrotherapy therapy is far superior than other types of therapy, including ‘usual care’, for improving QoL and physical activity in patients with RA.

The aim of this review was to synthesize the available literature on the efficacy of hydrotherapy in the management of patients with RA.

Materials and methods
Identification and selection criteria
An electronic database search of AMED, CINAHL, the Cochrane Library, EMBASE, MEDLINE, ProQuest, Pub Med, Science Direct and the Web of Science was conducted (1988 to May 2011). In order to standardize the patient sample included, the search was conducted from 1988 [which was the date of the publication of the American College of Rheumatology (ACR) criteria for RA] to May 2011 (Arnett et al., 1988). The search was limited to human adults (age >18 years) across all articles published in English. The keywords used were: ‘rheumatoid arthritis’, ‘hydrotherapy’, ‘aquatic physiotherapy’, ‘aqua therapy’ and ‘water therapy’. Keyword combinations were: ‘rheumatoid arthritis and hydrotherapy’, ‘rheumatoid arthritis and aquatic physiotherapy’, ‘rheumatoid arthritis and aqua therapy’ and ‘rheumatoid arthritis and water therapy’. Studies that used the following keywords were excluded from this literature search: ‘colon irrigation’, ‘water birth’, ‘Kneipp therapy’, ‘spa therapy’, ‘whirlpool therapy’, ‘contrast baths’ and ‘balneotherapy’. There is a lack of clarity in the usage of the terms ‘hydrotherapy’ and ‘balneotherapy’ (Bender et al., 2005). Hydrotherapy uses water as a treatment, while balneotherapy uses natural thermal mineral water (Bender et al., 2005). Although these terms have often been used interchangeably,
balneotherapy is not easily accessible to healthcare professionals and so studies involving this treatment were excluded.

Trials investigating solely the physiological responses (such as heart rate, blood pressure and renal function) of subjects immersed or exercising in water were also excluded.

The database search was supplemented by a manual search of: Clinical Journal of Rheumatology, Annals of the Rheumatic Disease, British Medical Journal, Physiotherapy, Arthritis and Rheumatism, Rheumatology and Journal of Rheumatology and Physical Therapy. Journals were searched from 1988 to May 2011. A further hand search of the bibliographic references in the extracted articles and existing reviews was also conducted to identify potential studies that were not captured by the electronic database searches. To ensure that all of the relevant articles were obtained, an iterative process was used.

Inclusion and exclusion criteria for considering studies for this review

Studies were included if:

- they were randomized controlled trials (RCTs);
- they were published in the English language;
- they included participants aged 18 years or above who had been diagnosed with RA according to the 1987 ACR criteria (Arnett et al., 1988) or they used the criteria of Steinbrocker (Steinbrocker et al., 1949);
- a water-based intervention (hydrotherapy) had been used in the study, and compared with the results without intervention;
- patients had received a minimum of four weeks of hydrotherapy intervention.
- they used one of the following outcome measures: pain, patient global assessment, activity of daily living (ADL), physical function, disease activity and QoL (Boers et al., 1994; Haigh et al., 2001).

Articles were excluded if:

- they had insufficient information available (abstract only);
- they did not involve an RCT;
- they were not adult trials (juvenile trials);
- they did not involve human trials;
- they included participants without rheumatic diseases;
- the treatment modality included balneotherapy, Kneipp therapy, mud therapy or sulphur therapy;
- they were not written in English (even if the abstract was in English);
- participants were primarily and predominantly diagnosed with osteoarthritis, fibromyalgia syndrome, back pain, neurological disease or osteoporosis.

Assessment of the validity of the study

Two reviewers (A.M.Y. and F.A.F.) made the decisions regarding the inclusion of the relevant articles in the present review. They independently applied the inclusion/exclusion criteria to papers identified by the literature search and classified the identified studies according to predetermined criteria. The abstracts were reviewed first and, if deemed appropriate, the full papers were then reviewed and scored. The methodological quality of each study was reviewed by using the Physiotherapy Evidence Database (PEDro) scale (Maher et al., 2003). A consensus method was used to solve any dispute regarding the inclusion or exclusion of a particular study. When there was disagreement, consensus was sought, but when disagreement persisted, a third independent reviewer (P.G.) made the final decision.

The PEDro scale contains 11 items (Table 1). The first item represents the external validity of the trial. This item is not included in the calculation of the total PEDro score (maximum 10); therefore, our score was based on items 2 to 11 and the PEDro score was thus a score out of 10. These items are scored either yes (1 point) or no (0 points). The individual item scores and the total PEDro scores have been shown to be reliable (Maher et al., 2003). A study that scores 7 (i.e. scores positive in seven out of ten criteria) is considered to have a high methodological quality, a score of 5–6 a moderate methodological quality and a score between 0 and 4 is regarded as poor quality (Kollen et al., 2009; Maher et al., 2003; Moseley et al., 2002). Although the PEDro scale is scored out of 10 (Maher et al., 2003; Sherrington et al., 2000), the maximum achievable score for a high-quality study is 8 because it is difficult to blind the therapist delivering the intervention or the participants in a trial of hydrotherapy rehabilitation (Maher et al., 2003; Sherrington et al., 2000). The PEDro scores for the present review ranged from 4 to 8 out of the maximum possible score of 10, without including the first item of the PEDro scale (see above).
Data collection and analysis

Articles fulfilling the inclusion criteria were subsequently assessed for methodological quality using the criteria list and operational instructions outlined and recommended by the PEDro for the quality assessment of RCTs (Maher et al., 2003; Sherrington et al., 2000), as shown in Table 1.

Methodological quality of the studies

The methodological quality of the studies ranged from 5 to 8 on the PEDro scale of internal validity (Table 3), with a mean score of 6.8. Four studies were of high quality, whereas two were of moderate quality. Two studies (Sanford-Smith et al., 1998; Stenstrom et al., 1991) failed to report or describe whether an intent-to-treat analysis or concealment of the treatment allocation was used. In three studies (Eversden et al., 2007; Hall et al., 1996; Sanford-Smith et al., 1998), the outcome assessor was blinded to the intervention. All of the participants were randomized in the included trials; however, only three studies (Bilberg et al., 2005; Eversden et al., 2007; Hall et al., 1996) specified the methods used. Two studies used optimal allocation using a computer program (Bilberg et al., 2005; Eversden et al., 2007) and one used block randomization (Hall et al., 1996).

Data extraction

The two reviewers (A.M.Y. and F.A.F.) independently extracted data using a standardized form regarding: the author(s), place and date of publication, study design, sample size and percentage of female sample, mean age, the interventions, type of outcome measures, and follow-up or failure to follow-up, to ensure that no significant information was omitted from the review. Meta-analysis or statistical pooling were not considered because of the heterogeneity among the studies, including the small sample size, variations in symptoms and duration, interventions and the reporting of the outcomes.

Participants

The six studies described above included both men and women (total no = 419); 326 (78%) of the participants were women. The participants’ age across the studies ranged from 18–80 years. The average number of participants in the treatment group post-randomization and before any withdrawals was 29 (range 12–57), with only three studies having groups with more than 30 participants (Eversden et al., 2007; Hall et al., 1996; Rintala et al., 1996).

Results

A total of 197 studies were identified, based on the key search terms and the hand search of bibliography references (CINAHL 12; Medline 42; PubMed 122; AMED 13; manual search eight). After the initial screening of the titles and abstracts, 32 studies were found to satisfy the inclusion criteria and were further scrutinized for the present systematic review (see Figure 1). From the six studies that were of high enough quality to analyse are presented in Table 2.
Outcome measures

RA affects physical, social and psychological aspects of patients’ health status or quality of life. The outcome measures that were used in the present review reflected one or more of the variables (Fitzpatrick et al., 1992; Hakala, 1997).

**Pain:** A pain scale was used in all the reviewed studies. Scores on these scales were measured before and after the intervention. Various instruments were used to measure sensory pain. The 10-cm visual analogue scale (VAS) was the tool used most commonly (Langley and Sheppeard, 1984). VAS was used in three studies (Eversden et al., 2007; Rintala et al., 1996; Stenstrom et al., 1991). Another instrument that was used, by Hall et al. (1996), to assess pain was the McGill Pain Questionnaire (Melzack, 1975). Moreover, pain subscales from a variety of self-reported questionnaires were used, such as the Arthritis Impact Measurement Scale (AIMS) (Meenan et al., 1980), Health Assessment Questionnaire (HAQ) (Bruce and Fries, 2005; Felson et al., 1993; Fries et al., 1980) and the Short Form-36 (SF-36) (Ware and Sherbourne, 1992). Rintala et al. (1996) used pain as a primary outcome measure and found that there was a statistically significant reduction in the level of pain after use of a water exercise programme in patients with RA. None of the studies used pain as an outcome measure for a power calculation to determine the sample size.
Table 2. Summary of studies meeting the selection criteria for inclusion in the systematic review for RA.

<table>
<thead>
<tr>
<th>Authors, country of study</th>
<th>Sample n (female)%</th>
<th>Study design</th>
<th>Drop-outs</th>
<th>Mean age (SD) in years</th>
<th>Intervention duration</th>
<th>Outcome measures</th>
<th>Patient assessment/follow-up</th>
<th>Results/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hall et al., 1996; UK</td>
<td>139 (66%)</td>
<td>RCT 1</td>
<td>1</td>
<td>58.2 (11.1)</td>
<td>A: aquatic exercise (n=35) 30 minutes twice weekly for 4 weeks, Pain; using McGill Questionnaire</td>
<td>Baseline</td>
<td>No significant differences between interventions in terms of pain (all patients demonstrated a significant pain reduction (p ≤ 0.005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B: land-based exercise (n=34) Ritchie articular index (RAI)</td>
<td>4 weeks</td>
<td>Significant reduction in joint tenderness in a number of tender joints in hydrotherapy group (p = 0.03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C: immersion (n=35) Morning stiffness duration</td>
<td>3 months</td>
<td>Grip strength, wrist ROM, duration of morning stiffness and CRP levels did not change significantly (p ≥ 0.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D: land relaxation (n=35) Grip strength (digital monitor inflated to 20 mm Hg)</td>
<td>Post treatment</td>
<td>Significant increase in knee ROM, mainly in women in hydrotherapy group (p ≤ 0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wrist and knee ROM; using a standard goniometer</td>
<td></td>
<td>Significant improvement in mood and tension occurred for all patients after treatment in both groups; the effect was most marked in women, with a greater effect in the hydrotherapy group (p = 0.003)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AIMS-2 for health status</td>
<td></td>
<td>All groups reported similar perceptions of the effectiveness of the interventions at pre-test and post-test (p ≤ 0.0001)</td>
<td></td>
</tr>
<tr>
<td>Sanford-Smith et al., 1998; Canada</td>
<td>24 (75%)</td>
<td>RCT 4</td>
<td>5</td>
<td>58.4 (11.6)</td>
<td>A: aquaerobics group 3 times/week for 10 weeks AJC</td>
<td>Baseline (one week prior)</td>
<td>There were no between-group differences; however, both groups showed a similar decrease in AJC and ESR (p ≥ 0.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B: ROM group ESR</td>
<td>Post treatment assessment occurred within one week after the completion of the 10-week exercise programme</td>
<td>Both groups demonstrated an improvement in grip strength (p ≤ 0.05), but there was no significant difference between the groups</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grip strength measured using Martin Vigorimeter (Hilside Medical Supplies Limited, Nottingham, UK)</td>
<td></td>
<td>Both group showed an increase in exercise tolerance (p ≤ 0.05)</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Design</td>
<td>Patients</td>
<td>Baseline</td>
<td>Intervention</td>
<td>Outcome Measures</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>--------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
</tr>
</tbody>
</table>
| Bilberg et al., 2005; Sweden |         | RCT   | 49       |          | Sub-maximum ergometer cycle (Astrand, Varberg, Sweden) for aerobic capacity as primary outcome measure | SF-36 for health status as primary outcome measure, HAQ for function | Baseline showed a statistically significant improvement in two components of HAQ in the control group ($p \leq 0.05$) and no significant improvement in the aqua-aerobics group ($p \geq 0.05$). No significant between-group effects for duration and peak workload on treadmill ($p \geq 0.05$).
<p>| Eversden et al., 2007; UK |         | RCT   | 55.2 (13.3) |          | Primary outcome measure was self-rated overall effects on a Likert 7-point scale | Secondary outcome measure include: VAS pain, ten-meter walk speed, HAQ, EQ-5D utility, EQ-VAS | There were no significant differences between groups in terms of changes to HAQ ($p = 0.09$), EQ-5D utility score ($p = 0.61$), EQ-VAS ($p = 0.57$) and pain VAS ($p = 0.40$). |</p>
<table>
<thead>
<tr>
<th>Authors, country, origin of study</th>
<th>Sample n (female)%</th>
<th>Study design</th>
<th>Drop outs</th>
<th>Mean age (SD) in years</th>
<th>Intervention</th>
<th>Outcome measures</th>
<th>Patient assessment/follow-up</th>
<th>Results/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rintala et al., 1996; Finland</td>
<td>34 (85%)</td>
<td>RCT</td>
<td>0</td>
<td>48 (10)</td>
<td>A: aquatic exercise (n=18) 45-60 minutes twice a week for 12 weeks, setting and pool temperature 30°C</td>
<td>VAS pain</td>
<td>Baseline</td>
<td>Pain more diminished in experimental group than in control group (p ≤ 0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B: no-treatment control (n=16)</td>
<td>Joint mobility by using signals of functional impairment</td>
<td>Post-treatment (12 weeks)</td>
<td>Joint mobility improved in experimental group (p ≤ 0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Muscle strength and endurance by using digital dynamometer</td>
<td></td>
<td>Muscle strength and endurance improved in experimental group compared with control group (p ≤ 0.001)</td>
</tr>
<tr>
<td>Stenstrom et al., 1991; Sweden</td>
<td>60 (86%)</td>
<td>RCT</td>
<td>5</td>
<td>52 (11.2)</td>
<td>A: training group (n=30), once weekly in group of 5, for 40 minutes, for 4 years in temperature of 34°C in hospital pool (each year there is a vacation for 2.5 months)</td>
<td>Ritchie’s articular index for disease activity</td>
<td>Post training (4 years)</td>
<td>No significant difference between the groups in Ritchie’s articular index, Larsen’s radiological index, soft tissue swelling or laboratory markers (p &gt; 0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B: comparison group (n=30)</td>
<td>Larsen radiological index</td>
<td>Laboratory inflammatory markers</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sphygmomanometer cuff for grip strength</td>
<td></td>
<td>Improvised right hand grip strength in training group (p ≤ 0.01); decreased grip strength in left hand of comparison group (p &gt; 0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VAS for pain and functional tests such as outdoor walking, indoor walking, lifting, learning forward and rising</td>
<td>No significant difference between the groups in VAS or functional tests (p &gt; 0.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activity level such as exercise habits two open-ended questions</td>
<td>Significant difference in activity levels between the groups in training group compared with comparison group (p ≤ 0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Two-year follow-up at the end of the training period; the difference between the training and comparison groups was significant (p ≤ 0.001)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AIMS-2, Arthritis Impact Measurement Scale version 2; CRP, C-reactive protein; EQ-SD, EuroQoL; EQ-VAS, health-related QoL; HAQ, Health Assessment Questionnaire, AJC, active joint count; ESR, erythrocyte sedimentation rate; QoL, quality of life; RCT, randomized controlled trial; ROM, range of motion; SD, standard deviation; SF-36, Short Form-36; VAS, visual analogue scale.
Physical function: HAQ was the most commonly used instrument in the reviewed studies (Bruce and Fries, 2005; Felson et al., 1993; Fries et al., 1980). It was used in three studies (Bilberg et al., 2005; Eversden et al., 2007; Sanford-Smith et al., 1998). However, only one of these (Bilberg et al., 2005) found a significant improvement in physical function compared with the control group. HAQ was used as a primary outcome measure in one study (Sanford-Smith et al., 1998). Sanford-Smith et al. (1998) showed a trend for an improvement in physical function using the total HAQ score in the aqua-aerobics group compared with the control group; however, this was not statistically significant.

Health status: The category of health status was investigated in the three studies using the EuroQoL (EQ-5D) (Bilberg et al., 2005; Eversden et al., 2007; Hall et al., 1996). Hurst et al. (1997) and Eversden et al. (2007) used the EQ-5D to examine the efficacy of hydrotherapy for improving health status. The findings from both studies showed that there was no statistically significant difference in health status between the hydrotherapy and control groups. Similarly, Bilberg et al. (2005) administered the SF-36 (Ware and Sherbourne, 1992) and showed that, while there was a significant within-group improvement from baseline to post-treatment in the hydrotherapy group, these differences were not statistically significant between the two groups. Hall et al. (1996) used AIMS-2 (Meenan et al., 1992) and demonstrated a statistically significant improvement for all of the participants in both groups in the category of mood and tension. Women in the hydrotherapy group showed a statistically significant reduction in the level of tension and mood compared with those in the control group.

Disease activity: In terms of disease activity, a variety of categories were measured separately in four studies (Bilberg et al., 2005; Hall et al., 1996; Sanford-Smith et al., 1998; Stenstrom et al., 1991), such as morning stiffness, joint tenderness, joint swelling, grip strength and laboratory markers [acute-phase reactants such as C-reactive protein (CRP)]. The results of Bilberg et al. (2005) indicated that grip strength of the left hand increased significantly in the training group compared with the control group between 0–3 months ($p < 0.001$). This contrasted with the findings of Hall et al. (1996) and Sanford-Smith et al. (1998), who did not find any significant difference between the groups in terms of grip strength, duration of morning stiffness and CRP.

<table>
<thead>
<tr>
<th>Study</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
<th>Item 9</th>
<th>Item 10</th>
<th>Item 11</th>
<th>Total score (/10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stenstrom et al., 1991; Sweden</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>YY</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>8/10 high quality</td>
</tr>
<tr>
<td>Hall et al., 1996; UK</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>8/10 high quality</td>
</tr>
<tr>
<td>Sanford-Smith et al., 1998; Canada</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>8/10 high quality</td>
</tr>
<tr>
<td>Bilberg et al., 2005; Sweden</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>8/10 high quality</td>
</tr>
<tr>
<td>Eversden et al., 2007; UK</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>8/10 high quality</td>
</tr>
<tr>
<td>Rintala et al., 1996; Finland</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>YY</td>
<td>8/10 high quality</td>
</tr>
</tbody>
</table>

Y, yes (= 1); N, no (= 0)

Table 3. Methodological quality using the Physiotherapy Evidence Database (PEDro) scale scoring the items out of 10

Al-Qubaeissy et al. The Effectiveness of Hydrotherapy in RA

level or erythrocyte sedimentation rate ($p > 0.05$). In Stenstrom et al. (1991), grip strength improved significantly in the right hand of training group participants ($p < 0.01$) while it deteriorated in the left hand of the control group ($p > 0.05$). Hall et al. (1996) also showed that there was a significant reduction in joint tenderness in the hydrotherapy group.

Patient perception: Patients’ perception of hydrotherapy treatment was investigated in two studies (Eversden et al., 2007; Hall et al., 1996). Hall et al. (1996) used a five-point Likert-type perception scale, which was designed by Langley and Sheppeard (1984); their findings were unexpected and showed that both groups reported similar perceptions of the effectiveness of the intervention. Eversden et al. (2007) used a seven-point scale and their findings showed that the largest set of significant clusters of feeling ‘very much better’ was in the hydrotherapy group compared with the land exercise group.

A Scandinavian study undertaken by Stenstrom et al. (1991) failed to show any statistically significant differences in pain rating, functional outcomes tests (Stenstrom et al., 1990), Ritchie’s articular index (Ritchie et al., 1968), Larsen’s radiological index (Larsen et al., 1977), soft tissue swelling or laboratory parameters between the training group and the control group. Perceptions of activity levels were measured in this study using self-reported questions (e.g. ‘what do you think is positive regarding the training?’ and ‘what do you think is negative regarding the training?’) recommended for use in patients with chronic pain (Dolce et al., 1986; Doleys et al., 1982). There was a significant difference in the perceptions of activity levels between the treatment group compared with the control group ($p < 0.01$). The two-year follow-up data showed that there was a statistically significant difference in the perception of activity levels between the treatment and control groups ($p < 0.001$).

Hall et al. (1996) showed that hydrotherapy was effective in improving physical and emotional aspects in patients with RA. This finding indicates that hydrotherapy provided greater benefits in terms of physical and psychological functioning in comparison with the control group. AIMS-2 measured mood and tension, and a significant improvement in psychological well-being was found during the follow-up period. However, the hydrotherapy group derived a significant improvement in joint tenderness and knee range of movement in women only.

Rintala et al. (1996) assessed the efficacy of a water-based exercise programme on chronic pain in patients with RA. Pain was assessed using VAS (Ekdahl et al., 1989; Fries, 1983). These authors also assessed ranges of movement by measuring joint mobility (Eberhardt et al., 1988), muscle strength and endurance (Talvitie, 1991). The researchers (Rintala and co-workers) randomly allocated 34 patients with RA to aquatic exercise ($n = 18$) or the control group ($n = 16$). The aquatic exercise group undertook muscle strength, endurance and joint mobility exercises in sessions lasting 45–60 minutes, twice a week for 12 weeks. The control group participated in their daily activity with no additional exercise during the study period. The major findings of this study were decreased pain, and increased muscle strength and endurance in the hydrotherapy group compared with the control group during the 12-week training period.

Sanford-Smith et al. (1998) recruited 24 participants (19 females and five males), with a mean age of 58.4 years, to participate in their study. Subjects were randomly allocated to the aqua-aerobic exercise group or the range of motion (ROM) exercise group. The aqua-aerobics sessions were held three times per week for ten weeks. Each session consisted of an hour of exercises performed in a hydrotherapy pool heated to $36^\circ$C. Fifteen minutes of warm-up aerobic stretches for the spine, chest and extremities was followed by 20–25 minutes of aerobics exercise. Subjects exercised to a maximum target heart rate of 70% exercise tolerance (Beals et al., 1985; Ekblohm et al., 1974; Minor et al., 1988; Nordmar et al., 1981). The control group participants received a ROM exercise and isometric strength exercises programme for ten weeks. Nonetheless, the results failed to reveal a differential effect between the intervention and control groups.

Bilberg et al. (2005) undertook a study in which they hypothesized that pool exercise for three months would improve patients’ aerobic capacity, functional ability and perception of physical health. Forty-seven participants (42 women and five men) were divided into two groups (the treatment group and the control group). The treatment group exercised twice a week for 12 weeks in groups of eight or nine patients in a temperate pool. The duration of each session was 45 minutes and the exercise was of moderate aerobic intensity. The patients in the control group continued with their usual daily activities, and provided a home exercise programme. The outcome measurements were carried out at baseline and at three months post-intervention for both groups. The patients in the training group were followed up to
six months after completion of the study. Aerobic capacity, estimated using a sub-maximum ergometer cycle (Åstrand, 2003), and the physical component of the SF-36 were chosen as the primary outcome measures. The study was unable to confirm whether the intervention was effective in improving aerobic capacity and quality of life. However, a significant improvement was found in the hydrotherapy group for the secondary outcome measures, isometric shoulder endurance, grip force, dynamic endurance of the lower extremities (chair test) and muscle function of the lower extremities, compared with the control group. The chair test was assessed by counting the maximum number of times that the patient was able to get up from a chair during one minute (Mannerkorpi and Ekdahl, 1997) and the isometric shoulder endurance test, which is used to measure the isometric endurance of the shoulder abductor muscles. This was measured as the maximum length of time that a person was able to hold his/her arm at 90-degree abduction with a 1-kg cuff attached proximally to the wrist joint (Mannerkorpi et al., 1999) at baseline and three months post-treatment. The difference in all of the primary and secondary outcome measures between baseline assessment and follow-up for the training group were statistically significant, with the exception of aerobic capacity.

Eversden et al. (2007) evaluated the effects of hydrotherapy with exercises versus land exercises on the overall response to treatment, physical function and QoL of patients with RA. These authors designed a programme of 30-minute hydrotherapy sessions once a week for six weeks (at 35 °C), with a control group on a land-based programme for six weeks. Patients were randomly allocated to hydrotherapy or land-based exercises using sealed opaque envelopes indicating their treatment allocation. The participants performed warm-up exercises for ten minutes using mobilizing and stretching exercises. The core exercises, repeated ten times a week, focused on joint mobility, muscle strength and functional activities.

The primary outcome measure applied in this study was self-rated QoL, in which the effect of treatment was measured as the change on a seven-point scale ranging from 1 (very much worse) to 7 (very much better) (Richards and Scott, 2002). Secondary outcomes were collected at baseline, on the day of the last treatment session and three months post-treatment. Pain was assessed using a 10-cm VAS, where 0 cm represented no pain and 10 cm represented severe pain (Langley and Sheppeard, 1984). Physical function was assessed using the HAQ (Bruce and Fries, 2005; Felson et al., 1993; Fries et al., 1980). The ten-metre walk speed was used to assess lower limb function; this primarily indicated in patients with neurological problems and had also been used by the authors who carried out the previous pilot study (Eversden et al., 2001; Wade et al., 1987). The EQ-5D valuation questionnaire comprised a self-report of health-related QoL (EQ-VAS) and a health status valuation (EQ-5D index or utility score) (Hurst et al., 1997). Eversden et al. (2007) showed that RA patients who attended outpatient clinics were more likely to report feeling much better or very much better if they were treated with hydrotherapy than if they were treated with exercises on land. This benefit was reported immediately after completion of the treatment; there was no difference between treatment groups in the secondary outcome measures.

Discussion

The objective(s) of the present systematic review was to evaluate the available evidence for the effectiveness of hydrotherapy in the treatment of RA patients. Our findings suggest that patients who received hydrotherapy treatment for RA gained some beneficial effects in improving their health status (e.g. reduced pain scores) compared with the control groups. Further additional benefits included a substantial increase in physical activity and emotional well-being in patients in the aquatic programmes compared with control groups in the short term. However, the long-term benefits were found to be inconclusive. There is no cure for RA, and it is therefore important to look into both disease prevention and non-pharmacological treatment that reduces the burden to patients and carers. A treatment for RA which reduces or slows down the inflammatory process would therefore be of great benefit, both from the health service perspective and also in terms of the perceived benefit to RA patients in improving their QoL.

The PEDro scores for all of the papers reviewed ranged from 5–8, and were regarded as being of moderate to high quality. The average methodological quality of all the studies was 6.8 and was regarded as moderate. However, all of the studies reviewed suffered from methodological flaws that limited their generalizability to the wider population of RA patients.
The six studies appraised differed in the frequency and duration of the hydrotherapy sessions given to participants: twice weekly over four weeks, once weekly over six weeks, three times weekly over ten weeks, twice weekly for 12 weeks and once weekly for four years (long term study); they also differed in the duration of hydrotherapy. Therefore, we were unable to determine from the present review the ideal number of hydrotherapy sessions that are needed for RA patients to derive clinically significant benefit from this intervention. A possible explanation for this might be that each study was designed with specific targets and goals, and different primary outcome measures. A recent national survey in the UK by Bryant et al. (2009) reported that the median optimal number sessions for the treatment of RA patients was six weeks.

**Methodological critique of the reviewed articles**

The choice of outcome measures used in the reviewed studies should be examined with caution. The HAQ was the most common instrument used to measure physical function. In terms of the efficacy of hydrotherapy, it was used as a primary outcome measure in one study only (Sanford-Smith et al., 1998). Significant improvements in health status (health-related QoL) were found in two studies (Bilberg et al., 2005; Hall et al., 1996) by using two different health-related QoL scales of measurement. This means that no standardized, specific scale, which was superior to another, was used when measuring health status or QoL in RA patients. Grip strength and joint tenderness were the most common disease activity indices, which were found to be statistically significant in hydrotherapy trials in comparison with other disease activity indices in patients with RA (Bilberg et al., 2005; Hall et al., 1996; Stenstrom et al., 1991). These findings should be interpreted with caution because few studies have investigated the disease activity domains in RA patients. The contradictory results of grip strength measures can be explained by the different types of assessment tools employed in the various studies. Hall et al. (1996) measured the grip strength of the dominant hand by using a digital grip strength monitor inflated to 20 mmHg (Lee et al., 1974; Rhind et al., 1980). The mean of three readings was recorded, whereas Bilberg et al. (2005) measured grip strength by using an electronic instrument (Grippit, AB Detektor, Göteborg, Sweden), recording the maximum and mean strength and the best performance of three (Nordenskiöld, 1990; Nordenskiöld and Grimby, 1993, 1997). Conversely, Stenstrom et al. (1991) measured grip strength manually by using a Sphygmomanometer cuff rolled up two turns and inflated to 20 mmHg (Lansbury, 1958). Sanford-Smith and colleagues (Sanford-Smith et al., 1998) did not report the method of assessment used to measure the grip strength. Therefore, future studies should consider using appropriate standardized procedures in measuring grip strength in patients with RA with malfunction of dexterity and pain.

The reduced joint tenderness observed in the hydrotherapy group of Hall et al. (1996) might be attributed to the reduction in joint loading supported by buoyancy. Furthermore, the hydrostatic pressure of water immersion is considered to be effective in reducing oedema (Poyhonen et al., 2000).

However, we noted many substantial methodological shortcomings in the research we reviewed, mainly in the inadequate reporting of interventions in terms of their setting, water temperature, depth of pool, and the type and intensity of the exercise programme. In addition, there were other methodological flaws relating to RCT design, such as inappropriate randomization, concealment of allocation to groups and the blinding procedure to the outcome measurements.

Overall, many of the studies involved in the present review had a relatively small sample size and lacked adequate statistical power to examine the effectiveness of hydrotherapy in the treatment of patients with RA. In addition, the studies reviewed used different primary outcome measures and a few studies had inadequate and variable follow-up periods.

The present review had several limitations. First, the review focused only on studies published in English; it is possible that potentially relevant articles published in other languages may have been missed. Such studies were excluded because of the limited resources available for the present review. Second, the searches were limited to published articles. Third, some of the studies did not give detailed information about their data analysis. This will have affected the conclusions drawn from these studies, so caution is required in the interpretation of their findings. Fourth, the variation in the dosages of the intervention in the six studies analysed makes it difficult to provide clear guidance in this area. Fifth, we did not investigate the cost-effectiveness of hydrotherapy. Unfortunately, none of the studies...
reviewed reported the cost-effectiveness of their intervention. Costs versus benefits assessments will become increasingly important in medical rehabilitation and physiotherapy research, as RA patients are more likely to continue to use healthcare services for a long period because of the chronic nature of the condition. Therefore, future studies should consider the cost-effectiveness of a hydrotherapy intervention. Finally, the present review focused on RCTs. It is therefore imperative that future studies assess the value of grey literature and case-controlled studies to evaluate the benefit of hydrotherapy for this patient group.

Implications for practice

The results of the present review indicate the beneficial effects of hydrotherapy compared with no intervention, or with other interventions. An important practical implication is that the outcome measures used to assess pain, physical function, disease activity and QoL scales are appropriate for the assessment of patients with RA. In addition, some of the studies reviewed showed hydrotherapy to be associated with improvements, particularly in regard to pain, disease activity (grip strength, joint tenderness) and health status (mood and tension). The evidence from this review might give further option for rheumatologists to refer appropriate RA patients for hydrotherapy treatment as part of their medical rehabilitation.

Implications for research

Few RCTs have examined the effects of a hydrotherapy intervention on RA. The present review indicates that there is no consistency in the literature in terms of the type of exercise and the dose (intensity, frequency and duration) used in hydrotherapy treatment for patients with RA. In addition, future studies should consider examining the cost-effectiveness of hydrotherapy and the optimal use of aquatic exercise for patients with RA. Considerably more work is needed to determine the effectiveness of hydrotherapy on disease activity, psychological aspects of RA (anxiety and depression) and physical function using appropriate outcome measures. Large, high-quality RCTs are needed which could provide more definitive evidence for the efficacy of hydrotherapy using rigorous methodology (e.g. an adequate sample size). In addition, case-controlled studies should be considered.

Conclusions

There is some evidence to suggest that hydrotherapy has a positive role in reducing pain and improving the health status of patients with RA in the short term. However, the long-term benefit is unknown. It is difficult to make specific recommendations at this stage because of lack of evidence (e.g. optimal duration and frequency) for clinical practice. Therefore, further studies are needed, using robust RCTs.

REFERENCES


