Physical Therapy
Journal of the American Physical Therapy Association and de Fysiotherapeut
Royal Dutch Society for Physical Therapy

PTJ Online First

Effect of Therapeutic Aquatic Exercise on Symptoms and Function Associated With Lower Limb Osteoarthritis: A Systematic Review With Meta-Analysis

Benjamin Waller, Anna Ogonowska-Slodownik, Manuel Vitor, Johan Lambeck, Daniel Daly, Urho M. Kujala and Ari Heinonen

PHYS THER. Published online June 5, 2014
Originally published online June 5, 2014
doi: 10.2522/ptj.20130417

The online version of this article, along with updated information and services, can be found online at: http://ptjournal.apta.org/content/early/2014/08/20/ptj.20130417

Collections	This article, along with others on similar topics, appears in the following collection(s): Evidence-Based Practice Injuries and Conditions: Lower Extremity Osteoarthritis Osteoarthritis Systematic Reviews/Meta-analyses Therapeutic Exercise
E-mail alerts	Sign up here to receive free e-mail alerts

Online First articles are published online before they appear in a regular issue of *Physical Therapy* (**PTJ**). **PTJ** publishes 2 types of Online First articles:

Author manuscripts: PDF versions of manuscripts that have been peer-reviewed and accepted for publication but have not yet been copyedited or typeset. This allows **PTJ** readers almost immediate access to accepted papers.

Page proofs: edited and typeset versions of articles that incorporate any author corrections and replace the original author manuscript.

Effect of Therapeutic Aquatic Exercise on Symptoms and Function Associated With Lower Limb Osteoarthritis: A Systematic Review With Meta-Analysis

Benjamin Waller, Anna Ogonowska-Slodownik, Manuel Vitor, Johan Lambeck, Daniel Daly, Urho M. Kujala, Ari Heinonen

Background. Current management of osteoarthritis (OA) focuses on pain control and maintaining physical function through pharmacological, nonpharmacological, and surgical treatments. Exercise, including therapeutic aquatic exercise (TAE), is considered one of the most important management options. Nevertheless, there is no up-to-date systematic review describing the effect of TAE on symptoms and function associated with lower limb OA.

Purpose. The purpose of this study was to conduct a systematic review with meta-analysis to determine the effect of TAE on symptoms and function associated with lower limb OA.

Data Sources. The data sources used in this study were: MEDLINE, PubMed, EMBASE, CINAHL, PEDro, and SPORTDiscus.

Study Selection. All studies selected for review were randomized controlled trials with an aquatic exercise group and a nontreatment control group. In total, 11 studies fulfilled the inclusion criteria and were included in the synthesis and meta-analysis.

Data Extraction. Data were extracted and checked for accuracy by 3 independent reviewers.

Data Synthesis. Standardized mean difference (SMD) with 95% confidence interval (95% CI) was calculated for all outcomes. The meta-analysis showed a significant TAE effect on pain (SMD=0.26 [95% CI=0.11, 0.41]), self-reported function (SMD=0.30 [95% CI=0.18, 0.43]), and physical functioning (SMD=0.22 [95% CI=0.07, 0.38]). Additionally, a significant effect was seen on stiffness (SMD=0.20 [95% CI=0.03, 0.36]) and quality of life (SMD=0.24 [95% CI=0.04, 0.45]).

Limitations. Heterogeneity of outcome measures and small sample sizes for many of the included trials imply that conclusions based on these results should be made with caution.

Conclusions. The results indicate that TAE is effective in managing symptoms associated with lower limb OA.

B. Waller, PT, MSc, Department of Health Sciences, University of Jyvaskyla, PO Box 35, FI-40014, Jyvaskyla, Finland. Address all correspondence to Mr Waller at: ben.waller@jyu.fi.

A. Ogonowska-Slodownik, MSc, Faculty of Rehabilitation, Jozef Pilsudski University of Physical Education, Warsaw, Poland.

M. Vitor, MSc, Faculty of Sport, University of Porto, Porto, Portugal.

J. Lambeck, PT, BHS, KU Leuven, Faculty of Kinesiology and Rehabilitation Sciences, Leuven, Belgium.

D. Daly, PhD, KU Leuven, Faculty of Kinesiology and Rehabilitation Sciences.

U.M. Kujala, MD, PhD, Department of Health Sciences, University of Jyvaskyla.

A. Heinonen, PhD, Department of Health Sciences, University of Jyvaskyla.

[Waller B, Ogonowska-Slodownik A, Vitor M, et al. Effect of therapeutic aquatic exercise on symptoms and function associated with lower limb osteoarthritis: a systematic review with meta-analysis. *Phys Ther*. 2014;94:xxx–xxx.]

Published Ahead of Print: June 5, 2014
Accepted: May 18, 2014
Submitted: September 13, 2013

Osteoarthritis (OA) is the most common form of arthritis,[1] with the knee, hand, hip, and spine being the most common symptomatic body parts affected.[2] People with OA of the lower limb are affected by symptoms in all levels of the *International Classification of Functioning, Disability and Health* (ICF).[3] There is currently no known cure for OA, and management focuses on nonpharmacological (eg, exercise, education, physical therapy, weight loss), pharmacological (eg, nonsteroidal anti-inflammatory drugs, acetaminophen, glucosamine, chondroitin), and surgical (eg, arthroplasty, osteotomy) treatments.[4–6]

Exercise in the form of land-based strengthening and aerobic conditioning or therapeutic aquatic exercise (TAE) and weight loss are considered the most central elements in the current nonpharmacological recommendations in the management of OA.[4,6] There is high-quality evidence supporting the prescription of land-based exercise for people with lower limb OA. Land-based strength and aerobic training has been shown to have a small to moderate effect size (ES) in pain (range=0.32–0.52) and functioning (range=0.32–0.46),[7–10] with no clear superiority between the 2 exercise regimens. Therapeutic aquatic exercise is recommended to people with lower limb OA because of the reduced loading on the joint as a result of buoyancy.[11] Although unproven,[12] the reduced loading is thought to protect the joints from further damage and allows more efficacious training for people who are unable to train effectively on land. However, like land-based training,[13] there is no consensus on what type of TAE is most effective in the management of lower limb OA.

To date, 2 systematic reviews have been published investigating the effect of TAE on OA: a Cochrane review[14] and a systematic review with meta-analysis.[15] The Cochrane review was limited to studies published up to May 2006 and included 6 studies, 5 of which had a control-comparison group. This review[14] indicated that TAE has a small effect on pain (ES=0.19 [95% CI=0.04, 0.35]), function (ES=0.26 [95% CI=0.11, 0.42]), walking ability (ES=0.18 [95% CI=−0.03, 0.39]), and quality of life (ES=0.32 [95% CI=0.03, 0.61]) compared with controls. The systematic review by Batterham et al[15] demonstrated that TAE and land-based exercise have similar effects on self-reported functioning and mobility. The authors concluded that neither approach appeared to be superior to the other. This review was based on studies published up to July 2010, including participants with OA or rheumatoid arthritis, or both.

Nevertheless, in recent years, there has been an increased number of randomized controlled trials (RCTs) investigating the effect of TAE on people with lower limb OA not yet integrated into a systematic review or meta-analysis. Therefore, the aim of this review was to investigate, through a systematic review and meta-analysis of RCTs, the effect of therapeutic aquatic exercise on symptoms and function in people with lower limb OA.

Method

Search Strategy and Study Selection

For this systematic review and meta-analysis, we performed a broad search of 6 databases (MEDLINE, PubMed, CINAHL, SPORTDiscus, PEDro, and EMBASE) using a comprehensive combination of key words: "hydrotherapy" or "water exercise" or "aquatic exercise" or "aquatic therapy" or "water rehabilitation" or "aquatic physical therapy" or "aquatic rehabilitation" or "aquatics" AND "osteoarthritis" or "OA" or "arthritis." The search included publications, in English, appearing before December 1, 2013 (eAppendix 1, available at ptjournal.apta.org). Additionally, a hand search of references was performed. Inclusion was based on assessment by 2 independent reviewers (B.W. and J.L.), and full agreement was required. Based on titles and abstracts, duplicates and nonaquatic exercise studies were excluded. Following this search, full-text manuscripts for the remaining studies were retrieved and read by each reviewer, and final selection was made. If needed, disagreements were resolved through discussion with and assistance from a third reviewer (D.D).

Studies included in our review had to have an RCT design, be published in English, and fulfill the following criteria according to the PICOS (Population, Intervention, Comparison, Outcome, and Study) system.[16] The study population included people with clinically diagnosed OA (as assessed with radiography[17,18] or according to American College of Rheumatology guidelines[19]) in one or more joints of the lower limb, with no age or sex restrictions. We included all interventions that could be classified as therapeutic aquatic exercise where there was full immersion of the body. No limitation was placed on the type of exercise (aerobic, range of motion [ROM], strength) or outcome measures used. Studies were excluded if the comparison group participated in an exercise intervention (land or water based) with or without an additional intervention (eg, home exercises, education). Furthermore, studies with a PEDro score of ≤5, indicating low methodological quality and a high risk of bias,[20] were excluded. The studies had to have a control-comparison group who continued usual care or participated in a sham intervention. Outcome data had to be reported for at least one outcome at baseline and postintervention.

Quality Assessment

Methodological quality or risk of bias was assessed using the 11-point PEDro scale,[21] which has been shown to be a reliable[20] and valid[22] assessment tool. The PEDro scale is based on the 9-point Delphi scale developed by Verhagen et al[23] and is used specifically with RCTs in physical therapy. The 11 quality assessment criteria are: eligibility criteria, random allocation, allocation concealment, baseline similarity, participant blinding, therapist blinding, assessor blinding, adequate follow-up, intention-to-treat analysis, between-group comparisons, and point and variability measures given. In the PEDro scale, the first criterion (eligibility) is not included in the final score, which ranges from 0 to 10. Each criterion is scored 1 (“yes”) or 0 (“no, don’t know/unclear”). Generally, the maximum a TAE study can be scored is 8 because of the difficulties in blinding the participants and therapist from the intervention in exercise studies.[24] A study with a score of ≥7 is considered to have high methodological quality, and studies scoring ≤5 have considered to have low methodological quality.[20] Assessment of methodological quality was performed independently by 2 reviewers (B.W. and A.O-S.) and compared. In case of disagreement, consensus was obtained by consulting a third reviewer (J.L.).

Data Extraction

Intervention description, inclusion and exclusion criteria, baseline data, and values for all outcomes at baseline, postintervention, and longer follow-up (3–6 months) were extracted by 2 reviewers (A.O-S. and M.V.) and checked for accuracy by a third reviewer (B.W.). Where possible, intention-to-treat data were extracted for follow-up measurements; otherwise, per protocol data were extracted. When data were not presented in the study as mean and standard deviation or were presented in a form that prevented calculation of mean and standard deviation, the original authors were contacted, and original data were requested. Standard deviation was calculated from 95% CI values for the study by Hale et al[25] and from standard error values for the study by Lund et al.[26] Mean and standard deviation were estimated from median values and interquartile ranges for all outcomes from the study by Foley et al.[27] In these cases, the median value was taken as best estimated mean, and interquartile ranges were divided by 1.35.[15,28] The original authors provided aggregate data for lower limb muscle strength measurements.[26] Postintervention scores for the Six-Minute Walk Test and isometric strength from the study by Foley et al[27] were requested, but no reply was received. Outcomes were divided into 5 groups: pain, stiffness, self-reported functioning, physical performance measures, and quality of life. In all cases, the ES between TAE and control groups was calculated as the standardized mean difference (SMD). Data were corrected so that effects in favor of TAE are described as positive ES values. In this study, an ES (SMD) of 0.2 to 0.5 was considered as small, 0.5 to 0.8 as medium, and ≥0.8 as a large effect.[29] For all analyses, we used an inverse-variance weighted random-effects model that incorporates heterogeneity into the model, and ES is presented as SMD (95% CI).

Results

In total, 1,234 potential studies were found; no additional studies where found by hand searching of references. Based on title and abstract content, 1,197 of these studies were excluded. The full texts of the remaining 37 studies were read, and a further 26 studies were excluded, resulting in 11 studies being retained in the qualitative and quantitative synthesis of this review (Fig. 1).

Methodological Quality and Risk of Bias of Included Studies

Methodological quality of the included studies is shown in Table 1, and scoring for each criterion is presented in eAppendix 2 (available at ptjournal.apta.org). Five studies achieved PEDro scores of 8/10.[25–27,30,31]

Participants

In total, data were extracted for 1,092 participants. Mean age ranged from 62 to 76 years, with an average body mass index range of 26.6 to 32.9 kg/m^2. Women comprised approximately 73% of the participants. Six studies included knee and hip OA,[25,27,30–33] 3 studies included only knee OA,[26,34,35] 1 study included hip OA only,[36] and 1 study included any lower limb OA.[37] The study by Patrick et al[37] has previously been classified as a knee/hip OA study, but on consultation with the authors, it was reclassified as any lower limb OA. A summary of the populations included is shown in Table 1.

Therapeutic Aquatic Exercise Interventions

Intervention duration, frequency, dose, intensity, exercise selection, and adherence were different among studies. The interventions, intensity, and dose for each study are described in Table 2. Planned exercise dose varied from 100 minutes per week to 180 minutes per week for 6 weeks to 52 weeks (in total, from 9 to 107 hours). Frequency of treatment was either 2 or 3 times a week (2 times a week was most popular, n=7). Additionally, for all but 4 studies,[30,31,33,35] it would have been impossible to accurately reproduce the intervention based on information given in or referenced in the text (Tab. 2).

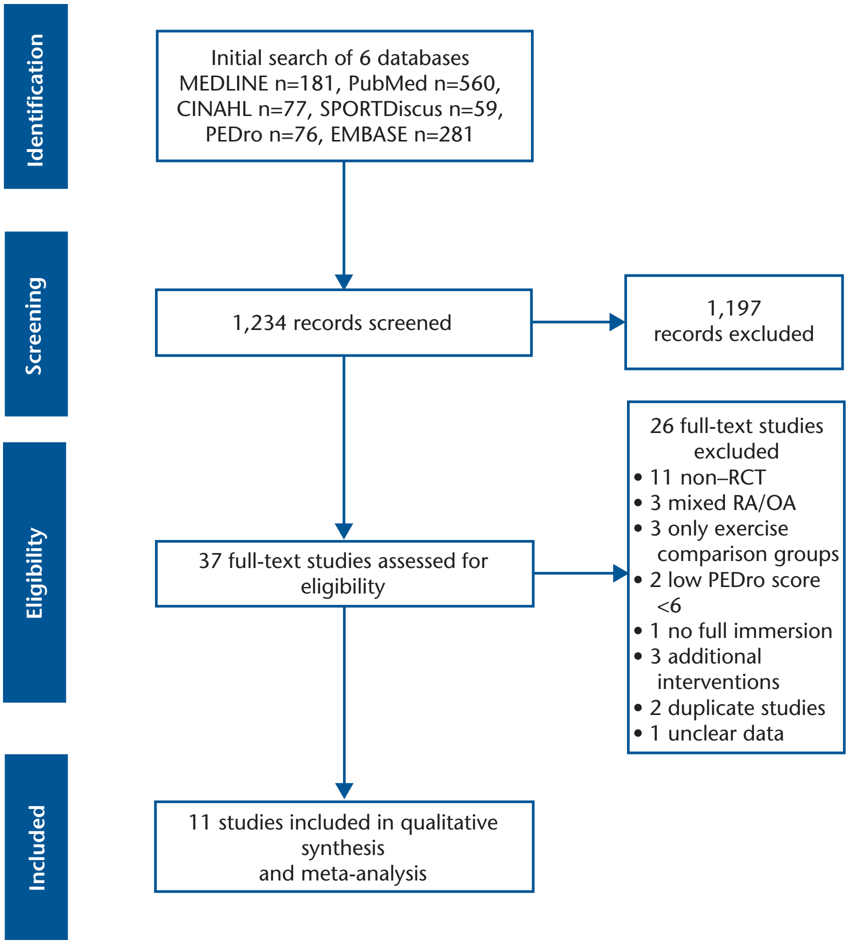


Figure 1.
Flow diagram showing screening process and search results. RCT=randomized controlled trial, RA=rheumatoid arthritis, OA=osteoarthritis.

Outcome Measurements

All outcomes with appropriately reported data were extracted and included in the qualitative and quantitative synthesis. Outcome measures were grouped according to their construct and design (Tab. 2). In cases where more than one outcome was used to measure a single construct in a single study, outcome selection was based on a predescribed hierarchy with the highest ranked outcome measure being included. Suitable recommendations were found to base selection for the constructs of pain[38] and self-reported functioning and quality of life[39] (eAppendix 3, available at ptjournal.apta.org). The mental component summary of the Medical Outcome Study 36-Item Short-Form Health Survey (SF-36) and the 12-Item Short-Form Health Survey (SF-12) was ranked higher than other scales for quality of life[39]; for the study by Cochrane et al,[32] we selected the data for SF-36 mental health over SF-36 role mental health. No selection for stiffness was required. Due to the wide variety of constructs covered by the physical functioning tests, we decided to first divide the constructs into activities, muscle strength, and joint ROM. Selection of outcome measure for activities, when possible, was based on the suggestions of Dobson et al.[40] In cases of disagreement, we selected the outcome that best covered different constructs related to activity (Timed "Up & Go" Test/stairs selected before walking ability). When isokinetic strength was measured (in newton-meters) using different angular velocities, the results for 60°/s were used. Unless data for both affected and unaffected sides were reported, measurements for the right side only were included here. Table 2 shows the full list of outcomes used in each study; outcomes in bold type indicate those used in the quantitative synthesis.[41]

Overall Effect of TAE on Lower Limb OA

Directly after intervention, TAE had a small but significant effect on pain (SMD=0.26 [95% CI=0.11, 0.41]) and stiffness (SMD=0.20 [95% CI=0.03, 0.36]) (Fig. 2). The effect of TAE on both self-reported and objectively measured physical functioning also was small but significant (SMD=0.30 [95% CI=0.18, 0.43] and SMD=0.22 [95% CI=0.07, 0.38], respectively). Therapeutic aquatic exercise had a small but significant effect on physical functioning at activity level (SMD=0.22 [95% CI=0.01, 0.42]) and ROM (SMD=0.56 [95% CI=0.14, 0.99]) and no significant effect on muscle strength (Fig. 3). Therapeutic aquatic exercise had a small but significant effect on quality of life (SMD=0.24 [95% CI=0.04, 0.45]).

Effect of TAE at 3- and 6-Month Follow-ups

One study carried out follow-up measurements at 3 months,[26] and one study performed a follow-up at 6 months after cessation of intervention.[32] Thus, pooling was not possible due to lack of data. In these studies, the effect of TAE had been lost at both follow-up measurement points.

Table 1.
Description of Included Studies, Population, Adverse Effects, and Dropouts[a]

Study	Location of OA	PEDro Score (/10)	Sample Size and Comparison	Age (y)	Male/Female	Participant Recruitment	Pain	Self-reported Function	Adverse Effects	Dropouts (%)[b]
Patrick et al,[37] 2001	Lower limb	6	TAE group (n=125) Control group (n=124)	65.7 66.1	18/107 16/108	Advertisement from local area	1.53 (0.60) 1.44 (0.61)	5.74 (1.62) 5.20 (1.73)	None reported	21 (17) 3 (2)
Foley et al,[27] 2003	Hip/knee (51/126)	8	TAE group (n=35) Control group (n=35)	73.0 (8.2) 69.8 (9.0)	20/15 15/20	Advertisement from local health services	10.0 (3.0)[c] 10.0 (4.0)	34.0 (16.0)[c] 37.0 (17.0)	2 reported increased pain	1 (3) 3 (9)
Cochrane et al,[32] 2005	Hip/knee (mixed)	7	TAE group (n=153) Control group (n=159)	69.9 (6.8) 69.6 (6.3)	56/97 60/99	Advertisement in health services	8.72 (3.62) 9.10 (3.14)	30.1 (13.1) 31.1 (11.2)	None reported	48 (31) 33 (21)
Wang et al,[33] 2007	Hip/knee	6	TAE group (n=21) Control group (n=21)	69.3 (13.3) 62.7 (10.7)	4/16 2/16	Advertisement in community sources	52.2 (23.8) 55.3 (24.6)	0.90 (0.4) 0.95 (0.5)	None reported	1 (5) 3 (14)
Fransen et al,[30] 2007	Hip/knee (mixed)	8	TAE group (n=55) Control group (n=41)	70.0 (6.3) 69.6 (6.1)	15/40 7/34	Advertisement in newspapers and community and physician referral	38.2 (17.4) 44.4 (17.0)	46.3 (20.4) 50.8 (19.3)	None reported	3 (5) 0 (0)
Hinman et al,[31] 2007	Hip/knee (16/55)	8	TAE group (n=36) Control group (n=35)	63.3 (9.5) 61.5 (7.8)	12/24 11/24	Advertisement	6 (2) 5 (2)	757 (327) 630 (315)	Only minor, no dropouts	1 (3) 4 (11)
Lund et al,[26] 2008	Knee	8	TAE group (n=27) Control group (n=27)	65 (12.6) 70 (9.9)	5/22 9/18	Recruited from outpatient and GPs and advertisement	59.8 (18.4) 48.5 (31.9)	44.7 (18.1) 39.6 (13.2)	None affecting participation	1 (4) 2 (7)
Lim et al,[34] 2010	Knee	7	TAE group (n=26) Control group (n=24)	65.7 (8.9) 63.3 (5.3)	3/23 3/21	Recruited from patients registered at hospital	4.41 (1.44) 4.12 (2.08)	35.1 (11.3) 30.40 (19.1)	None in TAE group	2 (8) 4 (17)
Arnold and Faulkner,[36] 2010	Hip	6	TAE group (n=28) Control group (n=25)	74.4 (7.5) 75.8 (6.2)	6/20 9/16	Advertisements and posters in clinics, recreational facilities	Not reported	70.4 (21.9) 65.3 (18.1)	1 fall, minor increased pain	5 (18) 6 (22)
Wang et al,[35] 2011	Knee	7	TAE group (n=28) Control group (n=28)	66.7 (5.6) 67.9 (5.9)	4/22 4/22	Advertisements in sports and community centers	61 (20) 66 (18)	73 (20) 70 (19)	1 reported dizziness	2 (7) 2 (7)
Hale et al,[25] 2012	Hip/knee	8	TAE group (n=23) Control group (n=16)	73.6 (1.5) 75.7 (1.1)	6/17 4/12	Volunteers recruited by advertisement	7.2 (5.81–8.62) 7.5 (6.67–8.39)	24.7 (21.0–28.5)[d] 27.8 (24.7–31.0)	1 in TAE group, increased leg pain	3 (13) 1 (6)

[a] OA=osteoarthritis, TAE=therapeutic aquatic exercise, GPs=general practitioners. Values presented as mean (SD) unless otherwise stated.
[b] Directly after intervention.
[c] Median and interquartile range.
[d] Mean and 95% confidence interval.

Effect of TAE on Specific Joint OA

Three studies[26,34,35] investigated the effect of TAE on knee OA, with no significant effect demonstrated (eAppendix 4, available at ptjournal.apta.org). One study included only hip OA[36]; therefore, no comparisions can be made. Six studies[25,27,30–33] included participants with either hip or knee joint OA, with effect of TAE reported in eAppendix 5 (available at ptjournal.apta.org).

Adverse Effects

Adverse effects were documented in all of the studies, with 5 studies reporting some form of adverse effect as a direct result of participation in the TAE intervention, including increase in pain.[25,27] In total, only 4 participants from 3 studies,[25,27,36] dropped out as a direct result of the adverse effect caused by the TAE program (Tab. 2).

Discussion

Our meta-analysis indicates that TAE is an effective treatment option for the management of symptoms and functional deficits as a result of lower limb OA compared with no treatment. The previous meta-analysis[14] contained 5 studies with a control-comparison group and included 661 participants, whereas our meta-analysis contains 11 high-quality studies with 1,092 participants, thus supporting the need for

Table 2.
Description of Interventions and Outcome Measures Used in Included Studies[a]

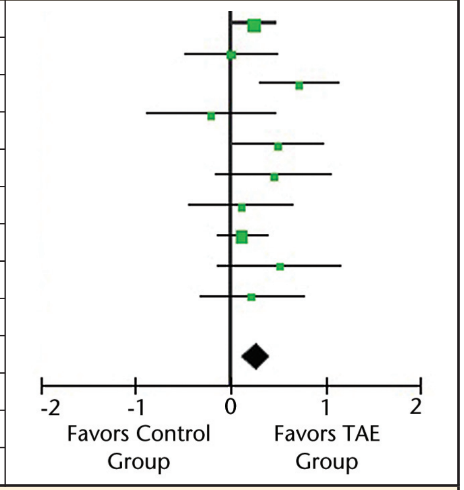
Study	Aquatic Intervention	Intensity	Pool (Depth/ Temperature)	Dose per Protocol[b]	Adherence	Outcome Measures			
						Pain and Stiffness	Self-reported Function	Physical Performance Measures	QoL
Patrick et al,[37] 2001	AFAP, 45–60 min of upper and lower limb flexibility, ROM, and strength exercises	Gentle exercises for flexibility; no other intensity stated	85°–92°F	20/40/52/2,080		**HAQ pain (0–3)**	**HAQ disability (0–3)**		**Perceived QoL scale (0–10)**, current health desirability rating CES-D (0–10), Quality of Well-being Scale
Foley et al,[27] 2003	Warm-up, lower limb strength exercises, cool-down	3 sets × 10–15 reps, progression with addition of gaiters		6/18/30/540	84%	**WOMAC pain (0–20) and stiffness (0–8)**	**WOMAC function (0–68)**, SF-12 PCS	6MWT, isometric knee extension strength	**SF-12 MCS (25–70)**, arthritis self-efficacy (symptoms, satisfaction)
Cochrane et al,[32] 2005	1 h of strength, ROM, endurance, stretches, balance, and coordination exercises	Increase speed and resistance, progressed every 6–8 wk	3 different pools	52/104/60/6,240		**WOMAC pain (0–20) and stiffness (0–8)**, SF-36 pain	**WOMAC function (0–68)**, SF-36 social functioning, physical function	8-ft walk test, **isometric knee extension (N)**/flexion strength, **stair climb ascending (s)**/descending	**SF-36 mental health (0–100)**, role mental health, vitality, general health, EuroQol-VAS
Wang et al,[33] 2007	AFAP flexibility, strength, and aerobic training	10–15 reps for strength, Borg (CR-10) 3–4 for aerobic training	30°–32°C	12/36/60/1,800	81.7%	**VAS pain (0–100)**	**HAQ (14 question) (0–3)**	**6MWT (m), isometric knee extension (kg)**/flexion strength, **ROM knee (degrees)** and hip	
Fransen et al,[30] 2007	Strength, lower limb, mobility, trunk work, and balance	10–20 reps		12/24/60/1,440		**WOMAC pain (0–100)**	**WOMAC function (0–100)**, SF 36 PCS	**TUG (s)**, 50-ft walk test, stair climb ascending	**SF-36 MCS (0–100)**, DASS-21
Hinman et al,[31] 2007	Warm-up, leg strength exercises, 6- to 10-min walking, cool-down	2–5 sets × 10 reps, individual progression	34°C	6/12/52/624	99%	**VAS pain (0–10)**, WOMAC pain	**WOMAC function (0–1,700)**	**TUG (s), isometric knee extension strength (kg)**, 6MWT, step test	**Assessment of Quality of Life Scale (range=−4 to 100)**
Lund et al,[26] 2008	10-min warm-up, 20-min strength and endurance exercises, 20-min balance, stretching, and cool-down	3.5 min per exercise, 30-s stretch per muscle group	33.5°C	8/16/50/800	92%	**VAS pain (0–100)** and rest, **KOOS symptoms (0–100)** and pain	**KOOS ADL (0–100)** and sports recreation	**Isokinetic knee extension (60°/s, N·m)**/flexion	**KOOS QoL (0–100)**
Lim et al,[34] 2010	5-min warm-up, 30-min strength and endurance exercises, 5-min cool-down	65% maximum heart rate	1.15 m/34°C	8/24/40/960	>67%	**Bodily pain mean (VAS 0–10)**/pain interference	**WOMAC global (0–96)**, SF-36 PCS	**Isokinetic knee extension (N·m)**/flexion	**SF-36 MCS (0–100)**
Arnold and Faulkner,[36] 2010	Warm-up, lower and upper limbs, trunk strengthening exercises, cool-down	Not stated	Variable depth/30°C	11/22/45/990	65%		**Activities and Balance Confidence Scale (0–100)**	Berg Balance Scale, 6MWT, **30-s chair stand**, $TUG_{cognitive}$	
Wang et al,[35] 2011	AFAP flexibility, and aerobic training	10–15 reps for strength, Borg (CR-10) 3–4 for aerobic training	30°C	12/36/60/2,160	86.4%	**KOOS pain (0–100)**	**KOOS ADL and sports recreation (0–100)**	**6MWT, ROM knee extension** and flexion	**KOOS QoL (0–100)**
Hale et al,[25] 2012	Warm-up (walking, stretching), balance exercises, cool-down (walking, stretching)	Weeks 1–3: 1 min per exercise Weeks 4–6: 1.5 min per exercise Weeks 7–12: 2 min per exercise	0.94–1.3 m/28°C	12/24/50/1,200	20%–100%	**WOMAC pain (0–20) and stiffness (0–8)**	**WOMAC function (0–68)**, AIMS-2, Activities and Balance Confidence Scale	**TUG (s), isometric knee extension strength (kg)**, step test, fall risk ratio	

[a] Bold type indicates measures included in meta-analysis, with scale range shown in parentheses. QoL=quality of life, ROM=range of motion, HAQ=Health Assessment Questionnaire, WOMAC=Western Ontario and McMaster Universities Osteoarthritis Index, KOOS=Knee Injury and Osteoarthritis Outcome Score, ADL=activities of daily living, SF-12=12-Item Short-Form Health Survey, SF-36=36-Item Short-Form Health Survey, PCS=physical component summary, MCS=mental component summary, DASS-21=Depression Anxiety Stress Scale, CES-D=Center for Epidemiological Studies Depression Scale, AIMS-2=Arthritis Impact Measurement Scales-2, AFAP= Arthritis Foundation Aquatic Program, CR-10=category ratio 1–10, 6MWT=Six-Minute Walk Test, TUG=Timed "Up & Go" Test, reps=repetitions.
[b] Duration (weeks)/number of sessions/session duration (minutes)/total dose per protocol (minutes).

Pain

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	X̄	SD	Total	X̄	SD	Total		
Cochrane et al,[32] 2005	−8.46	3.74	153	−9.35	3.54	158	25.4	0.24 (0.02, 0.47)
Foley et al,[27] 2003	−10	2.963	34	−10	2.963	32	8.1	0.00 (−0.48, 0.48)
Fransen et al,[30] 2007	−27.3	18.7	55	−40	16.2	41	10.3	0.71 (0.30, 1.13)
Hale et al,[25] 2012	−7.8	3.925	20	−7.1	1.917	15	4.4	−0.21 (−0.88, 0.46)
Hinman et al,[31] 2007	−4	2	36	−5	2	35	8.4	0.49 (0.02, 0.97)
Lim et al,[34] 2010	−3.27	1.67	24	−4.12	2.08	20	5.4	0.45 (−0.15, 1.05)
Lund et al,[26] 2008	−55.8	21	27	−58.1	20	25	6.5	0.11 (−0.43, 0.65)
Patrick et al,[37] 2001	−1.382	0.737	98	−1.462	0.619	117	20.1	0.12 (−0.15, 0.39)
Wang et al,[33] 2007	−43.5	18.6	20	−54.9	25.2	18	4.8	0.51 (−0.14, 1.16)
Wang et al,[35] 2011	72	18	26	68	18	26	6.5	0.22 (−0.33, 0.76)
Total (95% CI)			493			487	100	0.26 (0.11, 0.41)
Heterogeneity: tau^2=0.01, χ^2=10.81, *df*=9 (*P*=.29), I^2=17%								
Test for overall effect: *z*=3.48 (*P*=.0005)								

Standardized Mean Difference IV, Random (95% CI)



Stiffness

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	X̄	SD	Total	X̄	SD	Total		
Cochrane et al,[32] 2005	−3.88	1.67	152	−4.15	1.48	158	53.3	0.17 (−0.05, 0.39)
Foley et al,[27] 2003	−4	2.222	34	−4	2.222	32	11.4	0.00 (−0.48, 0.48)
Hale et al,[25] 2012	−3.7	1.392	20	−3.4	1.502	15	5.9	−0.20 (−0.87, 0.47)
Hinman et al,[31] 2007	−73	45	36	−95	44	35	11.9	0.49 (0.02, 0.96)
Lund et al,[26] 2008	64.6	12	26	61.4	12	25	8.7	0.26 (−0.29, 0.81)
Wang et al,[35] 2011	69	20	26	61	17	26	8.8	0.42 (−0.13, 0.97)
Total (95% CI)			294			291	100	0.20 (0.03, 0.36)
Heterogeneity: tau^2=0.00, χ^2=4.24, *df*=5 (*P*=.52), I^2=0%								
Test for overall effect: *z*=2.37 (*P*=.02)								

Standardized Mean Difference IV, Random (95% CI)

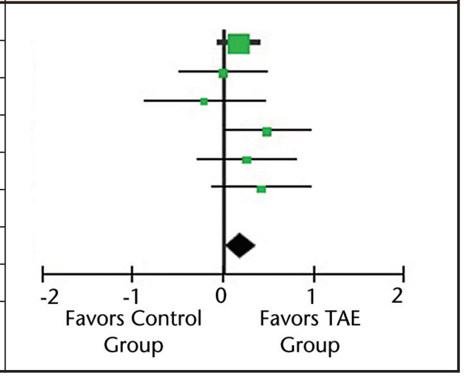


Figure 2.
Forest plots showing the effect of therapeutic aquatic exercise (TAE) on pain and stiffness. 95% CI=95% confidence interval.

Self reported functioning

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	$\overline{X}$	SD	Total	$\overline{X}$	SD	Total		
Arnold and Faulkner,[36] 2010	69.6	24.4	26	62.9	20.8	25	5.0	0.29 (−0.26, 0.84)
Cochrane et al,[32] 2005	−29.26	14.48	149	−32.42	13.25	156	29.9	0.23 (0.00, 0.45)
Foley et al,[27] 2003	−33	12.593	34	−37	9.63	32	6.4	0.35 (−0.14, 0.84)
Fransen et al,[30] 2007	−34.8	23.7	55	−49.9	19	41	8.7	0.69 (0.27, 1.10)
Hale et al,[25] 2012	−24	8.899	20	−24.9	7.015	15	3.4	0.11 (−0.56, 0.78)
Hinman et al,[31] 2007	−598	316	36	−656	373	35	7.0	0.17 (−0.30, 0.63)
Lim et al,[34] 2010	−20.9	9.9	24	−27.6	18.3	21	4.3	0.46 (−0.14, 1.05)
Lund et al,[26] 2008	62.7	12	27	61.1	11	25	5.1	0.14 (−0.41, 0.68)
Patrick et al,[37] 2001	−0.933	0.55	101	−1.127	0.671	121	21.5	0.31 (0.05, 0.58)
Wang et al,[33] 2007	−0.9	0.4	20	−1	0.5	18	3.7	0.22 (−0.42, 0.86)
Wang et al,[35] 2011	76	16	26	69	18	26	5.0	0.40 (−0.14, 0.95)
Total (95% CI)			518			515	100	0.30 (0.18, 0.43)
Heterogeneity: $tau^2=0.00$, $\chi^2=5.21$, $df=10$ ($P=.88$), $I^2=0\%$								
Test for overall effect: $z=4.81$ ($P<.00001$)								

Standardized Mean Difference IV, Random (95% CI)

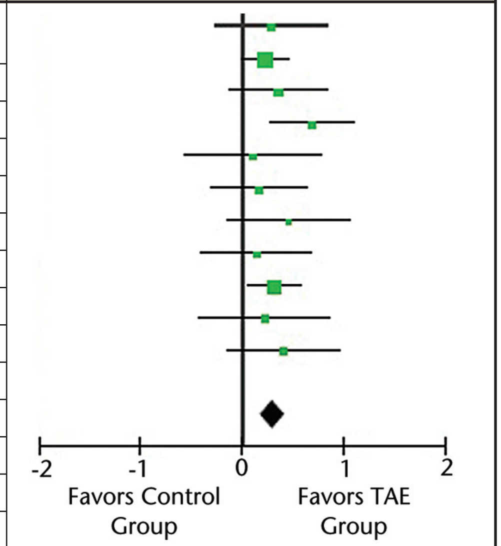


Figure 3.
Forest plots showing the effect of therapeutic aquatic exercise (TAE) on physical functioning. 95% CI=95% confidence interval.

Physical functioning tests

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	X̄	SD	Total	X̄	SD	Total		
Physical functioning tests (activities)								
Arnold et al,[36] 2010	7.5	3.9	26	8.1	2.6	25	5.7	−0.18 (−0.73, 0.37)
Cochrane et al,[32] 2005	−3.24	1.93	153	−3.78	2.72	159	14.1	0.23 (0.01, 0.45)
Fransen et al,[30] 2007	−8.2	1.7	55	−9.2	2.2	41	8.3	0.51 (0.10, 0.93)
Hale et al,[25] 2012	−10.1	2.9	20	−10.7	6.8	15	4.2	0.12 (−0.55, 0.79)
Hinman et al,[31] 2007	−10.32	1.94	36	−10.3	2.78	35	7.1	−0.01 (−0.47, 0.46)
Wang et al,[33] 2007	388.4	80.3	20	390.7	88.6	18	4.6	−0.03 (−0.66, 0.61)
Wang et al,[35] 2011	386	75.8	26	329.1	82.3	26	5.5	0.71 (0.15, 1.27)
Subtotal (95% CI)			336			319	49.6	0.22 (0.01, 0.42)
Heterogeneity: tau^2=0.02, χ^2=8.47, *df*=6 (*P*=.21), I^2=29% Test for overall effect: *z*=2.10 (*P*=.04)								
Physical functioning texts (muscle strength)								
Cochrane et al,[32] 2005	166.02	108.92	152	145.92	99.82	158	14.0	0.19 (−0.03, 0.42)
Hale et al,[25] 2012	18.8	8.465	20	19	12.528	15	4.2	−0.02 (−0.69, 0.65)
Hinman et al,[31] 2007	29.9	12.5	36	24.7	9.5	35	7.0	0.46 (−0.01, 0.93)
Lim et al,[34] 2010	50.4	19.3	24	63.1	24.1	21	5.0	−0.58 (−1.17, 0.02)
Lund et al,[26] 2008	73.22	25.86	26	75.08	33.5	25	5.7	−0.06 (−0.61, 0.49)
Wang et al,[33] 2007	16	4.5	20	12.9	3.8	18	4.3	0.73 (0.07, 1.39)
Subtotal (95% CI)			278			272	40.4	0.13 (−0.17, 0.43)
Heterogeneity: tau^2=0.07, χ^2=11.18, *df*=5 (*P*=.05), I^2=55% Test for overall effect: *z*=0.88 (*P*=.38)								
Physical functioning tests (joint range of motion)								
Wang et al,[33] 2007	−1.2	2.6	20	−3.2	3.5	18	4.4	0.64 (−0.01, 1.29)
Wang et al,[35] 2011	−2.4	2.2	26	−3.3	1.1	26	5.7	0.51 (−0.04, 1.06)
Subtotal (95% CI)			46			44	10.0	0.56 (0.14, 0.99)
Heterogeneity: tau^2=0.00, χ^2=0.09, *df*=1 (*P*=.77), I^2=0% Test for overall effect: *z*=2.62 (*P*=.009)								
Total (95% CI)			660			635	100	0.22 (0.07, 0.38)
Heterogeneity: tau^2=0.03, χ^2=22.72, *df*=14 (*P*=.06), I^2=38% Test for overall effect: *z*=2.79 (*P*=.005) Test for subgroup differences: χ^2=2.78, *df*=2 (*P*=.25), I^2=28.1%								

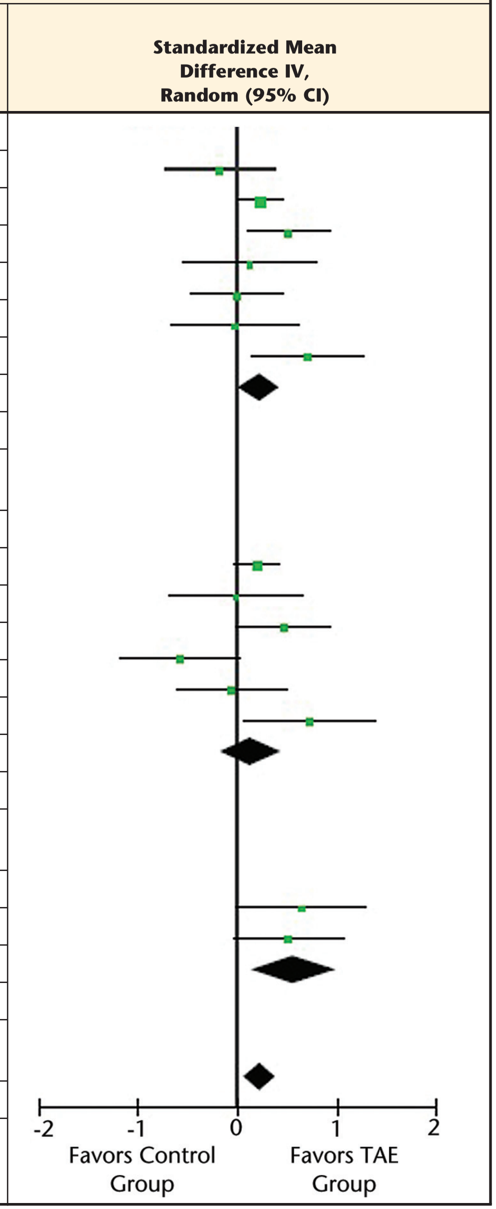


Figure 3.
Continued

Quality of Life

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	$\overline{X}$	SD	Total	$\overline{X}$	SD	Total		
Cochrane et al,[32] 2005	62.55	18.61	150	60.68	17.39	157	20.3	0.10 (−0.12, 0.33)
Foley et al,[27] 2003	53.3	11.481	34	50.5	10.37	32	10.7	0.25 (−0.23, 0.74)
Fransen et al,[30] 2007	54.6	8.9	55	48	11.4	41	12.8	0.65 (0.24, 1.07)
Hinman et al,[31] 2007	0.43	0.2	36	0.5	0.2	35	11.2	−0.35 (−0.82, 0.12)
Lim et al,[34] 2010	54.8	8.8	24	48.4	14.3	21	8.2	0.54 (−0.06, 1.14)
Lund et al,[26] 2008	43	12.47	27	43.1	11.18	25	9.3	−0.01 (−0.55, 0.54)
Patrick et al,[37] 2001	6.396	1.697	101	5.79	1.752	121	18.4	0.35 (0.08, 0.62)
Wang et al,[35] 2011	73	12	26	67	13	26	9.1	0.47 (−0.08, 1.02)
Total (95% CI)			453			458	100	0.24 (0.04, 0.45)
Heterogeneity: $\text{tau}^2=0.04$, $\chi^2=14.28$, $df=7$ ($P=.05$), $I^2=51\%$								
Test for overall effect: $z=2.33$ ($P=.02$)								

Standardized Mean Difference IV, Random (95% CI)

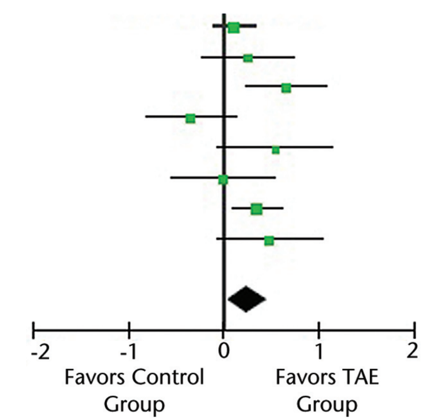


Figure 4.
Forest plots showing the effect of therapeutic aquatic exercise (TAE) on quality of life. 95% CI=95% confidence interval.

this update. The effects of TAE on pain and self-reported function were comparable to those achieved from land-based exercise or the use of acetaminophen and nonsteroidal anti-inflammatory drugs.[4,7–10] This finding, in combination with the high adherence to intervention, low dropout rate, and frequency of severe adverse effects, confirms that TAE should be considered a potentially effective treatment option for people with lower limb OA.

Pain is the most common reason for a person with lower limb OA to initially seek medical assistance, and a common belief is that water is a suitable environment for people who are unable or unwilling to train effectively on land due to pain. Our study demonstrated that TAE can have a small but significant effect on pain, thus strengthening findings from the earlier Cochrane review.[14] Although pain in OA has typically been considered nociceptive, there is growing evidence showing that 28% to 34% of people with hip or knee OA have neuropathic centralized pain.[42,43] The included outcome measures (Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC], Knee Injury and Osteoarthritis Outcome Score [KOOS], and visual analog scale for pain) may not truly capture the pain experienced by people with lower limb OA.[26,31] There is currently a lack of data identifying types of pain and patients who might benefit more from TAE over land-based exercises.

A small but significant ES for both self-reported functioning and physical functioning was demonstrated. The previous reviews[14,15] did not find any effect on physical function, whereas our study is the first to show TAE can have a small but significant effect on physical functioning at activity level. No significant effect of TAE, however, was seen for muscle strength. Care must be taken in interpreting these results, as only left-side and right-side results were reported here, without consideration of affected and unaffected sides. Based on the information provided in the manuscripts, only 1 of the 11 included studies[34] made reference to established exercise prescription guidelines in their intervention planning. Lund et al[26] suggested that the intensity of the aquatic intervention in their study could have been too low to stimulate positive changes. If the interventions were not of sufficient intensity or duration to cause physiological changes at the level of muscle structure and function, it could partially explain why the effects of TAE were lost even at short-term follow-up. One possible advantage of the aquatic environment is that this population may be able to train at higher intensities than on land.[44,45] Nevertheless, due to limited reporting of actual intensities used and limited use of exercise guidelines in the intervention design, it is difficult to confirm the hypothesis. Moreover, Juhl et al[46] indicated that, for best results, the exercise programs for knee OA should have one aim alone, whereas the TAE interventions included a mixture of strength, aerobic, and flexibility exercises. Furthermore, the ES for self-report functioning was similar to the ES of pain and could explain the differences between perceived functional ability and actual functional ability as measured with physical performance tests.

The small effect of TAE on quality of life was of equivalent size to that seen in the previous review investigating the effect of TAE on lower limb OA[14] and is in line with the findings from other studies investigating the effects of exercise on quality of life.[47] The small ES could be explained by the limited improvements in physical functioning found in our review. Small changes in quality-of-life measures have been reported in association with small or nonsignificant changes in physical functioning.[48] Moreover, interpretation of the results has to be done with caution, as the outcome measures used may not accurately represent the true changes in quality of life within this population. Only 2 studies[26,37] used an OA-specific measure (KOOS quality of life). Commonly used generic instruments, such as the SF-36, which was used in 4 studies,[27,30,32,34] tend to be less responsive than OA-specific measures.[49–51]

This review has both strengths and weaknesses. Strengths include selection of studies, all with moderate to high methodological quality of the PEDro score,[21] and in combining studies with small sample sizes, we provided the most accurate effect of TAE. The advantage of using SMD to report ES is that it allows the synthesis of different outcome measures, but the disadvantage is that it is difficult to apply in a clinical situation. Furthermore, the use of minimal clinically important difference (MCID) and responder criteria has been suggested and recommendations made for this population group[51,52]; however, only one study[30] reported the number of participants reaching MCID. Although we controlled for methodological quality, we did not exclude studies that failed to recruit sufficient participants to meet their power calculation.[25,26] Removal of these studies from the overall analysis did not affect overall outcome for the main synthesis but could explain the results for the knee OA-only analysis. Our study did not demonstrate a significant effect of TAE on populations containing individuals with only knee OA or hip OA.[36] It is our view that the lack of significant effect in these analyses was a result of a lack of internal validity (eg, sample size and differences at baseline) in addition to insufficient intervention intensity. Moreover, it is recognized that the symptoms and treatments associated with hip OA are different from those of knee OA. However, we can support the pooling of studies from diverse populations because exercise performed in water generally affects the musculoskeletal system globally and not only locally at one particular joint. All but one outcome measure (KOOS) found in this review can be appropriately used with both hip and knee OA. Our results and other aquatic-based studies in general do not provide a clear indication of which OA patient populations would benefit more from TAE over a land-based intervention, but these data might permit more optimal prescription of what is often a limited and high cost treatment. Additionally, our review was unable to demonstrate the optimal intervention dose, type of exercise, and training intensity for this population group.

In conclusion, this meta-analysis confirmed that TAE is an effective treatment option for people with lower limb OA and should be considered a frontline management option. Researchers planning an aquatic intervention study should ensure that all aspects of the disease, not just pain and self-reported functioning, are taken into consideration. The diverse outcome measures used suggest there is a need for researchers to refer to current recommendations when designing new projects to facilitate more specific between-study comparisons. Future research should identify the patient groups that would benefit the most from TAE and the effects of aquatic exercise on cartilage and examine methods to maintain the training effect and increase physical activity following the treatment period. Nevertheless, investigation of the effect that TAE has on this population in clinical situations is needed using pragmatic study designs and large sample sizes.

Mr Waller, Mr Lambeck, Professor Daly, Professor Kujala, and Professor Heinonen provided concept/idea/research design and writing. Mr Waller, Mrs Ogonowska-Slodownik, Mr Vitor, and Mr Lambeck provided data collection. Mr Waller, Mrs Ogonowska-Slodownik, Professor Daly, Professor Kujala, and Professor Heinonen provided data analysis. Professor Heinonen provided project management, facilities/equipment, and institutional liaisons. Mrs Ogonowska-Slodownik, Professor Kujala, and Professor Heinonen provided consultation (including review of the manuscript before submission).

This research was presented as part of an oral presentation at the World Aquatic Health Conference; October 2013; Indianapolis, Indiana, and as an e-poster at the 2nd World Congress on Controversies, Debates and Consensus in Bone, Muscle and Joint Diseases (BMJD); November 21–24, 2013; Brussels, Belgium.

This work was financed jointly by the Academy of Finland and the Social Insurance Institution of Finland.

DOI: 10.2522/ptj.20130417

References

1 Suri P, Morgenroth DC, Hunter DJ. Epidemiology of osteoarthritis and associated comorbidities. *PM&R*. 2012;4(5 suppl): S10–S19.

2 Lawrence RC, Felson DT, Helmick CG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States, part II. *Arthritis Rheum*. 2008;58:26–35.

3 Dreinhofer K, Stucki G, Ewert T, et al. ICF core sets for osteoarthritis. *J Rehabil Med*. 2004;36(44 suppl):75–80.

4 Zhang W, Nuki G, Moskowitz RW, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part III: changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage*. 2010;18:476–499.

5 Roddy E, Zhang W, Doherty M, et al. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee: the MOVE consensus. *Rheumatology (Oxford)*. 2005;44:67–73.

6 Hochberg MC, Altman RD, April KT, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res*. 2012;64:465–474.

7 Roddy E, Zhang W, Doherty M. Aerobic walking or strengthening exercise for osteoarthritis of the knee? A systematic review. *Ann Rheum Dis*. 2005;64:544–548.

8 Hernandez-Molina G, Reichenbach S, Zhang B, et al. Effect of therapeutic exercise for hip osteoarthritis pain: results of a meta-analysis. *Arthritis Rheum*. 2008;59: 1221–1228.

9 Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev*. 2009;(3):CD007912.

10 Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev*. 2008;(4):CD004376.

11 Harrison RA, Hillman M, Bulstrode S. Loading of the lower limb when walking partially immersed: implications for clinical practice. *Physiotherapy*. 1992;78:164–166.

12 Multanen J, Nieminen MT, Hakkinen A, et al. Effects of high-impact training on bone and articular cartilage: 12-month randomized controlled quantitative MRI study. *J Bone Miner Res*. 2014:29:192–201.

13 Bennell KL, Hinman RS. A review of the clinical evidence for exercise in osteoarthritis of the hip and knee. *J Sci Med Sport*. 2011;14:4–9.

14 Bartels EM, Lund H, Hagen KB, et al. Aquatic exercise for the treatment of knee and hip osteoarthritis. *Cochrane Database Syst Rev*. 2007;(4):CD005523.

15 Batterham SI, Heywood S, Keating JL. Systematic review and meta-analysis comparing land and aquatic exercise for people with hip or knee arthritis on function, mobility and other health outcomes. *BMC Musculoskelet Disord*. 2011;12:123.

16 Rios LP, Ye C, Thabane L. Association between framing of the research question using the PICOT format and reporting quality of randomized controlled trials. *BMC Med Res Methodol*. 2010;10:11.

17 Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16:494–502.

18 Kellgren JH, Lawrence JS. Radiological assessment of rheumatoid arthritis. *Ann Rheum Dis*. 1957;16:485–493.

19 Altman R, Asch E, Bloch D, et al; Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. *Arthritis Rheum*. 1986;29:1039–1049.

20 Maher CG, Sherrington C, Herbert RD, et al. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther*. 2003;83:713–721.

21 Sherrington C, Herbert RD, Maher CG, Moseley AM. PEDro: a database of randomized trials and systematic reviews in physiotherapy. *Man Ther*. 2000;5:223–226.

22 de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother*. 2009;55:129–133.

23 Verhagen AP, de Vet HC, de Bie RA, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol*. 1998;51:1235–1241.

24 van Tulder M, Furlan A, Bombardier C, Bouter L; Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine*. 2003;28:1290–1299.

25 Hale LA, Waters D, Herbison P. A randomized controlled trial to investigate the effects of water-based exercise to improve falls risk and physical function in older adults with lower-extremity osteoarthritis. *Arch Phys Med Rehabil*. 2012;93:27–34.

26 Lund H, Weile U, Christensen R, et al. A randomized controlled trial of aquatic and land-based exercise in patients with knee osteoarthritis. *J Rehabil Med*. 2008;40:137–144.

27 Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis: a randomised controlled trial comparing a gym-based and a hydrotherapy-based strengthening programme. *Ann Rheum Dis*. 2003;62:1162–1167.

28 Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13.

29 Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988:567.

30 Fransen M, Nairn L, Winstanley J, et al. Physical activity for osteoarthritis management: a randomized controlled clinical trial evaluating hydrotherapy or tai chi classes. *Arthritis Rheum*. 2007;57:407–414.

31 Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. *Phys Ther*. 2007;87:32–43.

32 Cochrane T, Davey RC, Matthes Edwards SM. Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis. *Health Technol Assess*. 2005;9:iii–iv, ix–xi, 1–114.

33 Wang TJ, Belza B, Elaine Thompson F, et al. Effects of aquatic exercise on flexibility, strength and aerobic fitness in adults with osteoarthritis of the hip or knee. *J Adv Nurs*. 2007;57:141–152.

34 Lim JY, Tchai E, Jang SN. Effectiveness of aquatic exercise for obese patients with knee osteoarthritis: a randomized controlled trial. *PM&R*. 2010;2:723–731.

35 Wang T, Lee S, Liang S, et al. Comparing the efficacy of aquatic exercises and land-based exercises for patients with knee osteoarthritis. *J Clin Nurs*. 2011;20:2609–2622.

36 Arnold CM, Faulkner RA. The effect of aquatic exercise and education on lowering fall risk in older adults with hip osteoarthritis. *J Aging Phys Act*. 2010;18:245–260.

37 Patrick DL, Ramsey SD, Spencer AC, et al. Economic evaluation of aquatic exercise for persons with osteoarthritis. *Med Care*. 2001;39:413–424.

38 Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol*. 2006;20:721–740.

39 Veenhof C, Bijlsma JW, van den Ende CH, et al. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. *Arthritis Rheum*. 2006;55:480–492.

40 Dobson F, Bennell K, Hinman R, et al. OARSI recommended performance-based tests to assess physical function in osteoarthritis of the hip or knee: authors' reply. *Osteoarthritis Cartilage*. 2013;21:1625–1626.

41 Messier SP, Mihalko SL, Legault C, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. *JAMA*. 2013;310:1263–1273.

42 Hochman JR, Gagliese L, Davis AM, Hawker GA. Neuropathic pain symptoms in a community knee OA cohort. *Osteoarthritis Cartilage*. 2011;19:647–654.

43 Hochman JR, French MR, Bermingham SL, Hawker GA. The nerve of osteoarthritis pain. *Arthritis Care Res (Hoboken)*. 2010;62:1019–1023.

44 Roper JA, Bressel E, Tillman MD. Acute aquatic treadmill exercise improves gait and pain in people with knee osteoarthritis. *Arch Phys Med Rehabil*. 2013;94:419–425.

45 Valtonen A, Poyhonen T, Sipila S, Heinonen A. Effects of aquatic resistance training on mobility limitation and lower-limb impairments after knee replacement. *Arch Phys Med Rehabil*. 2010;91:833–839.

46 Juhl C, Christensen R, Roos EM, et al. Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. *Arthritis Rheumatol*. 2014;66:622–636.

47 Kelley GA, Kelley KS, Hootman JM, Jones DL. Exercise and health-related quality of life in older community-dwelling adults: a meta-analysis of randomized controlled trials. *J Appl Gerontol*. 2009;28:369.

48 Karinkanta S, Nupponen R, Heinonen A, et al. Effects of exercise on health-related quality of life and fear of falling in home-dwelling older women. *J Aging Phys Act*. 2012;20:198–214.

49 Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care*. 1989;27(3 suppl):S217–S232.

50 Busija L, Osborne RH, Roberts C, Buchbinder R. Systematic review showed measures of individual burden of osteoarthritis poorly capture the patient experience. *J Clin Epidemiol*. 2013;66:826–837.

51 Angst F, Aeschlimann A, Steiner W, Stucki G. Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. *Ann Rheum Dis*. 2001;60:834–840.

52 Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthritis Cartilage*. 2004;12:389–399.

eAppendix 1.

Example of PubMed Search

((((“osteoarthritis”[MeSH Terms] OR “osteoarthritis”[All Fields]) AND “humans”[MeSH Terms]) OR (OA[All Fields] AND “humans”[MeSH Terms])) OR ((“arthritis”[MeSH Terms] OR “arthritis”[All Fields]) AND “humans”[MeSH Terms]) AND “humans”[MeSH Terms]) AND ((((((aquatics[All Fields] OR (aquatic[All Fields] AND (“rehabilitation” [Subheading] OR “rehabilitation”[All Fields] OR “rehabilitation”[MeSH Terms]))) OR (aquatic[All Fields] AND (“physical therapy modalities”[MeSH Terms] OR (“physical”[All Fields] AND “therapy”[All Fields] AND “modalities” [All Fields]) OR “physical therapy modalities”[All Fields] OR (“physical”[All Fields] AND “therapy”[All Fields]) OR “physical therapy”[All Fields]))) OR ((“water”[MeSH Terms] OR “water”[All Fields] OR “drinking water”[MeSH Terms] OR (“drinking”[All Fields] AND “water”[All Fields]) OR “drinking water”[All Fields]) AND (“rehabilitation” [Subheading] OR “rehabilitation”[All Fields] OR “rehabilitation”[MeSH Terms]))) OR (aquatic[All Fields] AND (“therapy”[Subheading] OR “therapy”[All Fields] OR “therapeutics”[MeSH Terms] OR “therapeutics”[All Fields]))) OR (aquatic[All Fields] AND (“exercise”[MeSH Terms] OR “exercise”[All Fields]))) OR (“hydrotherapy”[MeSH Terms] OR “hydrotherapy”[All Fields]) AND “humans”[MeSH Terms]) AND “humans”[MeSH Terms] AND (“0001/01/01”[PDAT] : “2013/11/30”[PDAT])

eAppendix 2.
Quality and Risk of Bias Assessment (PEDro)

Study	Random Allocation	Allocation Concealment	Baseline Similarity	Participant Blinding	Therapist Blinding	Assessor Blinding	Adequate Follow Up	Intention-to-Treat Analysis	Between Group Comparisons	Point & Variability Measures	Final PEDro Score (/10)
Patrick et al,[37] 2001	1	0	1	0	0	0	1	1	1	1	6
Foley et al,[27] 2003	1	1	1	0	0	1	1	1	1	1	8
Cochrane et al,[32] 2005	1	1	1	0	0	1	0	1	1	1	7
Wang et al,[33] 2007	1	0	1	0	0	0	1	1	1	1	6
Fransen et al,[30] 2007	1	1	1	0	0	1	1	1	1	1	8
Hinman et al,[31] 2007	1	1	1	0	0	1	1	1	1	1	8
Lund et al,[26] 2008	1	1	1	0	0	1	1	1	1	1	8
Lim et al,[34] 2010	1	0	1	0	0	1	1	1	1	1	7
Arnold and Faulkner,[36] 2010	1	0	1	0	0	1	0	1	1	1	6
Wang et al,[35] 2011	1	1	1	0	0	1	1	0	1	1	7
Hale et al,[25] 2012	1	1	1	0	0	1	1	1	1	1	8

eAppendix 3.

Hierarchy of Outcome Measures

Hierarchy of continuous pain-related outcomes[38]	
Ranking	**Outcome**
1	Global pain score
2	Pain on walking
3	WOMAC pain subscore
4	Composite pain scores other than WOMAC
5	Pain on activities other than walking
6	WOMAC global score
7	Lesquesne osteoarthritis index global score
8	Other algofunctional composite scores
9	Patient's global assessment
10	Physician's global assessment
Hierarchy of continuous self-reported functioning and quality-of-life outcomes (Adapted From the Findings of Veenhof et al[39])	
Ranking	**Outcome**
1	WOMAC VAS, version 3.0, or numerical scale
2	WOMAC Likert scale or VAS, version 3.0, modified
3	HOOS or KOOS or Lesquesne modified
4	SF-36 (component summary)
5	Health Assessment Questionnaire
6	Arthritis Impact Measurement Scales-2 Short Form (AIMS-2-SF)
7	SF-36 physical function or role–physical
8	Activities and Balance Confidence Scale

[a] WOMAC=Western Ontario and McMaster Universities Osteoarthritis Index, VAS=visual analog scale, HOOS=Hip Injury and Osteoarthritis Outcome Score, KOOS=Knee Injury and Osteoarthritis Outcome Score, SF-36=36-Item Short-Form Health Survey, AIMS-2=Arthritis Impact Measurement Scales-2 Short Form.

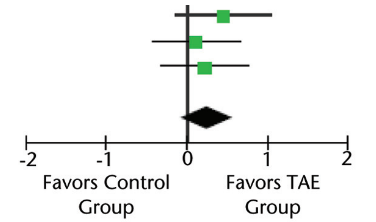
eAppendix 4.

Meta-Analysis for the Effect of Therapeutic Aquatic Exercise (TAE) on Knee Osteoarthritis[a]

2.1 Pain

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	$\bar{X}$	SD	Total	$\bar{X}$	SD	Total		
Lim et al,[34] 2010	−3.27	1.67	24	−4.12	2.08	20	29.1	0.45 (−0.15, 1.05)
Lund et al,[26] 2008	−55.8	21	27	−58.1	20	25	35.5	0.11 (−0.43, 0.65)
Wang et al,[35] 2011	72	18	26	68	18	26	35.4	0.22 (−0.33, 0.78)
Total (95% CI)			77			71	100	0.25 (−0.08, 0.57)
Heterogeneity: tau^2=0.00, χ^2=0.68, *df*=2 (*P*=.71), I^2=0%								
Test for overall effect: *z*=1.49 (*P*=0.14)								

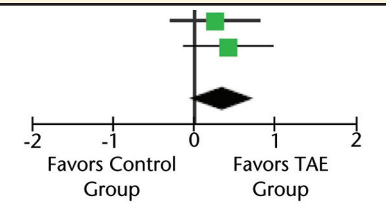
Standardized Mean Difference IV, Random (95% CI)



2.2 Stiffness

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	$\bar{X}$	SD	Total	$\bar{X}$	SD	Total		
Lund et al,[26] 2008	64.6	12	26	61.4	12	25	49.9	0.26 (−0.29, 0.81)
Wang et al,[35] 2011	69	20	26	61	17	26	50.1	0.42 (−0.13, 0.97)
Total (95% CI)			52			51	100	0.34 (−0.05, 0.73)
Heterogeneity: tau^2=0.00, χ^2=0.17, *df*=1 (*P*=.68), I^2=0%								
Test for overall effect: *z*=1.73 (*P*=.08)								

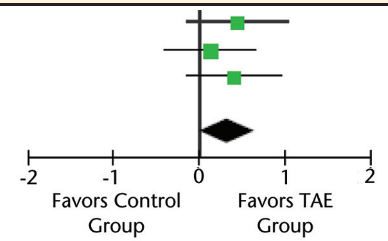
Standardized Mean Difference IV, Random (95% CI)



2.3 Self reported functioning

Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)
	$\bar{X}$	SD	Total	$\bar{X}$	SD	Total		
Lim et al,[34] 2010	−20.9	9.9	24	−27.6	18.3	21	29.4	0.46 (−0.14, 1.05)
Lund et al,[26] 2008	62.7	12	27	61.1	11	27	36.3	0.14 (−0.40, 0.67)
Wang et al,[35] 2011	76	16	26	69	18	26	34.3	0.40 (−0.14, 0.95)
Total (95% CI)			77			74	100	0.32 (0.00, 0.64)
Heterogeneity: tau^2=0.00, χ^2=0.74, *df*=2 (*P*=.69), I^2=0%								
Test for overall effect: *z*=1.96 (*P*=.05)								

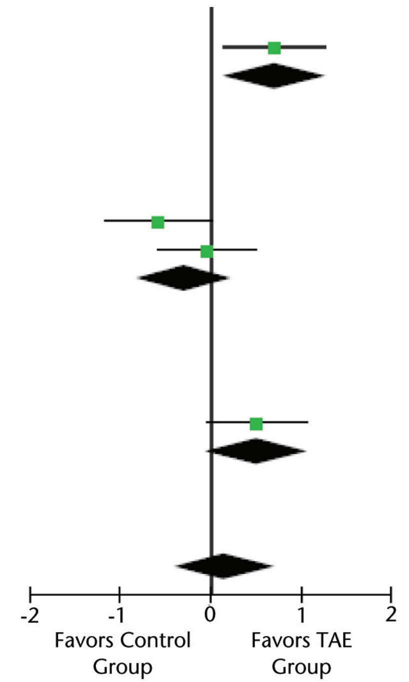
Standardized Mean Difference IV, Random (95% CI)



(Continued)

eAppendix 4.
Continued

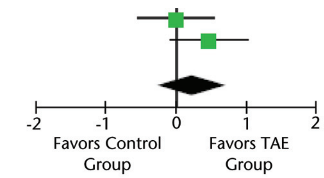
2.4 Physical functioning tests									
Study or Subgroup	**TAE Group**			**Control Group**			**Weight (%)**	**Standardized Mean Difference IV, Random (95% CI)**	**Standardized Mean Difference IV, Random (95% CI)**
	$\bar{X}$	**SD**	**Total**	**$\bar{X}$**	**SD**	**Total**			
2.4.1 Physical functioning tests (activities)									
Wang et al,[35] 2011	386	75.8	26	329.1	82.3	26	25.1	0.71 (0.15, 1.27)	
Subtotal (95% CI)			26			26	25.1	0.71 (0.15, 1.27)	
Heterogeneity: Not applicable									
Test for overall effect: z=2.47 (P=.01)									
2.4.2 Physical functioning tests (muscle strength)									
Lim et al,[34] 2010	50.4	19.3	24	63.1	24.1	21	24.2	−0.58 (−1.17, 0.02)	
Lund et al,[26] 2008	39.17	15.61	26	39.94	19.25	25	25.4	−0.04 (−0.59, 0.51)	
Subtotal (95% CI)			50			46	49.6	−0.30 (−0.82, 0.23)	
Heterogeneity: tau^2=0.06, χ^2=1.65, df=1 (P=.20), I^2=39%									
Test for overall effect: z=1.11 (P=.27)									
2.4.3 Physical functioning tests (joint range of motion)									
Wang et al,[35] 2011	−2.4	2.2	26	−3.3	1.1	26	25.3	0.51 (−0.04, 1.06)	
Subtotal (95% CI)			26			26	25.3	0.51 (−0.04, 1.06)	
Heterogeneity: not applicable									
Test for overall effect: z=1.81 (P=.07)									
Total (95% CI)			102			98	100	0.16 (−0.40, 0.71)	
Heterogeneity: tau^2=0.24, χ^2=11.52, df=3 (P=.009), I^2=74%									
Test for overall effect: z=0.55 (P=.58)									
Test for subgroup differences: χ^2=7.58, df=2 (P=.02), I^2=73.6%									



(Continued)

eAppendix 4.
Continued

2.5 Quality of Life

Study or Subgroup	TAE			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)	Standardized Mean Difference IV, Random (95% CI)
	$\overline{X}$	SD	Total	$\overline{X}$	SD	Total			
Lund et al,[26] 2008	43	12.47	27	43.1	11.18	25	50.5	−0.01 (−0.55, 0.54)	
Wang et al,[35] 2011	73	12	26	67	13	26	49.5	0.47 (−0.08, 1.02)	
Total (95% CI)			53			51	100	0.23 (−0.24, 0.70)	
Heterogeneity: $tau^2=0.04$, $\chi^2=1.48$, $df=1$ ($P=.22$), $I^2=32\%$									
Test for overall effect: $z=0.96$ ($P=.34$)									

[a] 95% CI=95% confidence interval.

eAppendix 5.

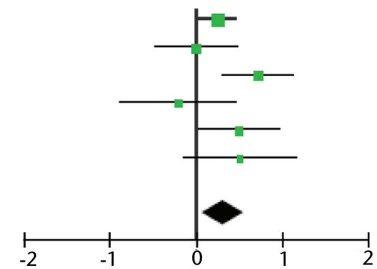
Meta-Analysis for the Effect of Therapeutic Aquatic Exercise (TAE) on Hip and Knee Osteoarthritis[a]

6 Knee and Hip

6.1 Pain

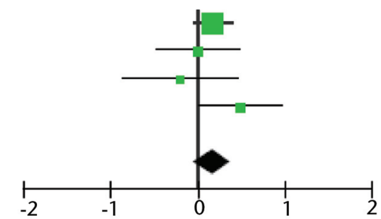
Study or Subgroup	TAE $\bar{X}$	TAE SD	TAE Total	Control Group $\bar{X}$	Control Group SD	Control Group Total	Weight (%)	Standardized Mean Difference IV, Random (95% CI)
Cochrane et al,[32] 2005	−8.46	3.74	153	−9.35	3.54	158	30.3	0.24 (0.02, 0.47)
Foley et al,[27] 2003	−10	2.222	34	−10	2.963	32	15.4	0.00 (−0.48, 0.48)
Fransen et al,[30] 2007	−27.3	18.7	55	−40	16.2	41	18.3	0.71 (0.30, 1.13)
Hale et al,[25] 2012	−7.8	3.925	20	−7.1	1.917	15	9.7	−0.21 (−0.88, 0.46)
Hinman et al,[31] 2007	−4	2	36	−5	2	35	15.8	0.49 (0.02, 0.97)
Wang et al,[33] 2007	−43.5	18.6	20	−54.9	25.2	18	10.3	0.51 (−0.14, 1.16)
Total (95% CI)			318			299	100	0.31 (0.07, 0.55)
Heterogeneity: tau^2=0.04, χ^2=8.76, df=5 (P=.12), I^2=43%								
Test for overall effect: z=2.57 (P=.01)								

Standardized Mean Difference IV, Random (95% CI)



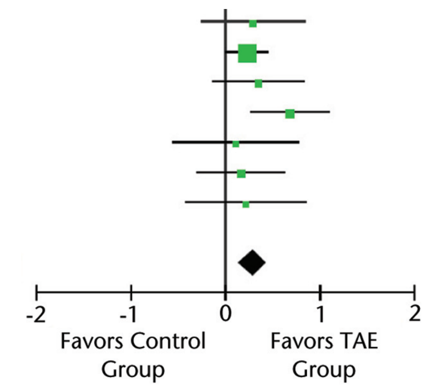
6.2 Stiffness

Study or Subgroup	TAE $\bar{X}$	TAE SD	TAE Total	Control Group $\bar{X}$	Control Group SD	Control Group Total	Weight (%)	Standardized Mean Difference IV, Random (95% CI)
Cochrane et al,[32] 2005	−3.88	1.67	152	−4.15	1.48	158	57.3	0.17 (−0.05, 0.39)
Foley et al,[27] 2003	−4	2.222	34	−4	2.222	32	16.6	0.00 (−0.48, 0.48)
Hale et al,[25] 2012	−3.7	1.392	20	−3.4	1.502	15	9.0	−0.20 (−0.87, 0.47)
Hinman et al,[31] 2007	−73	45	36	−95	44	35	17.2	0.49 (0.02, 0.96)
Total (95% CI)			242			240	100	0.16 (−0.04, 0.37)
Heterogeneity: tau^2=0.01, χ^2=3.41, df=3 (P=.33), I^2=12%								
Test for overall effect: z=1.55 (P=.12)								



(Continued)

eAppendix 5.
Continued

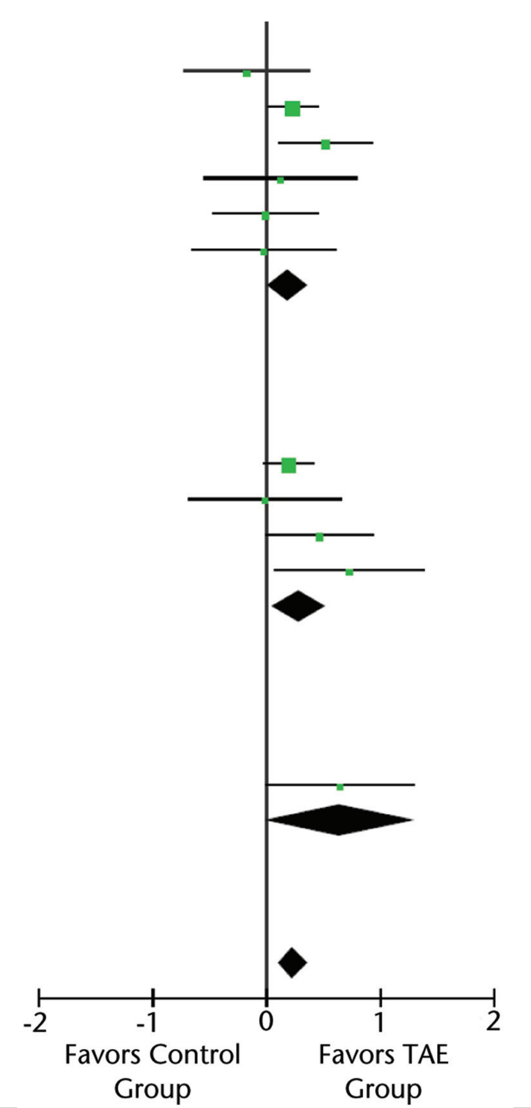
6.3 Self-reported functioning									
Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)	Standardized Mean Difference IV, Random (95% CI)
	X̄	SD	Total	X̄	SD	Total			
Arnold and Faulkner,[36] 2010	69.6	24.4	26	62.9	20.8	25	7.8	0.29 (−0.26, 0.84)	
Cochrane et al,[32] 2005	−29.26	14.48	149	−32.42	13.25	156	46.6	0.23 (0.00, 0.45)	
Foley et al,[27] 2003	−33	12.593	34	−37	9.63	32	10.0	0.35 (−0.14, 0.84)	
Fransen et al,[30] 2007	−34.8	23.7	55	−49.9	19	41	13.6	0.69 (0.27, 1.10)	
Hale et al,[25] 2012	−24	8.899	20	−24.9	7.015	15	5.3	0.11 (−0.56, 0.78)	
Hinman et al,[31] 2007	−598	316	36	−656	373	35	10.9	0.17 (−0.30, 0.63)	
Wang et al,[33] 2007	−0.9	0.4	20	−1	0.5	18	5.8	0.22 (−0.42, 0.86)	
Total (95% CI)			340			322	100	0.29 (0.14, 0.45)	
Heterogeneity: tau^2=0.00, χ^2=4.44, df=6 (P=.62), I^2=0%									
Test for overall effect: z=3.74 (P=.0002)									

(Continued)

eAppendix 5.
Continued

6.4 Physical functioning tests								
	TAE Group			**Control Group**				
Study or Subgroup	**$\overline{X}$**	**SD**	**Total**	**$\overline{X}$**	**SD**	**Total**	**Weight (%)**	**Standardized Mean Difference IV, Random (95% CI)**
6.4.1 Physical functioning tests (activities)								
Arnold et al,[36] 2010	7.5	3.9	26	8.1	2.6	25	5.4	−0.18 (−0.73, 0.37)
Cochrane et al,[32] 2005	−3.24	1.93	153	−3.78	2.72	159	26.0	0.23 (0.01, 0.45)
Fransen et al,[30] 2007	−8.2	1.7	55	−9.2	2.2	41	9.2	0.51 (0.10, 0.93)
Hale et al,[25] 2012	−10.1	2.9	20	−10.7	6.8	15	3.7	0.12 (−0.55, 0.79)
Hinman et al,[31] 2007	−10.32	1.94	36	−10.3	2.78	35	7.4	−0.01 (−0.47, 0.46)
Wang et al,[33] 2007	388.4	80.3	20	390.7	88.6	18	4.1	−0.03 (−0.66, 0.61)
Subtotal (95% CI)			310			293	55.6	0.18 (0.01, 0.35)
Heterogeneity: tau^2=0.00, χ^2=5.40, *df*=5 (*P*=.37), I^2=7%								
Test for overall effect: *z*=2.02 (*P*=.04)								
6.4.2 Physical functioning tests (muscle strength)								
Cochrane et al,[32] 2005	166.02	108.92	152	145.92	99.82	158	25.9	0.19 (−0.03, 0.42)
Hale et al,[25] 2012	18.8	8.465	20	19	12.528	15	3.7	−0.02 (−0.69, 0.65)
Hinman et al,[31] 2007	29.9	12.5	36	24.7	9.5	35	7.2	0.46 (−0.01, 0.93)
Wang et al,[33] 2007	16	4.5	20	12.9	3.8	18	3.8	0.73 (0.07, 1.39)
Subtotal (95% CI)			228			226	40.5	0.28 (0.05, 0.52)
Heterogeneity: tau^2=0.01, χ^2=3.64, *df*=3 (*P*=.30), I^2=18%								
Test for overall effect: *z*=2.39 (*P*=.02)								
6.4.3 Physical functioning tests (joint range of motion)								
Wang et al,[33] 2007	−1.2	2.6	20	−3.2	3.5	18	3.8	0.64 (−0.01, 1.29)
Subtotal (95% CI)			20			18	3.8	0.64 (−0.01, 1.29)
Heterogeneity: not applicable								
Test for overall effect: *z*=1.92 (*P*=.06)								
Total (95% CI)			558			537	100	0.23 (0.10, 0.36)
Heterogeneity: tau=0.00, χ^2=10.93, *df*=10 (*P*=.36), I^2=9%								
Test for overall effect: *z*=3.51 (*P*=.0006)								
Test for subgroup differences: χ^2=2.03, *df*=2 (*P*=.36), I^2=1.4%								

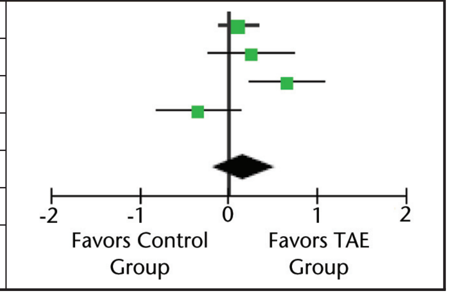
Standardized Mean Difference IV, Random (95% CI)



(Continued)

eAppendix 5.
Continued

3.5 Quality of life									
Study or Subgroup	TAE Group			Control Group			Weight (%)	Standardized Mean Difference IV, Random (95% CI)	Standardized Mean Difference IV, Random (95% CI)
	$\bar{X}$	SD	Total	$\bar{X}$	SD	Total			
Cochrane et al,[32] 2005	62.55	18.61	150	60.68	17.39	157	31.8	0.10 (−0.12, 0.33)	
Foley et al,[27] 2003	53.3	11.481	34	50.5	10.37	32	21.7	0.25 (−0.23, 0.74)	
Fransen et al,[30] 2007	54.6	8.9	55	48	11.4	41	24.3	0.65 (0.24, 1.07)	
Hinman et al,[31] 2007	0.43	0.2	36	0.5	0.2	35	22.3	−0.35 (−0.82, 0.12)	
Total (95% CI)			275			265	100	0.17 (−0.19, 0.52)	
Heterogeneity: tau^2=0.01, χ^2=10.25, *df*=3 (*P*=.02), I^2=71%									
Test for overall effect: *z*=0.93 (*P*=.35)									



[a] 95% CI=95% confidence interval.

Physical Therapy
Journal of the American Physical Therapy Association and de Fysiotherapeut
Royal Dutch Society for Physical Therapy

PTJ Online First

Effect of Therapeutic Aquatic Exercise on Symptoms and Function Associated With Lower Limb Osteoarthritis: A Systematic Review With Meta-Analysis

Benjamin Waller, Anna Ogonowska-Slodownik, Manuel Vitor, Johan Lambeck, Daniel Daly, Urho M. Kujala and Ari Heinonen

PHYS THER. Published online June 5, 2014
Originally published online June 5, 2014
doi: 10.2522/ptj.20130417