

BMJ Open Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a systematic review with meta-analysis of randomised controlled trials

George A Kelley,¹ Kristi S Kelley,² Leigh F Callahan³

To cite: Kelley GA, Kelley KS, Callahan LF. Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a systematic review with meta-analysis of randomised controlled trials. *BMJ Open* 2018;**8**:e019138. doi:10.1136/bmjopen-2017-019138

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-019138>).

Received 16 August 2017
Revised 3 November 2017
Accepted 29 November 2017



¹Department of Biostatistics, School of Public Health, West Virginia University, Morgantown, West Virginia, USA

²Department of Biostatistics, School of Public Health, Robert C Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia, USA

³Departments of Social Medicine and Orthopaedics, Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, USA

Correspondence to

Professor George A Kelley;
gkelley@hsc.wvu.edu

ABSTRACT

Background/purpose Given conflicting findings, the purpose of this study was to use the meta-analytic approach to examine the effects of exercise (aerobic, strength training or both) on anxiety in adults with arthritis and other rheumatic diseases (AORD).

Methods Randomised controlled exercise intervention trials ≥ 4 weeks in adults ≥ 18 years of age with osteoarthritis, rheumatoid arthritis or fibromyalgia were included. Studies were located by searching eight electronic databases, cross-referencing and expert review. Dual selection and data abstraction of studies were performed. Hedge's standardised effect size (ES) was calculated for each result and pooled using the recently developed inverse heterogeneity model. Two-tailed z -alpha values ≤ 0.05 and non-overlapping 95% CI were considered statistically significant. Heterogeneity was estimated using Q and I^2 with alpha values ≤ 0.10 for Q considered statistically significant. Small-study effects were examined using funnel plots and Egger's regression test. In addition, the number needed to treat (NNT), percentile improvement and meta-regression were conducted.

Results Of the 639 citations screened, 14 studies representing 926 initially enrolled participants (539 exercise, 387 control) met the criteria for inclusion. Length of training (mean \pm SD) averaged 15.8 \pm 6.7 weeks, frequency 3.3 \pm 1.3 times per week and duration 28.8 \pm 14.3 min per session. Overall, statistically significant reductions in anxiety were found (exercise minus control changes ES= -0.40 , 95% CI -0.65 to -0.15 , tau²=0.14; $Q=40.3$, $P=0.0004$; $I^2=62.8\%$). The NNT was 6 with a percentile improvement of 15.5% and an estimated 5.3 million inactive US adults with AORD improving their anxiety if they started exercising regularly. Statistically significant small-study effects were observed ($P<0.0001$).

Conclusions Exercise is associated with reductions in anxiety among adults with selected types of AORD. However, a need exists for additional, well-designed, randomised controlled trials on this topic.

PROSPERO registration number CRD42016048728.

Strengths and limitations of this study

- To the best of the investigative team's knowledge, this is the first systematic review with meta-analysis to examine the effects of exercise on anxiety as a primary outcome in adults with arthritis and other rheumatic diseases.
- The use of the recently developed inverse heterogeneity model appears to provide more robust estimates than those derived from other models.
- Common to all aggregate data meta-analyses, the possibility of ecological fallacy exists and the meta-regression analyses conducted do not allow for causal inferences.
- Given that no adjustments for multiple testing were made because of concerns about missing possibly important findings that could be tested in original trials, the possibility of chance findings exists.

INTRODUCTION

Rationale

Arthritis and other rheumatic diseases (AORD) are major public health problems in the USA. Based on combined 2013–2015 data from the National Health Interview Survey, the annual prevalence of doctor-diagnosed arthritis in the civilian, non-institutionalised US population aged 18 years or older was 22.7% (54.4 million), with prevalence higher among women (23.5%) than men (18.1%).¹ By 2040, it is estimated that 78.4 million (25.9%) US adults 18 years of age and older will have doctor-diagnosed arthritis.² Compared with combined 2013–2015 data,¹ this represents an increase of approximately 24 million adults. Not surprisingly, the financial costs associated with AORD in the USA are high. In 2003, the total costs attributable to AORD were estimated to be approximately \$128 billion: \$80.8 billion in direct costs (medical expenditures)

and \$47 billion in indirect costs (lost earnings).³ This represents an increase of 24% between 1997 and 2003 and was primarily the result of an increase in the number of people with AORD.³

Elevated and sustained levels of anxiety can result in a number of deleterious consequences. These include, but are not limited to: (1) an increased risk for coronary heart disease as a result of heightened arousal leading to an increased risk for hypertension and a proinflammatory state,⁴⁻⁶ (2) an increased risk for cardiac death,⁵ and (3) poorer health-related quality of life.⁷ While it is well recognised that depression is a common comorbidity among adults with AORD, recent research suggests that the prevalence of anxiety among US adults with arthritis is approximately twice as high as depression.⁸ Using data from the Arthritis Conditions Health Effects Survey, the prevalence of anxiety and depression among US adults with doctor-diagnosed arthritis was estimated to be almost twice as high for anxiety (30.5%) versus depression (17.5%), with US population estimates of 11.5 million for anxiety and 6.6 million for depression.⁸ In addition, an 18-country study found the prevalence of anxiety to be greater in adults with AORD versus those without AORD.⁹ This included, separately, generalised anxiety disorder, social phobia, agoraphobia/panic disorder and post-traumatic stress disorder.⁹ Given the prevalence of anxiety, it has been recommended that healthcare providers screen people with AORD for anxiety.⁸

Exercise is an intervention that is generally safe and appropriate for most persons with various types of AORD.^{10 11} While there is no firm consensus regarding the mechanisms associated with the anxiolytic effects of exercise in adults with AORD as well as other chronic diseases, a recent review on this topic has reported several potential mechanisms, exclusive of specific chronic disease status, that may be responsible for such.¹² These include physiological (decrease in sympathetic nervous system and hypothalamic-pituitary-adrenal axis reactivity) as well as psychological (improvements in anxiety sensitivity, self-efficacy and distraction) processes.¹²

Recent meta-analytic work has shown that community-deliverable exercise interventions reduce depressive symptoms in adults, with an estimated 3.1 million inactive US adults with AORD improving their depressive symptoms if they began and maintained a regular exercise programme.¹³ However, the effects of community-deliverable exercise on anxiety as a primary outcome are not known given a plethora of conflicting randomised controlled trials on this topic as well as a lack of studies that assess both depression and anxiety within the same study (only 44.8% based on a previous meta-analysis).¹³ Most importantly, a recent systematic review of previous meta-analyses, not to be confused with an original systematic review with meta-analyses, found that no meta-analysis of randomised controlled trials has examined the effects of community-deliverable exercise on anxiety as a primary outcome in adults with AORD.¹⁴ Community-deliverable exercise may be especially

important given its potential convenience, ability to reach a large amount of people and cost.¹⁵⁻¹⁸ In addition, a systematic review found that greater levels of anxiety were a barrier to exercise adherence in those with AORD.¹⁹ Thus, reductions in anxiety may increase adherence to exercise in adults with AORD, a population in which meeting current exercise guidelines is significantly lower than in those without AORD, approximately 40% vs 50%.²⁰ Clearly, it is critically important to develop a better understanding of the overall magnitude of effect, as well as factors associated with exercise for improving anxiety in adults with AORD.

Objective

The primary objective of this study was to conduct a systematic review with an aggregate data meta-analysis of randomised controlled trials to determine the effects of community-deliverable exercise interventions on anxiety in adults with AORD.

METHODS

Overview

This study followed the guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement for meta-analyses of healthcare interventions.²¹ The protocol for this study is registered in PROSPERO (trial registration number CRD42016048728) and has been previously published in *BMJ Open*.²²

Eligibility criteria

Studies that met the following criteria were included: (1) randomised controlled trials with the unit of assignment at the participant level, (2) community-deliverable exercise-only intervention group (aerobic, strength training or both), (3) interventions ≥ 4 weeks, (4) comparative control group (non-intervention, wait list control, usual care, attention control); (5) adults ≥ 18 years of age with doctor-diagnosed osteoarthritis, rheumatoid arthritis or fibromyalgia, (6) studies (published and unpublished in the form of master's theses and dissertations) in any language, assuming an English-language abstract was available, from 1 January 1981 onwards, and (7) data for anxiety, as defined by the authors of the original studies. Studies were limited to those with osteoarthritis, rheumatoid arthritis or fibromyalgia based on previous meta-analytic research showing a lack of exercise intervention studies for other types of AORD.¹³ Studies were limited to randomised trials because it is the only way to control for confounders that are not known or measured as well as the observation that non-randomised controlled trials tend to overestimate the effects of healthcare interventions.^{23 24} Aerobic and progressive resistance (strength training) exercises were defined according to section C2 of the 2008 Physical Activity Guidelines Advisory Committee Report.²⁵ Specifically, aerobic exercise is defined as any 'exercise that primarily uses the aerobic energy-producing systems, can improve the capacity and

efficiency of these systems, and is effective for improving cardiorespiratory endurance,' while strength training is defined as 'exercise training primarily designed to increase skeletal muscle strength, power, endurance, and mass.'²⁵

For this proposed project, community-deliverable exercise interventions were considered to be those that could be performed, or have the potential to be adapted and performed, by persons in a community setting (recreation or senior centres, in the home or neighbourhood, etc.) and meet the implementation guidelines for physical activity interventions recently recommended by the Arthritis Program at the Centers for Disease Control and Prevention: (1) no academic degree required for a leader/implementer but leader training available, if needed, (2) no special facilities beyond a community room (except a warm pool for aquatic exercise), (3) inexpensive equipment, (4) cost to participants less than \$50.00, (5) implementation guide available, (6) supporting structures judged to be adequate to support widespread implementation.²⁶ An exercise duration of at least 4 weeks was chosen based on previous research in which statistically significant improvements in anxiety occurred as a result of as little as 4 weeks of exercise training.²⁷ There was no maximum limit on the length of any interventions for the studies included in our proposed project. We limited our studies to adults aged 18 and older because the inclusion of children and adolescents posed additional confounding problems congruent with the many developmental changes that occur during this period. In addition, the prevalence of AORD is more common in adults than children and adolescents. We restricted our studies to published articles, dissertations and master's theses and examined for potential small-study effects such as publication bias when limited to published articles in peer-reviewed journals. The year 1981 was chosen as the starting point for eligibility based on a preliminary PubMed search in which it was found that this was the first year that a randomised controlled trial on exercise and arthritis was published.²⁸ Studies from both English and non-English-language sources were included with the latter translated into English by the second author using the freely available web-based Babelfish and Bing translators. Finally, while acknowledging that the mechanisms for changes in anxiety may differ for different types of AORD, for example, fibromyalgia versus osteoarthritis and rheumatoid arthritis, all three were included in the current systematic review with meta-analysis because the focus was on determining whether exercise reduces anxiety in adults with AORD, not why it reduces anxiety. In addition, community-deliverable exercise programmes, and exercise programmes in general, are not traditionally focused on one type of AORD, but rather, across all types of AORD.¹⁷ Furthermore, metaregression (see Data Synthesis section) was conducted to examine for any potential associations between type of AORD and changes in anxiety. Finally, there has been recent criticism by others regarding the publication of systematic

reviews with meta-analysis using the 'least publishable unit' approach.²⁹

Information sources

Eight electronic databases were searched for potentially eligible studies in any language published from 1 January 1981 onwards. The last searches were conducted on 6 January 2017. Databases searched included: (1) PubMed, (2) SPORTDiscus, (3) Cochrane Central Register of Controlled Clinical Trials (CENTRAL), (4) Cumulative Index to Nursing and Allied Health Literature (CINAHL), (5) PsycINFO, (6) Web of Science, (7) Scopus and (8) ProQuest (master's theses and dissertations). In addition, cross-referencing from retrieved studies was conducted. Furthermore, the third author, an expert on physical activity and AORD, reviewed the reference list for completeness.

Search strategy

Search strategies were developed using text words as well as Medical Subject Headings associated with the effects of exercise on anxiety in adults with AORD. The second author conducted all electronic database searches. A copy of the search strategy for one of the databases searched (PubMed) is shown in online supplementary file 1.

Study records

Study selection

All studies to potentially be screened were imported into EndNote (EndNote, V.X8. New York, NY: Thomson Reuters, 2016). Duplicates were then removed both electronically and manually by the second author. A copy of the database was then provided to the first author for duplicate screening. The first two authors selected all studies, independent of each other. The full report for each article was obtained for all titles and abstracts that appeared to meet the inclusion criteria or where there was any uncertainty. Multiple reports of the same study were handled by including the most recently published article as well as drawing from previous reports, assuming similar methods and sample sizes. Neither of the screeners were blinded to the journal titles or to the study authors or institutions. Reasons for exclusion were coded as one or more of the following: (1) inappropriate population, (2) inappropriate intervention, (3) inappropriate comparison(s), (4) inappropriate outcome(s), (5) inappropriate study design, (6) other. Upon completion, the first two authors met and reviewed their selections. Discrepancies were resolved by consensus. If consensus could not be reached, the third author provided a recommendation. Using Cohen's kappa statistic (κ),³⁰ the overall agreement rate prior to correcting disagreements was 0.90. After identifying the final number of studies to be included, the overall precision of the searches was calculated by dividing the number of studies included by the total number of studies screened

after removing duplicates.³¹ The number needed to read (NNR) was then calculated as the inverse of the precision.³¹

Data abstraction

Prior to the abstraction of data, a codebook that could hold more than 200 items per study was developed using Microsoft Excel (Microsoft Excel, V.2010, Redmond, WA). The codebook was developed by the first two authors with input from the third author. The major categories of variables coded included: (1) study characteristics (author, journal, year, etc.), (2) participant characteristics (age, height, body weight, etc.), (3) intervention characteristics (type, length, frequency, intensity, duration, compliance, etc.), and (4) outcome characteristics for anxiety (sample sizes, baseline and postexercise means and SD, etc.). The first two authors abstracted data from all studies, independent of each other, using separate codebooks in Microsoft Excel. Upon completion of coding, the codebooks were merged into one primary codebook for review. Both authors then met and reviewed every item for agreement. Discrepancies were resolved by consensus. If consensus could not be reached, the third author provided a recommendation. Using Cohen's kappa statistic (κ),³⁰ the overall agreement rate prior to correcting disagreements was 0.95.

Outcomes and prioritisation

The primary outcome in this study was changes in anxiety. Secondary outcomes included changes in physical function, pain, depression, quality of life, body mass index (BMI, in $\text{kg}\cdot\text{m}^{-2}$), maximum oxygen consumption ($\text{VO}_{2\text{max}}$, in $\text{mL}/\text{kg}/\text{min}$) and muscular strength.

Risk of bias assessment in individual studies

Risk of bias was assessed at the study level using the Cochrane Risk of Bias instrument,³² with a focus on the primary outcome of interest, changes in anxiety. Bias was evaluated across six domains: (1) random sequence generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, (6) selective reporting, and (7) whether participants were exercising regularly, as defined by the original study authors, prior to taking part in the study. Each item was classified as having either a high, low or unclear risk of bias. Because it is virtually impossible to blind participants to group assignment in exercise intervention protocols, all studies were classified as high risk with respect to the category 'blinding of participants and personnel'. Based on previous research, no study was excluded based on the results of the risk of bias assessment.³³ The first two authors conducted all assessments, independent of each other. Both authors then met and reviewed all selections for agreement. Discrepancies were resolved by consensus, and if necessary, consultation with the third author. Using Cohen's kappa statistic (κ),³⁰ the overall agreement rate

for risk of bias assessment prior to correcting disagreements was 0.71.

Data synthesis

Calculation of effect sizes

The primary outcome for this proposed project was changes in anxiety, calculated as Hedge's standardised mean difference effect size (ES), adjusted for small-sample bias.³⁴ This was calculated by subtracting the change outcome difference in the exercise group minus the change outcome difference in the control group, and then dividing by the pooled SD of the change outcome for the exercise and control groups. If change score SDs were not available, they were calculated from reported change outcomes, treatment effect 95% confidence intervals (CI's), or pre and post-SD values according to procedures developed by Follmann *et al.*³⁵ Secondary outcomes (physical function, pain, quality of life, depression, muscular strength) were calculated using the same procedures as for the primary outcome while BMI and $\text{VO}_{2\text{max}}$ were calculated using the original metric. For studies in which outcomes were assessed at multiple time points, differences between baseline values and the final time point closest to cessation of the exercise intervention were used.

Pooled estimates for changes in outcomes

ES changes in anxiety and all secondary outcomes were pooled using the recently developed inverse heterogeneity (IVhet) model.³⁶ The IVhet model is a quasilielihood model that is computed by: (1) calculating weights that sum to 1 from each study, (2) pooling effects from all the studies, and (3) calculating the variance of the pooled ES. The IVhet model has been shown to be more robust than the original random effects, method of moments model of DerSimonian and Laird,³⁷ the most common random effects model used to pool aggregate data meta-analytic results.³⁸ Specifically, simulation studies have shown that the IVhet model retains correct coverage probabilities as well as a lower observed variance than the random effects model, regardless of heterogeneity.³⁶ Two-tailed z -alpha values ≤ 0.05 and non-overlapping 95% CIs were considered statistically significant.

Heterogeneity and *inconsistency* for each pooled outcome were estimated using the Q ³⁹ and I^2 statistics,⁴⁰ respectively. An alpha level of ≤ 0.10 for Q was considered to represent statistically significant heterogeneity while inconsistency was categorised as very low ($<25\%$), low (25% to $<50\%$), moderate (50% to $<75\%$) or large ($\geq 75\%$).⁴⁰ To improve practical relevance with respect to potential improvements in anxiety and all secondary outcomes, percentile gain in the exercise groups was calculated using Cohen's U_3 index.⁴¹ In addition, the number needed to treat (NNT) was estimated using the approach suggested by the Cochrane Collaboration³² and a control group risk of 30% based on previous research on anxiety.⁴² This was accomplished by converting the standardised mean differences into a log OR, OR, assumed

control risk (30%) and then NNT. Extending the NNT, estimates of the number of US adults with AORD who could improve their anxiety levels but were not currently meeting physical activity recommendations were calculated. These were based on the reciprocal of the NNT multiplied by the number of adults in the USA with doctor-diagnosed arthritis who were not currently meeting exercise guidelines, currently approximately 31.6million.^{1 43} *Influence analysis* was conducted with each study deleted from the model once while *cumulative meta-analysis*, ranked by year, was used to examine the accumulation of results over time.

The primary pooling of results was based on group level findings. However, findings were also examined by collapsing results so that only one ES represented each study. In addition, for those outcomes that were assessed using multiple instruments, results from these instruments were pooled into one ES. The rationale for this was based on the lack of consensus regarding the most valid and reliable instrument for the outcome and population of interest as well as the need to maintain as much independence as possible. For those studies that reported both per-protocol and intention-to-treat analyses, results were also pooled into one ES.

Exploratory *metaregression* based on the IVhet model was used to examine the relationship between changes in our primary outcome (anxiety) and selected covariates.³⁶ This was accomplished by: (1) conducting simple metaregression for statistically significant associations between selected covariates and changes in anxiety, (2) examining for multicollinearity between covariates ($r > 0.80$), and (3) building a multiple metaregression model. These models used a multiplicative versus additive component for residual heterogeneity. To achieve matching error variances, robust Huber-Ecker-White-sandwich error variances were used to account for the underestimated dispersion. Such errors were expected to calculate the correct SEs for heterogeneous data that are traditionally heteroskedastic. Because of the small sample size, a post hoc decision was made to include no more than three potential predictors based on simple regression analyses. Multivariate imputation using chained equations⁴⁴ was used to impute missing values for minutes of training per week (five imputations) and changes in depression (one imputation) so that the maximum sample size could be achieved. Because of the small sample size, a post hoc decision was made to conduct metaregression with results for osteoarthritis and rheumatoid arthritis combined.

Based on the recommendations of Rothman,⁴⁵ no adjustments for multiple testing were made because of concerns about missing possibly important findings that could be pursued in future randomised controlled trials. While this could be viewed as a 'fishing expedition', such analyses are important for providing investigators with potential direction for future randomised controlled trials, one of the very reasons for conducting a systematic review with meta-analysis.⁴⁶ This approach is especially

appropriate for meta-analysis since covariates are not randomly assigned in meta-analysis and thus, such analyses are considered to be observational in nature.⁴⁷ As a result, causal inferences cannot be derived from metaregression. However, any observed associations can provide direction for future research.

Meta-biases

Small-study effects (publication bias, etc.) for primary and secondary outcomes were assessed following current guidelines.⁴⁸ This included qualitative analysis using funnel plots as well as quantitative analysis using Egger's regression intercept test for continuous data.⁴⁸ An alpha level $\leq 0.05\%$ and 95% CI that did not include zero (0) were considered to represent statistically significant small-study effects.

Confidence in cumulative evidence

Strength of findings for our primary outcome (anxiety) was evaluated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) for meta-analysis.⁴⁹ The quality of evidence was assessed across the domains of risk of bias, consistency, directness, precision and publication bias. Quality was judged as high (further research is very unlikely to change our confidence in the estimate of effect), moderate (further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate), low (further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate), or very low (very uncertain about the estimate of effect).

Software used for data synthesis

All data were analysed using Stata (V.14.1; Stata/SE for Windows, College Station, Texas: Stata, 2015), Microsoft Excel V.2010 and three add-ins for Excel: Meta XL (V.5.3; 2016), SSC-Stat (V.2.18, University of Reading, UK: Statistical Services Center, 2007) and EZ-Analyze (V.3.0, Boston, MA: Tim Poynton, 2007).

RESULTS

Study characteristics

A flow diagram that depicts the search process for study selection is shown in [figure 1](#). After initially identifying 639 citations and removing 240 duplicates both electronically and manually, 399 citations were screened. Of these, 14 studies representing 30 groups (16 exercise, 14 control) and 926 initially enrolled participants (539 exercise, 387 control) met the criteria for inclusion.⁵⁰⁻⁶³ Two studies included more than one intervention group.^{58 61} One included a pool and walking group⁵⁸ while the other included a long and short bout exercise group.⁶¹ The major reasons for exclusion were: (1) inappropriate intervention (47.8%), (2) inappropriate study design (32.2%), (3) inappropriate population (11.9%), (4) inappropriate comparison group (4.7%),

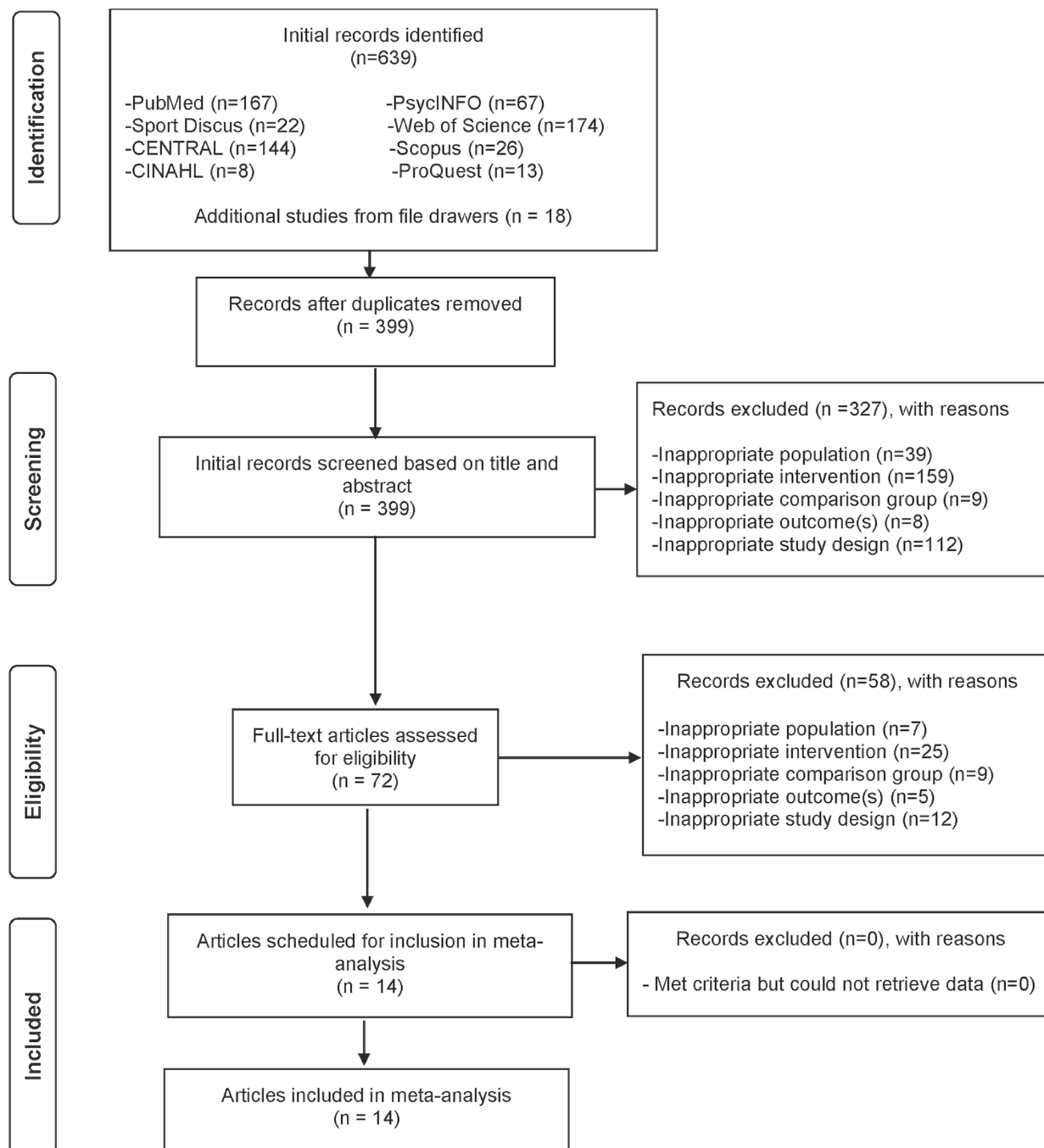


Figure 1 Flow diagram depicting the search process.

and (5) inappropriate outcome (3.4%). The precision of the search, excluding duplicates, was 3.5% while the NNR was 29. A list of excluded studies, including the reasons for exclusion, can be found in online supplementary file 2.

With respect to source, 1 study (7.1%) was published as a dissertation⁵¹ while the remaining 13 (92.9%) were published in peer-reviewed journals.^{50 52-63} For the 12 studies (85.7%) in which data were available, the 5-year impact factor of the journals in which studies were published ranged from 1.7 to 7.9 ($\bar{X}\pm\text{SD}$, 4.1 ± 2.4 , median=2.9). Five studies (35.7%) were published in the USA,^{51 52 56-58} two each in either Canada,^{55 61} Portugal^{62 63}

or Spain,^{59 60} and one each in either Australia,⁵³ Brazil⁵⁰ or Ireland.⁵⁴ Thirteen of the 14 studies (92.9%) were published in English-language journals⁵⁰⁻⁶² while 1 (7.1%) was published in Spanish.⁶³ With respect to types of control groups, four studies used either a wait list control^{50 51 53 54} or what appeared to be some type of attention control group,^{52 56 58 61} while three each used either a non-intervention^{55 57 62} or usual care group.^{59 60 63} Three studies (21.4%) reported using matching procedures according to either sex,⁵⁵ age, sex and BMI,⁵⁶ or type of arthritis (rheumatoid or osteoarthritis).⁵⁸ None of the studies reported using any type of crossover design.⁵⁰⁻⁶³ For data analysis, seven studies (50.0%) reported using

the per-protocol approach,^{51 56–58 60 62 63} four (28.6%) used intention to treat,^{50 52–54} while the remaining three (21.4%) used both.^{55 59 61} Eight studies (57.1%) provided sample size estimates.^{50 52–54 56 60 61 63} With respect to funding, 12 studies (85.7%) reported receiving funding from government, university or private sources for their work.^{50 52–62}

Participant characteristics

A description of the physical characteristics of participants for those studies that reported data is shown in [tables 1 and 2](#). On average, participants were overweight and had low cardiorespiratory fitness. The majority of participants were women, although seven studies (50.0%) also included a small number of men.^{52–58} Nine studies (64.3%) were limited to participants with fibromyalgia,^{50 51 55 56 59–63} three (21.4%) with osteoarthritis,^{52–54} one with rheumatoid arthritis⁵⁷ and one with either rheumatoid arthritis or osteoarthritis.⁵⁸ For those studies in which data were available, the number of years in which rheumatic symptoms were present ranged from 3 to 24 ($\bar{X}\pm\text{SD}$, 13.1 \pm 7.0, median=12.0),^{54–56 59 61–63} while years since diagnosis ranged from 3 to 7 ($\bar{X}\pm\text{SD}$, 4.1 \pm 2.0, median=3).^{55 61–63} Nine of the 14 studies (64.3%) reported that one or more participants were taking some type of medication for their condition.^{51 53–55 57–59 62 63} These included, but were not necessarily limited to, non-steroidal anti-inflammatory drugs, analgesics, narcotics and non-narcotics for pain, muscle relaxants, antidepressants and anxiolytics. However, a lack of specific data was available on exposure to medications both before and during the study, including any changes during the intervention period. In addition, a lack of data was available on pharmacological equivalence between study arms both before and during the study. For cigarette smoking, one study reported cigarette smoking in some of the participants.⁶¹ Participant withdrawals or removal in the exercise groups ranged from 0% to 50% ($\bar{X}\pm\text{SD}$, 17.4 \pm 13.4, median=17.0) while withdrawals or removal in the control groups ranged from 0% to 41% ($\bar{X}\pm\text{SD}$, 11.5 \pm 12.7, median=6.3). Reasons for withdrawals or removal included such things as family issues, pain from exercising, injuries, personal issues, time, unhappiness with group assignment, transportation issues, moving, employment commitments, boredom with exercise routine, not enough room or privacy to perform exercise, failure to complete lab assessments, not attending a specific percentage of the exercise sessions, or changing medications that could affect mood.

Exercise intervention characteristics

Exercise intervention characteristics for each group from each study are shown in [tables 1 and 3](#). Intensity of training, categorised according to American College of Sports Medicine⁶⁴ and limited to aerobic exercise only, included very light, light, moderate and vigorous exercises for the nine (64.3%) studies that reported such information.^{51 55 56 58–63} For mode of training, six studies (42.9%) focused on aerobic types of exercise,^{50 51 55 58 60 61} one (7.1%) on

weight training⁵⁷ and seven (50.0%) on both.^{52–54 56 59 62 63} Specific types of aerobic activities included such things as aquatic exercise, walking, jogging, exercising to music and cycling. For the few studies that provided detailed information on resistance training, the number of sets ranged from 1 to 3,^{57 59} the number of repetitions from 8 to 15^{57 59} and the number of exercises from 2 to 10 or more.^{56 57 59} One study reported a rest period for 30s between sets.⁵⁷ For those studies that reported data, the equipment used for resistance training included free weights and elastic bands.^{52 56 57 62 63} Seven studies (50.0%) reported supervised exercise,^{55 56 58–60 62 63} five (35.7%) reported both supervised and unsupervised exercises,^{50–54} while the remaining two (14.3%) reported unsupervised exercise.^{57 61} Three of the 14 studies (21.4%) reported some type of adverse event.^{51 52 61} Reasons included one participant dropping out due to pain after the first exercise session,⁵¹ two participants due to a history of a herniated disk and low back and leg pain⁵² and one participant due to a metatarsal fracture.⁶¹ Another study reported that one participant withdrew because of the exacerbation of back pain⁵³ while a final study reported the exclusion of two participants because of severe coronary artery disease.⁵⁷ None of the studies provided data on the costs associated with conducting the intervention.^{50–63}

Risk of bias assessment

Results for risk of bias assessment using the Cochrane Risk of Bias Assessment Instrument are shown in [figure 2](#) and online supplementary file 3. As can be seen, greater than 50% of the studies were at an unclear or high risk of bias with respect to: (1) incomplete outcome reporting (78.6%), (2) allocation concealment (78.6%), and (3) blinding of outcome assessors (57.1%). Given the inability to truly blind participants in exercise intervention trials, all studies (100%) were considered to be at a high risk of bias for the category ‘blinding of participants and personnel’.

Data synthesis

Overall results for primary outcome (anxiety)

Overall results for changes in anxiety for the 14 included studies^{50–63} are shown in [table 4](#) and [figure 3](#). As can be seen, statistically significant ($P=0.002$) reductions in anxiety were observed. In addition, statistically significant heterogeneity was observed while overall inconsistency was categorised as moderate (range=low to large). The NNT was 6 with a percentile improvement of 15.5% and an estimated 5.3million inactive US adults with AORD improving their anxiety if they started exercising regularly. Statistically significant small-study effects were observed ($P<0.0001$) ([figure 4](#)). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from -0.44 (95% CI -0.71 to -0.18) to -0.35 (95% CI -0.57 to -0.12). Cumulative meta-analysis, ranked by year, demonstrated that results have been statistically significant since the first study was conducted in 1989,⁵⁸ but with a trend towards smaller improvements in anxiety with each accumulating year ([figure 5](#)). Reductions

Table 1 Study characteristics

Study and year	Country	Participants	Exercise intervention	Anxiety assessment
Baptista <i>et al</i> , 2012 ⁵⁰	Brazil	Women (n=80) 18–65 years of age with FM assigned to exercise (n=40, age, \bar{X} =49.5 years) or control (n=40, age, \bar{X} =49.1 years) group	Dance exercise 2 times/week, 60 min/session, for 12 weeks	STAI
Beltran, 2003 ⁵¹	USA	Women (n=21) 22–65 years of age with FM assigned to exercise (n=11, age, \bar{X} ±SD=50.1±12.6 years) or control (n=10, age, \bar{X} ±SD=53.7±6.9 years) group	Aerobic aquatic exercise, 3 times/week, 25 min/session, 60%–85% MHR, for 10 weeks	AIMS
Cheung <i>et al</i> , 2017 ⁵²	USA	Men and women (n=51) >65 years of age with knee OA assigned to exercise (n=28, age, \bar{X} ±SD=74.4±7.5 years) or control (n=23, age, \bar{X} ±SD=71.8±8.0 years) group	Aerobic exercise, 5 times/week, 20–30 min/session for 8 weeks, and strength exercise 3 times/week, 30 min/session for 8 weeks (10+ exercises)	HADS
Fransen <i>et al</i> , 2007 ⁵³	Australia	Men and women (n=96) 59–85 years of age with hip or knee OA assigned to hydrotherapy (n=55, age, \bar{X} ±SD=70±6.3 years) or wait list control (n=41, age, \bar{X} ±SD=69.6±6.1 years) group	Hydrotherapy exercises in warm water, 2 times/week, 60 min/session, for 12 weeks	DASS21
French <i>et al</i> , 2013 ⁵⁴	Ireland	Men and women (n=88) 40–80 years of age with hip OA assigned to exercise (n=45, age, \bar{X} ±SD=61.8±9.7 years) or control (n=43, age, \bar{X} ±SD=60.8±9.7 years) group	Aerobic and strength training, 30 min/session, for 8 weeks, up to 5 strength exercises	HADS
Gowans <i>et al</i> , 2001 ⁵⁵	Canada	Men and women (n=57) with FM assigned to an exercise (n=30, age, \bar{X} ±SD=44.6±8.7 years) or control (n=27, age, \bar{X} ±SD=49.8±7.3 years) group	Aerobic exercise (2 walking/jogging classes in a gym, 1 pool class), 3 times/week, 30 min/session, 60%–75% MHR, for 23 weeks	MHI, STAI
Jones <i>et al</i> , 2008 ⁵⁶	USA	Men and women (n=101) 18–65 years of age with FM assigned to an exercise (n=47, age, \bar{X} ±SD=49.6±7.7 years) or control (n=54, age, \bar{X} ±SD=49.8±7.9 years) group	Aerobic exercise 3 times/week, 30 min/session, 40%–50% MHR and strength exercise, 3 times/week, 10 min/session, for 24 weeks	FIQ (Anxiety)
Komatireddy <i>et al</i> , 1997 ⁵⁷	USA	Men and women (n=49) 35–76 years of age with RA assigned to exercise (n=25, \bar{X} ±SD=57.7±9.8 years of age, range 40–72 years) or control (n=24, \bar{X} ±SD=60.5±11 years of age, range 35–76 years) group	Circuit weight training with light loads and high repetitions, 7 exercises, 2–3 circuits/session, 12–15 reps, 30 s rest between sets, 20–27 min/session, ≥3 times/week, RPE of 3–4, for 12 weeks	AIMS
Minor <i>et al</i> , 1989 ⁵⁸	USA	Men and women (n=115) 21–83 years of age with RA or OA assigned to a pool (n=47), walking (n=36) or control group (n=32)	Aerobic aquatics or walking group, 3 times/week, 60 min/session (30 min of this was aerobic), 60%–80% MHR, for 12 weeks	AIMS
Munguía-Izquierdo and Legaz-Arrese, 2008 ⁵⁹	Spain	Men and women (n=60) 18–60 years of age with FM assigned to exercise (n=35, \bar{X} ±SD=50.0±7.0 years of age) or control (n=25, \bar{X} ±SD=46.0±8.0 years of age) group	Aerobic aquatic exercise, 3 times/week, 20–30 min/session, 50%–80% MHR, and strength exercise 3 times/week, 8–20 min/session, 1–3 sets, 8–15 reps, for 16 weeks	STAI

Continued

Table 1 Continued

Study and year	Country	Participants	Exercise intervention	Anxiety assessment
Sañudo <i>et al</i> , 2015 ⁶⁰	Spain	Women (n=32) with FM assigned to exercise (n=16, $\bar{X}\pm SD=55.0\pm 8.0$ years of age) or control (n=16, $\bar{X}\pm SD=58.0\pm 6.9$ years of age) group	Aerobic/interval training, 2 times/week, 30–35 min/session, 60%–80% MHR, for 24 weeks	VAS
Schachter <i>et al</i> , 2003 ⁶¹	Canada	Women (n=143) 20–55 years of age with FM assigned to a short bout (n=56, $\bar{X}\pm SD=41.9\pm 8.6$ years of age), long bout (n=51, $\bar{X}\pm SD=41.3\pm 8.7$ years of age) or control group (n=36, $\bar{X}\pm SD=42.5\pm 6.7$ years of age)	Low-impact, videotape-based aerobic exercise to music. Short bout, 2 times/day, 7.1 times/week, 12.3 min/session, 60% HRR; long bout group: 1 times/day, 3.6 times/week, 24.5 min/session, 60% HRR, for 16 weeks	FIQ (Anxiety)
Tomas-Carus <i>et al</i> , 2008 ⁶²	Portugal	Women with FM (n=33) assigned to aquatic exercise (n=17, $\bar{X}\pm SD=50.7\pm 10.6$ years of age) or control group (n=16, $\bar{X}\pm SD=50.9\pm 6.7$ years of age)	Pool exercises performed in warm water, 3 times/week, 20 min, aerobic phase, 60%–65% MHR, strength exercise, 20 min, 4 sets, 10 reps, for 32 weeks	FIQ (Anxiety), STAI
Tomas-Carus <i>et al</i> , 2007 ⁶³	Portugal	Women with FM (n=34) assigned to aquatic exercise (n=17, $\bar{X}\pm SD=51\pm 10.0$ years of age) or control group (n=17, $\bar{X}\pm SD=51\pm 9.0$ years of age)	Pool exercises performed in warm water, 3 times/week, 20 min, aerobic phase, 60%–65% MHR, strength exercises, 20 min, 4 sets, 10 reps, for 12 weeks	FIQ (Anxiety)

Description of groups from each study limited to those that met the criteria for inclusion.

$\bar{X}\pm SD$, mean $\pm SD$; AIMS, Arthritis Impact and Measurement Scale; DASS21, Depression, Anxiety and Stress Scale; FIQ, Fibromyalgia Impact Questionnaire; FM, fibromyalgia; HADS, Hospital Anxiety and Depression Scale; HRR, heart rate reserve; MHI, Mental Health Inventory; MHR, maximum heart rate; OA, osteoarthritis; RA, rheumatoid arthritis; reps, repetitions; RPE, rating of perceived exertion; STAI, State-Trait Anxiety Inventory; VAS, visual analogue scale.

in anxiety were similar to group-level results when collapsed so that only one ES represented each study (ES, -0.40 , 95% CI -0.67 to -0.13 , $P=0.004$; $Q=39.6$, $P=0.0002$; $I^2=67.2\%$, 95% CI 42.6% to 81.3%, $\tau^2=0.15$). With six outliers deleted from the model, overall reductions in anxiety were similar but heterogeneity was no longer statistically significant and overall inconsistency was reduced to a level categorised as small (ES, -0.40 , 95% CI -0.62 to -0.18 , $P=0.0004$; $Q=11.8$, $P=0.22$; $I^2=24.0\%$, 95% CI 0% to 63.1%, $\tau^2=0.03$).

Metaregression results for anxiety

Simple metaregression results are shown in online supplementary file 4. For study characteristics, greater reductions in anxiety were associated with: (1) earlier publication year, (2) studies at an unclear versus low risk of bias, (3) studies in which sample size estimates were

not provided, (4) trials in which a larger percentage of participants initially agreed to participate in, and (5) studies that were not funded versus funded. For participant characteristics, greater improvements were associated with a larger per cent dropout in the exercise groups as well as younger age. For exercise intervention characteristics, greater reductions were associated with: (1) aerobic/strength training versus aerobic and strength training combined, (2) fewer minutes of exercise per session, (3) minutes of exercise per week, and (4) total minutes of exercise for the entire intervention period. Greater reductions were also associated with: (1) supervised and unsupervised exercises versus both, (2) facility and home-based exercise versus both, and (3) group and self-exercise versus both. No statistically significant

Table 2 Baseline characteristics of participants

Variable	Exercise				Control			
	Groups/ participants	$\bar{X}\pm SD$	Median	Range	Groups/ participants	$\bar{X}\pm SD$	Median	Range
Age (years)	14/458	53.4 \pm 9.7	50	41–74	13/349	54.9 \pm 8.8	50	43–72
BMI (kg·m ²)	7/204	29.0 \pm 1.5	29	27–31	7/170	28.4 \pm 1.7	28	27–31
VO _{2max} (mL/kg/min)	7/245	21.3 \pm 2.3	21	19–24	5/101	20.7 \pm 2.7	21	17–24

Groups represent number of exercise and control groups reporting data.

$\bar{X}\pm SD$, mean $\pm SD$; BMI, body mass index; VO_{2max}, maximum oxygen consumption.

Table 3 Exercise programme characteristics

Variable	Groups/participants	$\bar{X}\pm SD$	Median	Range
Length (weeks)	16/514	16 \pm 7	14	8–32
Frequency (times/week)	14/469	3.3 \pm 1.3	3	2–7
Duration (min/session)	14/450	28.8 \pm 14.3	30	10–60
Min/week	11/407	85.5 \pm 21.2	88	60–120
Min/week (adj)	7/277	61.3 \pm 22.7	58	33–99
Compliance (%)	8/312	74.3 \pm 19	80	38–97

Groups represent number of exercise groups reporting data. $\bar{X}\pm SD$, mean \pm SD; adj, adjusted for compliance.

association ($P=0.71$) was observed between rheumatoid/osteoarthritis ($ES=-0.35$) and fibromyalgia ($ES=-0.44$) while the overall ES for those studies in which the participants had rheumatoid arthritis was -0.54 and the overall ES for those in which the majority had osteoarthritis was -0.32 .

Results for the final multiple regression model are shown in table 5. The overall model was statistically significant ($F=13.4$, $P=0.004$). Earlier year of publication and decreases in depression were statistically significant predictors for greater reductions in anxiety. However, minutes of exercise per week were no longer statistically significant.

GRADE findings for changes in anxiety

An evidence profile for changes in anxiety is shown in online supplementary file 5. As can be seen, the outcome (anxiety) was considered critical and the overall strength of the finding was considered high, with future additional studies unlikely to have an effect on the overall direction of findings.

Results for secondary outcomes

Overall results for changes in secondary outcomes are shown in table 4. *Physical function* was assessed in 10 studies^{50 52–55 57 58 61–63} using the 10 m walk test, 50-foot walk, 6min walk, Arthritis Impact and Measurement

Scale (AIMS), Fibromyalgia Impact Questionnaire (FIQ), Short Physical Performance Battery, sit-to-stand test, stair-climbing test, up and go test, and the Western Ontario and McMaster Osteoarthritis Index (WOMAC). As can be seen, statistically significant ($P<0.001$) improvements in *physical function* were observed. Statistically significant heterogeneity was observed while inconsistency was categorised as moderate (range=low to large). The NNT was 4 with a percentile improvement of 24.5% and an estimated 9 million inactive US adults with AORD improving their physical function if they started exercising regularly. No statistically significant small-study effects were observed ($P=0.17$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from 0.54 (95% CI 0.30 to 0.78) to 0.73 (95% CI 0.39 to 1.06). Cumulative meta-analysis, ranked by year, demonstrated that results have been statistically significant since the first study was conducted in 1989,⁵⁸ but with a trend towards smaller improvements in physical function with each accumulating year (from 0.86 in 1989 to 0.66 in 2016). Improvements in physical function were similar to group-level results when collapsed so that only one ES represented each study (ES 0.66, 95% CI 0.31 to 1.01, $P=0.0002$; $Q=35.6$, $P<0.001$; $I^2=74.7\%$, 95% CI 52.8% to 86.5%, $\tau^2=0.21$). With four outliers deleted from the model, overall improvements in physical function were

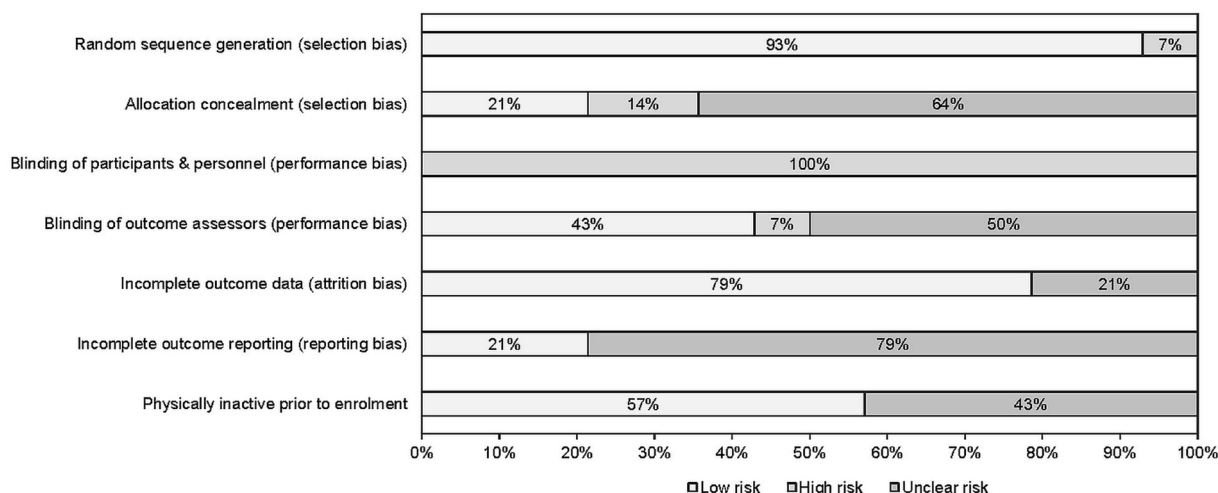


Figure 2 Risk of bias results using the Cochrane Risk of Bias Assessment Instrument.

Table 4 Results for primary and secondary outcomes (data reported as standardised effect size unless otherwise noted)

Variable	ES (n)	Participants (n)	\bar{X} (95% CI)	Q (P)	I^2 (95% CI)	τ^2
Primary outcome						
Anxiety	16	883	-0.40 (-0.65 to 0.15)*	40.3 (<0.001)**	62.8 (36.2 to 78.3)	0.14
Secondary outcomes						
Physical function	12	677	0.66 (0.34 to 0.97)*	36.0 (0.0002)**	69.4 (44.5 to 83.1)	0.19
Pain	15	803	-0.62 (-1.12 to 0.11)*	128.6 (<0.001)**	89.1 (83.7 to 92.7)	0.75
QOL	13	730	0.63 (0.35 to 0.91)*	32.4 (0.001)**	63.0 (32.7 to 79.7)	0.15
Depression	15	813	-0.38 (-0.67 to 0.10)*	46.3 (<0.001)**	69.7 (48.6 to 82.2)	0.20
VO _{2max} (mL/kg/min)	7	346	2.01 (0.85 to 3.2)*	20.2 (0.003)**	70.3 (35.0 to 86.4)	1.40
Muscular strength	6	261	0.59 (0.33 to 0.85)*	3.9 (0.6)	0 (0 to 67.1)	0

*Statistically significant (two-tailed alpha value ≤ 0.05 and non-overlapping 95% CI).

**Statistically significant (alpha value ≤ 0.10).

\bar{X} , mean effect size; ES, effect size; Q (P), Cochran Q statistic and alpha value for Q; QOL, quality of life; VO_{2max}, maximum oxygen consumption.

statistically significant but slightly smaller, heterogeneity was no longer statistically significant and overall inconsistency was reduced to a level categorised as very low (ES 0.57, 95% CI 0.36 to 0.78, $P < 0.001$; $Q = 7.2$, $P = 0.41$; $I^2 = 2.7\%$, 95% CI 0% to 68.5%, $\tau^2 = 0.003$).

For *pain*, assessment was conducted in 13 studies^{50-54 56-63} using the AIMS, FIQ, numerical rating scale, tender point count, visual analogue scale (VAS) and the WOMAC. As can be seen in table 4, statistically significant ($P = 0.017$) decreases in pain were found. Statistically significant heterogeneity was observed while inconsistency was

categorised as large, including both CIs. The NNT was 6 with a percentile improvement of 23.1% and an estimated 5.6 million inactive US adults with AORD decreasing their pain if they started exercising regularly. No statistically significant small-study effects were observed ($P = 0.34$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from -0.70 (95% CI -1.21 to -0.19) to -0.47 (95% CI -0.80 to -0.15). Cumulative meta-analysis, ranked by year, demonstrated that results have been statistically significant, and remained statistically

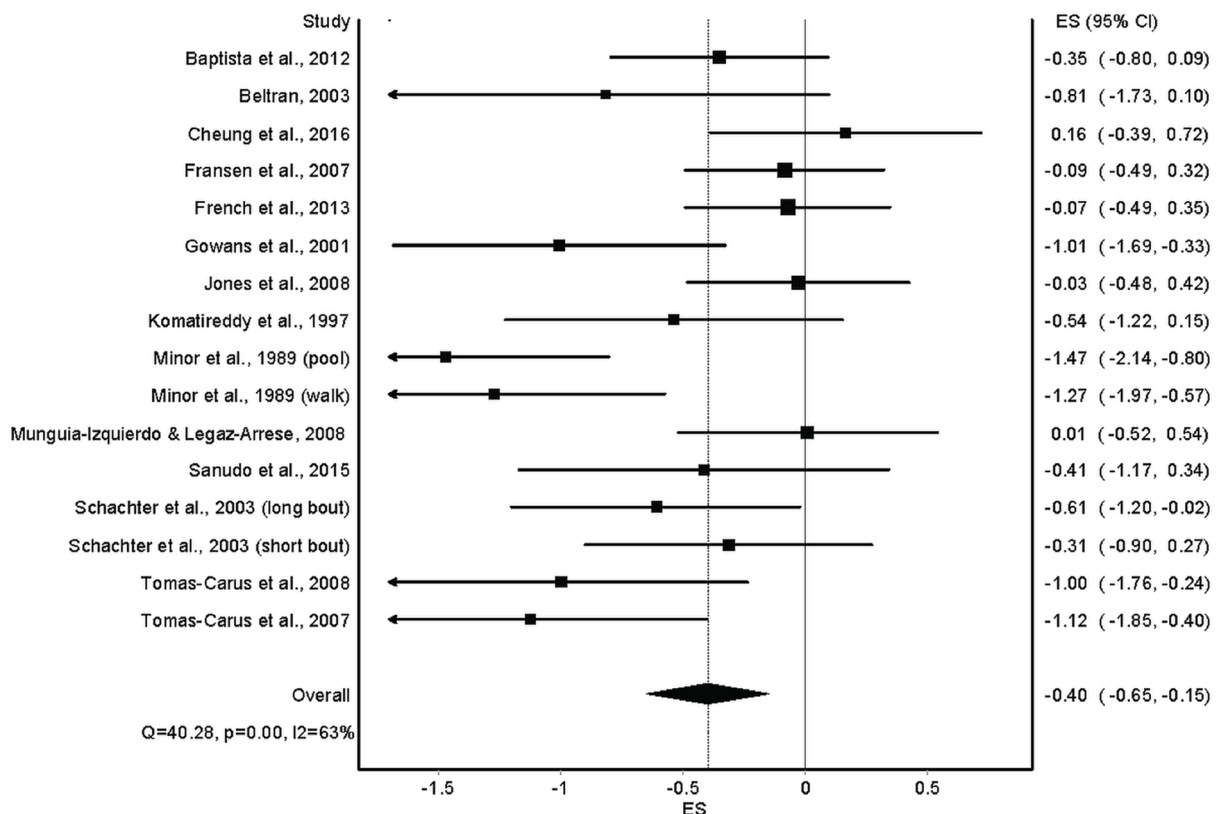


Figure 3 Forest plot for changes in anxiety. ES, effect size.

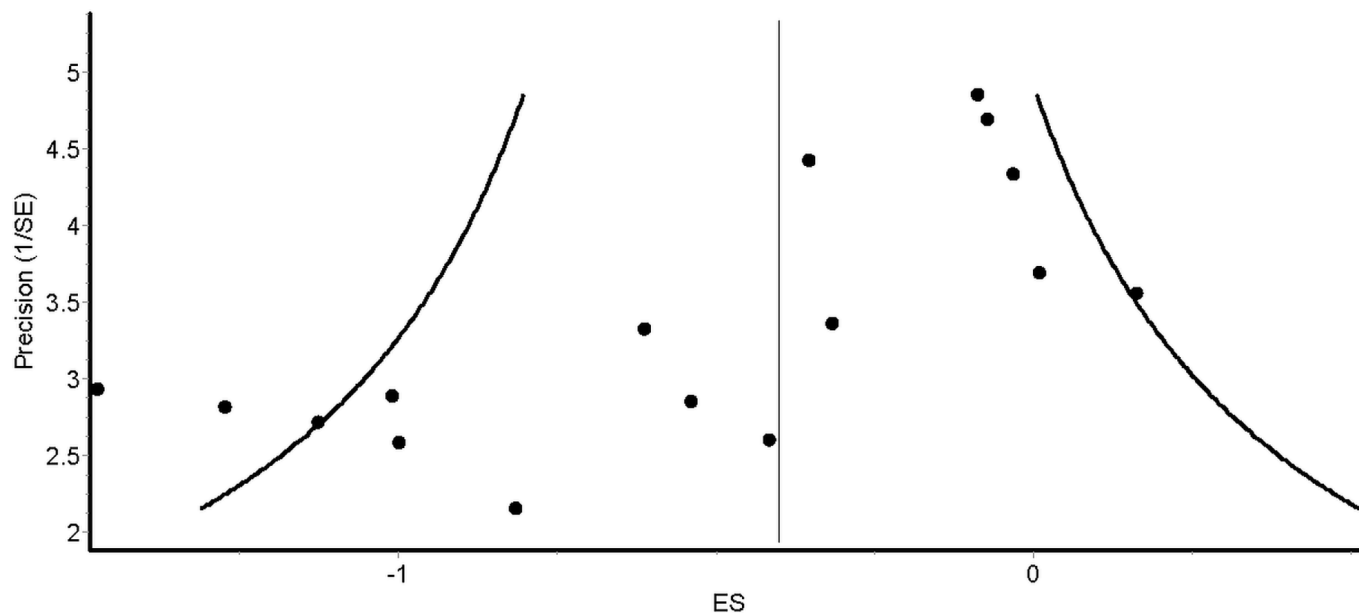


Figure 4 Funnel plot for changes in anxiety. ES, effect size.

significant, since only 2013. Decreases in pain were similar to group-level results when collapsed so that only one ES represented each study (ES -0.62 , 95% CI -1.16 to -0.07 , $P=0.003$; $Q=123.7$, $P<0.001$; $I^2=90.3\%$, 95% CI 85.3% to 93.6%, $\tau^2=0.75$). With 10 outliers deleted from the model, decreases in pain remained statistically significant but smaller, heterogeneity was no longer statistically significant and overall inconsistency was reduced to a level categorised as very low (ES -0.44 , 95% CI -0.70 to -0.18 , $P=0.001$; $Q=1.5$, $P=0.68$; $I^2=0\%$, 95% CI 0% to 69.5%, $\tau^2=0$).

Depression was assessed in 13 studies^{50-58 60-63} using the AIMS, Beck Depression Inventory, Center for Epidemiologic Studies Depression Scale, Depression, Anxiety and Stress Scale-21, FIQ, Hospital Anxiety and Depression Scale, Mental Health Inventory and VAS. As can be seen in [table 4](#), statistically significant ($P=0.009$) decreases in depression were found. Statistically significant

heterogeneity was observed while inconsistency was categorised as moderate (95% CI=low to large). The NNT was 6 with a percentile improvement of 15% and an estimated 5.1 million inactive US adults with AORD reducing their depression if they started exercising regularly. Small-study effects were not statistically significant ($P=0.08$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from -0.52 (95% CI -0.68 to -0.37) to -0.32 (95% CI -0.61 to -0.03). Cumulative meta-analysis, ranked by year, demonstrated that improvements have been statistically significant and remained stable since 2013. Decreases in depression were similar to group-level results when collapsed so that only one ES represented each study (ES -0.38 , 95% CI -0.70 to -0.07 , $P=0.02$; $Q=45.6$, $P<0.001$; $I^2=73.7\%$, 95% CI 54.3% to 84.8%, $\tau^2=0.21$). With three outliers deleted from the model, improvements in depression remained statistically significant, slightly

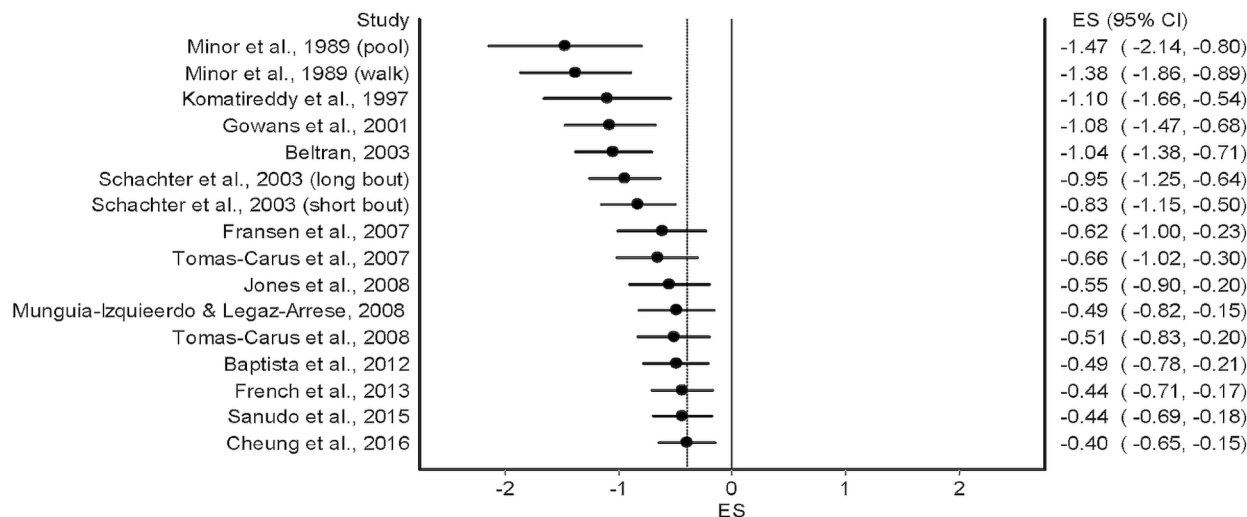


Figure 5 Cumulative meta-analysis for changes in anxiety. ES, effect size.

Table 5 Final multiple regression model for changes in anxiety (n=16)

Variable	Coefficient±SE	t (P)	95% CI
Year of publication	0.035±0.010	3.50 (0.006)*	0.012 to 0.057
Changes in depression	0.316±0.758	4.16 (0.01)*	0.110 to 0.521
Minutes of training per week	0.007±0.003	2.40 (0.06)**	-0.0002 to 0.016
Intercept	-70.718±19.74	-3.58 (0.006)*	-115.19 to -26.24

P, alpha value for t.

*statistically significant (two-tailed alpha value ≤ 0.05 and non-overlapping 95% CI).

**trend for statistical significance ($p \leq 0.10$).

larger, but with non-significant heterogeneity and overall inconsistency categorised as very low (ES -0.43, 95% CI -0.60 to -0.26, $P < 0.001$; $Q = 5.7$, $P = 0.89$; $I^2 = 0\%$, 95% CI 0% to 19.3%, $\tau^2 = 0$).

Quality of life was assessed in 12 studies^{50-57 59 61-63} using the FIQ, Health Assessment Questionnaire, Quality of Life Scale, Short Form (SF)-12 and SF-36. As shown in table 4, statistically significant ($P < 0.0001$) improvements in quality of life were found. Statistically significant heterogeneity was observed while inconsistency was categorised as large (95% CI=low to large). The NNT was 4 with a percentile improvement of 23.1% and an estimated 8.6 million inactive US adults with AORD improving their quality of life if they started exercising regularly. Statistically significant small-study effects were observed ($P < 0.001$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from 0.58 (95% CI 0.35 to 0.81) to 0.70 (95% CI 0.41 to 0.99). Cumulative meta-analysis, ranked by year, demonstrated that improvements have been statistically significant since 2001, but with a trend towards a decrease in the magnitude of effect. Increases in quality of life were similar to group-level results when collapsed so that only one ES represented each study (ES 0.63, 95% CI 0.34 to 0.92, $P < 0.001$; $Q = 30.9$, $P = 0.001$; $I^2 = 64.3\%$, 95% CI 33.9% to 80.8%, $\tau^2 = 0.14$). With four outliers deleted from the model, improvements in quality of life remained statistically significant, similar in magnitude, but with non-significant heterogeneity and overall inconsistency categorised as low (ES 0.64, 95% CI 0.39 to 0.90, $P < 0.001$; $Q = 11.4$, $P = 0.18$; $I^2 = 29.8\%$, 95% CI 0% to 67.5%, $\tau^2 = 0.04$).

VO_{2max} in mL/kg/min was assessed in five studies^{56-58 61 62} using various maximal treadmill tests while one study used a submaximal step test (Canadian Aerobic Fitness Test). As can be seen in table 4, statistically significant ($P = 0.001$) improvements in VO_{2max} were observed. Statistically significant heterogeneity was observed while inconsistency was categorised as moderate (95% CI=low to large). The NNT was 3 with a percentile improvement of 25.7% and an estimated 9.5 million inactive US adults with AORD improving their relative VO_{2max} if they started exercising regularly. No statistically significant small-study effects were observed ($P = 0.18$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from 1.80 mL/kg/min

(95% CI 0.58 to 3.01) to 2.47 mL/kg/min (95% CI 1.25 to 3.69). Cumulative meta-analysis, ranked by year, demonstrated that improvements have been statistically significant since 2003, but with a trend towards a decrease in the magnitude of effect over time. Increases in VO_{2max} were similar to group-level results when collapsed so that only one ES represented each study (ES 2.01 mL/kg/min, 95% CI 0.59 to 3.44, $P = 0.01$; $Q = 19.4$, $P = 0.001$; $I^2 = 79.4\%$, 95% CI 51.2% to 91.3%, $\tau^2 = 1.7$). With outliers deleted from the model, improvements in VO_{2max} were smaller, with non-significant heterogeneity and overall inconsistency categorised as low (ES 1.93 mL/kg/min, 95% CI 0.83 to 3.02, $P = 0.001$; $Q = 0.03$, $P = 0.86$; $I^2 = 0\%$, 95% CI 0% to 0%, $\tau^2 = 0$).

Upper and lower body muscular strength was assessed in five studies^{55 57-59 62} using free weights, grip strength, isokinetic strength and the sit-to-stand test with free weights. As shown in table 4, statistically significant ($P < 0.001$) improvements in strength were observed. No statistically significant heterogeneity or mean inconsistency was observed (95% CI=none to moderate). The NNT was 4 with a percentile improvement of 22.2% and an estimated 8 million inactive US adults with AORD improving their strength if they started exercising regularly. No statistically significant small-study effects were observed ($P = 0.65$). With each result deleted from the model once, results remained statistically significant across all deletions, ranging from 0.50 (95% CI 0.21 to 0.79) to 0.68 (95% CI 0.39 to 0.96). Cumulative meta-analysis, ranked by year, demonstrated that improvements in strength have been statistically significant since the first included study in 1989.⁵⁸ Changes in strength were similar to group-level results when collapsed so that only one ES represented each study (ES 0.59, 95% CI 0.33 to 0.85, $P < 0.001$; $Q = 3.5$, $P = 0.48$; $I^2 = 0\%$, 95% CI 0% to 76.2%, $\tau^2 = 0$). There were no outliers.

Insufficient data were available to analyse BMI. The one study that did provide change outcome results for BMI reported no statistically significant changes.⁵⁶

DISCUSSION

Overall findings

The primary purpose of the current systematic review with meta-analysis was to examine the effects of exercise (aerobic, strength training or both) on anxiety in adults

with AORD. The overall findings, a primary purpose of meta-analysis,⁶⁵ suggest that exercise is associated with both statistically significant and practically important reductions in anxiety among adults with selected types of AORD. These findings are supported by: (1) a magnitude of effect comparable to or greater than anxiolytics,⁶⁶ (2) a large percentile improvement of 15.5, (3) a NNT of only 6, (4) a large number of physically inactive US adults with AORD who could benefit from exercising regularly (5.3 million), (5) similar findings when examined at the study versus group level, (6) continued existence of a statistically significant effect when each result was deleted from the analysis once, (7) similar findings when outliers were deleted and statistical heterogeneity was reduced to a non-significant effect and overall inconsistency to 0, and (8) a consistent finding of improvements in anxiety since the first included study was reported in 1989.⁵⁸ Based on GRADE, it was concluded that anxiety was a critical outcome and that further research would unlikely change the direction of effect.

While the current findings are encouraging, it is interesting to note that cumulative meta-analysis revealed a distinct trend for a reduction in the pooled ES over time (from -1.47 in 1989 to -0.40 in 2016) and based on metaregression, a statistically significant association between greater reductions in anxiety with older versus more recent studies. While the specific reasons for this could not be determined, it may be that the experimental design and conduct of studies have improved over time. However, from the investigative team's perspective, it is highly unlikely that the results will become non-significant in future years. The former notwithstanding, the general conclusion that exercise is associated with reductions in anxiety among adults with selected types of AORD may need to be viewed with some caution given that the majority of included studies consisted of participants with fibromyalgia^{50 51 55 56 59-63} while the remaining studies included those with osteoarthritis and/or rheumatoid arthritis.^{52-54 57 58} However, metaregression revealed no statistically significant association between type of AORD and changes in anxiety, thus allowing for the pooling of findings and a resultant increase in statistical power.

The final metaregression model resulted in two statistically significant variables being included with earlier year of publication and changes in depression associated with greater reductions in anxiety. These findings further reinforce the influence of year of publication on changes in anxiety in the current systematic review with meta-analysis, a potential reason for such having been previously mentioned. The greater reductions in anxiety associated with reductions in depression and quality of life based on simple metaregression as well as depression in the multiple regression model illustrate the potential interaction between these factors and the difficulty in identifying such, especially in a systematic review with meta-analysis. Finally, for those studies in which data were available,⁵⁰⁻⁶³ it is worthy to note that while the focus was on community-deliverable exercise or exercise delivered

in other settings in which the exercise intervention could be adapted for delivery in the community, two were delivered in either a university^{50 56} or hospital^{54 55} and one each in either a hospital and home,⁵³ university and home,⁵⁷ Young Mens' Christian Association (YMCA),⁵¹ or home only.⁶¹

In addition to statistically significant and practically important improvements in anxiety, similar improvements were also observed for all secondary outcomes assessed (physical function, pain, depression, quality of life, VO_{2max} in mL/kg/min, muscular strength). These findings are important because unlike pharmacological interventions that are usually targeted to address one condition, exercise has the potential to improve multiple physiological and psychological outcomes. Given the former, it would seem plausible to suggest that continued efforts be made to increase the exercise and physical activity levels of adults with AORD.

Implications for research

There are at least eight implications for the conduct and reporting of future research on exercise and anxiety in adults with AORD. First, since only five of the studies included participants with osteoarthritis and/or rheumatoid arthritis, future randomised controlled trials on exercise and anxiety in these populations appear warranted. Second, exercise was performed indoors in the majority of included studies. Given that previous research has suggested the exercise performed outdoors may have better mood-enhancing effects than indoor exercise,^{67 68} future research examining this phenomenon with respect to anxiety as well as other outcomes in adults with AORD seems appropriate. Third, future studies should focus on examining the dose-response effects of exercise on anxiety in adults with AORD and report complete information on the characteristics of the intervention, including intervention fidelity.⁶⁹ Such information is critical for the development of evidence-based recommendations aimed at practitioners. Fourth, since more than half of the studies were considered to be at high or unclear risk of bias with respect to allocation concealment, blinding of outcome assessors and incomplete outcome reporting, future studies should address these issues in their experimental design and report such information. Notably, while all included studies were considered to be at a high risk of bias with respect to blinding of participants, this is difficult for researchers to address since unlike pharmacological studies, it is almost impossible to blind intervention participants to group assignment in exercise intervention studies. Fifth, given the lack of data on exposure to psychotropic and analgesic medications before and during the studies, including any changes in medication during the study, it is suggested that future trials track and report this information, including pharmacological equivalence between the intervention and control groups. Sixth, while a lack of data was reported on how and what level anxiety was assessed, all the studies used instruments apparently focused on generalised anxiety.

However, anxiety is a heterogeneous construct that includes generalised anxiety disorder, separation anxiety disorder, selective mutism, specific phobia, social phobia, panic disorder and agoraphobia.⁷⁰ Given the former, it is suggested that future studies report detailed information on the specific types of anxiety they are assessing so as to better identify what exact type(s) might be affected. Seventh, given previous meta-analytic research by others,⁷¹ a focus on group versus individualised exercise may be preferable because of increased contact and social support. This may be especially true for ethnic and racial minorities.⁷² Finally, and collectively, a study in which one might have clinical confidence in the results might consist of a randomised controlled trial that: (1) uses random sequence generation, for example, computer random generator, to assign participants to an exercise intervention and attention control group, (2) conceals allocation to group assignment using an approach such as sequentially numbered, sealed, opaque envelopes, (3) blinds personnel and outcome assessors to group assignment while noting the inherent risk that all exercise studies are traditionally unable to blind participants to group assignment, (4) accounts for incomplete data (dropouts, etc.) by conducting intention-to-treat analyses, (5) avoids selective reporting of data by providing results for both statistically significant and non-significant results, (6) includes sedentary participants with a selected type or types of arthritis and elevated baseline levels of anxiety, (7) includes valid and reliable instruments for the assessment of the different types of anxiety based on the most recent version of the Diagnostic and Statistical Manual of Mental Disorders,⁷⁰ (8) accounts for potential confounders (age, gender, race, ethnicity, education, income, pain, physical function, depression, comorbidities, baseline levels of anxiety, sleep, medication use before and during the study, other non-exercise and non-medication uptake before and during the study), (9) based on a theoretical model, conducts group exercise intervention sessions (aerobic, strength training or both) based on the recent Consensus on Exercise Reporting Template,⁷³ (10) examines the mechanisms for changes in anxiety, and (11) conducts cost-effectiveness analysis of the intervention. While these recommendations are focused on exercise, alternative treatments exist. Thus, the most clinically useful study might consist of a mega-randomised controlled trial that addresses all currently available treatments. However, this is probably not realistic. Therefore, an alternative approach might be to conduct a network meta-analysis that includes both direct and indirect evidence from randomised trials of all available interventions (exercise, pharmacologic, etc.).⁷⁴ To the best of the authors' knowledge, no such study currently exists.

Implications for practice

The results of this study appear to have important implications for practice. First, since changes in both anxiety and secondary outcomes resulted in statistically significant

and practically important improvements, exercise may be more vital than any other intervention given the apparent multiple benefits of such. In addition, the magnitude of effect (-0.40), NNT (6) and percentile improvement (15.5) observed in the current study were equivalent to a previous meta-analysis on the pharmacological treatment of generalised anxiety disorder in which the overall standardised mean difference ES was -0.39 , the calculated NNT was 6 and the percentile improvement was 15.2.⁶⁶ While the current systematic review with meta-analysis was unable to establish with any degree of certainty the dose-response effects of exercise on anxiety in adults with AORD, it would appear plausible, given the numerous other benefits that can be derived as well as the minimal adverse events associated with exercise, to adhere to the Active Adult or Active Older Adult Guidelines from the Physical Activity Guidelines for Americans Physical Activity Guidelines²⁵ as recommended by the Centers for Disease Control and Prevention.⁷⁵ Broadly, this includes aerobic activities on most, if not all, days of the week, muscle strengthening activities at least 2 days/week, balance exercises at least 3 days/week and flexibility exercises on a daily basis. However, these recommendations as applied to exercise and anxiety should be considered with respect to factors such as the lack of transparency with respect to risk of bias in the included studies.

Strengths and limitations

There are at least two apparent *strengths* of the current study. First, to the best of our knowledge, this is the first systematic review with meta-analysis to examine, as a primary outcome, the effects of exercise on anxiety in adults with AORD. This is important given: (1) the prevalence of AORD,¹ (2) the expected future increase in the prevalence of AORD,² (3) the financial costs associated with AORD,³ and (4) the previously reported finding that the prevalence of anxiety is almost twice that of depression.⁸ Second, a novel and recently developed approach, the IVhet model, was used to pool findings.³⁶ This resulted in more robust estimates than those derived from the traditional random effects model.³⁷ This is noteworthy given the need to provide the most accurate results possible when examining the effects of an intervention on selected outcome(s).

While there are several strengths to this study, there are also at least four potential *limitations*. First, similar to any aggregate data meta-analysis, the potential for ecological fallacy exists. Thus, it may be that the observed findings in the current study would not apply at the individual participant level. Second, and also common to any type of aggregate data meta-analysis, meta-regression results do not support causal inferences because the included studies are not randomly assigned to covariates.⁴⁷ Therefore, the associations observed in the current investigation would need to be assessed in appropriately powered randomised trials. Third, while a large number of statistical tests were conducted, no adjustments were made for such. Thus, some findings could have been nothing more than the

play of chance. However, common to most aggregate data meta-analyses, no adjustments for multiple testing were made because of concerns about missing possibly important findings that could be tested in original trials.⁴⁵ Fourth, since anxiety was assessed using self-report instruments,^{50–63} the possibility of reporting bias in the original studies existed.

CONCLUSIONS

Exercise is associated with reductions in anxiety among adults with selected types of AORD. However, a need exists for additional, well-designed, randomised controlled trials on this topic.

Contributors GAK was responsible for the conception and design, acquisition of data, analysis and interpretation of data, drafting the initial manuscript and revising it critically for important intellectual content. KSK was responsible for the conception and design, acquisition of data, drafting the initial manuscript and revising all drafts critically for important intellectual content. LFC was responsible for the conception and design, acquisition of data, drafting the initial manuscript and revising all drafts critically for important intellectual content. All authors read and approved the final manuscript.

Funding This meta-analysis was funded by the National Institutes of Health, National Institute for Arthritis, Musculoskeletal and Skin Diseases, grant number R01AR061346 (GAK, principal investigator).

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All data are available upon request from the corresponding author.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Barbour KE, Helmick CG, Boring M, *et al.* Vital signs: prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation - United States, 2013–2015. *MMWR Morb Mortal Wkly Rep* 2017;66:246–53.
- Hootman JM, Helmick CG, Barbour KE, *et al.* Updated projected prevalence of self-reported doctor-diagnosed arthritis and arthritis-attributable activity limitation among US adults, 2015–2040. *Arthritis Rheumatol* 2016;68:1582–7.
- Yelin E, Murphy L, Cisternas MG, *et al.* Medical care expenditures and earnings losses among persons with arthritis and other rheumatic conditions in 2003, and comparisons with 1997. *Arthritis Rheum* 2007;56:1397–407.
- Player MS, Peterson LE. Anxiety disorders, hypertension, and cardiovascular risk: a review. *Int J Psychiatry Med* 2011;41:365–77.
- Roest AM, Martens EJ, de Jonge P, *et al.* Anxiety and risk of incident coronary heart disease: a meta-analysis. *J Am Coll Cardiol* 2010;56:38–46.
- Janszky I, Ahnve S, Lundberg I, *et al.* Early-onset depression, anxiety, and risk of subsequent coronary heart disease: 37-year follow-up of 49,321 young Swedish men. *J Am Coll Cardiol* 2010;56:31–7.
- Blakemore A, Dickens C, Guthrie E, *et al.* Depression and anxiety predict health-related quality of life in chronic obstructive pulmonary disease: systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis* 2014;9:501–12.
- Murphy LB, Sacks JJ, Brady TJ, *et al.* Anxiety and depression among US adults with arthritis: prevalence and correlates. *Arthritis Care Res* 2012;64:968–76.
- He Y, Zhang M, Lin EH, *et al.* Mental disorders among persons with arthritis: results from the World Mental Health Surveys. *Psychol Med* 2008;38:1639–50.
- Busch AJ, Schachter CL, Overend TJ, *et al.* Exercise for fibromyalgia: a systematic review. *J Rheumatol* 2008;35:1130–44.
- Hurkmans E, van der Giesen FJ, Vliet Vlieland TP, *et al.* Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database Syst Rev* 2009;4:CD006853.
- Anderson E, Shivakumar G. Effects of exercise and physical activity on anxiety. *Front Psychiatry* 2013;4:27.
- Kelley GA, Kelley KS, Hootman JM. Effects of exercise on depression in adults with arthritis: a systematic review with meta-analysis of randomized controlled trials. *Arthritis Res Ther* 2015;17:21.
- Kelley GA, Kelley KS. Effects of exercise on anxiety in adults with arthritis and other rheumatic disease: a systematic review of meta-analyses. *J Nov Physiother* 2014;04:1–5.
- Bruno M, Cummins S, Audiano L, *et al.* Effectiveness of two arthritis foundation programs: walk with ease, and you can break the pain cycle. *Clin Interv Aging* 2006;1:295–306.
- Callahan LF, Ambrose KR. Physical activity and osteoarthritis - considerations at the population and clinical level. *Osteoarthritis Cartilage* 2015;23:31–3.
- Hootman JM, Helmick CG, Brady TJ. A public health approach to addressing arthritis in older adults: the most common cause of disability. *Am J Public Health* 2012;102:426–33.
- Zgibor JC, Ye L, Boudreau RM, *et al.* Community-based healthy aging interventions for older adults with arthritis and multimorbidity. *J Community Health* 2017;42:390–9.
- Jack K, McLean SM, Moffett JK, *et al.* Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. *Man Ther* 2010;15:220–8.
- Fontaine KR, Heo M, Bathon J. Are US adults with arthritis meeting public health recommendations for physical activity? *Arthritis Rheum* 2004;50:624–8.
- Liberati A, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009;151:W65–W94.
- Kelley GA, Kelley KS, Callahan LF. Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a protocol for a systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2017;7:e014957.
- Sacks H, Chalmers TC, Smith H. Randomized versus historical controls for clinical trials. *Am J Med* 1982;72:233–40.
- Schulz KF, Chalmers I, Hayes RJ, *et al.* Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
- Physical Activity Guidelines Advisory Committee. *Physical activity guidelines advisory report*. Washington, DC: US: Department of Health and Human Services, 2008.
- Brady TJ, Jernick SL, Hootman JM, *et al.* Public health interventions for arthritis: expanding the toolbox of evidence-based interventions. *J Womens Health* 2009;18:1905–17.
- Ide MR, Laurindo IMM, RODRIGUES-JUNIOR AL, *et al.* Effect of aquatic respiratory exercise-based program in patients with fibromyalgia. *Int J Rheum Dis* 2008;11:131–40.
- Waggoner CD, LeLievre RB. A method to increase compliance to exercise regimens in rheumatoid arthritis patients. *J Behav Med* 1981;4:191–201.
- Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485–514.
- Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–20.
- Lee E, Dobbins M, Decorby K, *et al.* An optimal search filter for retrieving systematic reviews and meta-analyses. *BMC Med Res Methodol* 2012;12:51.
- Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions version 5.1.0*. London, UK: The Cochrane Collaboration, 2011.
- Ahn S, Becker BJ. Incorporating quality scores in meta-analysis. *J Educ Behav Stat* 2011;36:555–85.
- Hedges LV, Olkin I. *Statistical methods for meta-analysis*. San Diego, CA: Academic Press, 1985.

35. Follmann D, Elliott P, Suh I, *et al.* Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45:769–73.
36. Doi SA, Barendregt JJ, Khan S, *et al.* Advances in the meta-analysis of heterogeneous clinical trials I: The inverse variance heterogeneity model. *Contemp Clin Trials* 2015;45:130–8.
37. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
38. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials* 2015;45(Pt A):139–45.
39. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101–29.
40. Higgins JP, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
41. Cohen J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1988.
42. Khan A, Khan S, Brown WA. Are placebo controls necessary to test new antidepressants and anxiolytics? *Int J Neuropsychopharmacol* 2002;5:193–7.
43. Murphy LB, Hootman JM, Boring MA, *et al.* Leisure time physical activity among U.S. adults with arthritis, 2008–2015. *Am J Prev Med* 2017;53:345–54.
44. Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol* 2010;171:624–32.
45. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–6.
46. Sacks HS, Berrier J, Reitman D, *et al.* Meta-analyses of randomized controlled trials. *N Engl J Med* 1987;316:450–5.
47. Littell JH, Corcoran J, Pillai V. *Systematic reviews and meta-analysis*. New York: Oxford University Press, 2008:1–202.
48. Sterne JA, Sutton AJ, Ioannidis JP, *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
49. Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
50. Baptista AS, Villela AL, Jones A, *et al.* Effectiveness of dance in patients with fibromyalgia: a randomized, single-blind, controlled study. *Clin Exp Rheumatol* 2012;30(6 Suppl 74):18–23.
51. Beltran R. *The effects of a supervised group aerobic exercise program and a chronobiological oriented treatment protocol on symptomatology and mood in women with fibromyalgia [Dissertation]*. San Diego, California: Alliant University, 2003.
52. Cheung C, Wyman JF, Bronas U, *et al.* Managing knee osteoarthritis with yoga or aerobic/strengthening exercise programs in older adults: a pilot randomized controlled trial. *Rheumatol Int* 2017;37:389–98.
53. 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;57:407–14.
54. French HP, Cusack T, Brennan A, *et al.* Exercise and manual physiotherapy arthritis research trial (EMPART) for osteoarthritis of the hip: a multicenter randomized controlled trial. *Arch Phys Med Rehabil* 2013;94:302–14.
55. Gowans SE, deHueck A, Voss S, *et al.* Effect of a randomized, controlled trial of exercise on mood and physical function in individuals with fibromyalgia. *Arthritis Rheum* 2001;45:519–29.
56. Jones KD, Burckhardt CS, Deodhar AA, *et al.* A six-month randomized controlled trial of exercise and pyridostigmine in the treatment of fibromyalgia. *Arthritis Rheum* 2008;58:612–22.
57. Komatireddy GR, Leitch RW, Cella K, *et al.* Efficacy of low load resistive muscle training in patients with rheumatoid arthritis functional class II and III. *J Rheumatol* 1997;24:1531–9.
58. Minor MA, Hewett JE, Webel RR, *et al.* Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 1989;32:1396–405.
59. Munguía-Izquierdo D, Legaz-Arrese A. Assessment of the effects of aquatic therapy on global symptomatology in patients with fibromyalgia syndrome: a randomized controlled trial. *Arch Phys Med Rehabil* 2008;89:2250–7.
60. Sañudo B, Carrasco L, de Hoyo M, *et al.* Vagal modulation and symptomatology following a 6-month aerobic exercise program for women with fibromyalgia. *Clin Exp Rheumatol* 2015;33(1 Suppl 88):S41–5.
61. Schachter CL, Busch AJ, Peloso PM, *et al.* Effects of short versus long bouts of aerobic exercise in sedentary women with fibromyalgia: a randomized controlled trial. *Phys Ther* 2003;83:340–58.
62. Tomas-Carus P, Gusi N, Häkkinen A, *et al.* Eight months of physical training in warm water improves physical and mental health in women with fibromyalgia: a randomized controlled trial. *J Rehabil Med* 2008;40:248–52.
63. Tomas-Carus P, Gusi N, Leal A, *et al.* [The fibromyalgia treatment with physical exercise in warm water reduces the impact of the disease on female patients' physical and mental health]. *Rheumatol Clin* 2007;3:33–7.
64. Garber CE, Blissmer B, Deschenes MR, *et al.* American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc* 2011;43:1334–59.
65. Glass GV, McGaw B, Smith ML. *Meta-analysis in social research*. Newbury Park, California: Sage, 1981.
66. Hidalgo RB, Tupler LA, Davidson JR. An effect-size analysis of pharmacologic treatments for generalized anxiety disorder. *J Psychopharmacol* 2007;21:864–72.
67. Rogerson M, Gladwell VF, Gallagher DJ, *et al.* Influences of green outdoors versus indoors environmental settings on psychological and social outcomes of controlled exercise. *Int J Environ Res Public Health* 2016;13:363.
68. Pasanen TP, Tyrväinen L, Korpela KM. The relationship between perceived health and physical activity indoors, outdoors in built environments, and outdoors in nature. *Appl Psychol Health Well Being* 2014;6:324–46.
69. Resnick B, Inguito P, Orwig D, *et al.* Treatment fidelity in behavior change research: a case example. *Nurs Res* 2005;54:139–43.
70. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th edn. Missouri, USA: American Psychiatric Association, 2013.
71. Burke SM, Carron AV, Eys MA, *et al.* Group versus individual approach? A meta-analysis of the effectiveness of interventions to promote physical activity. *Sport Exerc Psychol Rev* 2006;2:19–55.
72. Saint Onge JM, Krueger PM. Education and racial-ethnic differences in types of exercise in the United States. *J Health Soc Behav* 2011;52:197–211.
73. Slade SC, Dionne CE, Underwood M, *et al.* Consensus on Exercise Reporting Template (CERT): modified delphi study. *Phys Ther* 2016;96:1514–24.
74. Rouse B, Chaimani A, Li T. Network meta-analysis: an introduction for clinicians. *Intern Emerg Med* 2017;12:103–11.
75. CDC. Physical activity for arthritis Atlanta, georgia: centers for disease control and prevention;. 2017 <https://www.cdc.gov/arthritis/basics/physical-activity-overview.html> (accessed 5 July 2017).