Journal Pre-proof

Effectiveness of exercise on fatigue and sleep quality in fibromyalgia: a systematic review and meta-analysis of randomised trials

Fernando Estévez-López, Cristina Maestre-Cascales, Deborrah Russell, Inmaculada C. Álvarez-Gallardo, María Rodriguez-Ayllon, Ciara M. Hughes, Gareth W. Davison, Borja Sañudo, Joseph G. McVeigh

PII: S0003-9993(20)30434-2

DOI: https://doi.org/10.1016/j.apmr.2020.06.019

Reference: YAPMR 57905

To appear in: *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*

Received Date: 28 March 2020

Revised Date: 26 June 2020

Accepted Date: 29 June 2020

Effectiveness of exercise on fatigue and sleep quality in fibromyalgia: a systematic review and meta-analysis of randomised trials

Running head: Exercise, fatigue and sleep in fibromyalgia

Fernando Estévez-López [1,*], Cristina Maestre-Cascales [2,*], Deborrah Russell [3], Inmaculada C. Álvarez-Gallardo [4], María Rodriguez-Ayllon [5], Ciara M. Hughes [3], Gareth W. Davison [6], Borja Sañudo [7], Joseph G. McVeigh [8]

[1] Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC University Medical Center, Rotterdam, the Netherlands; [2] LFE Research Group, Department of Health and Human Performance, Universidad Politécnica de Madrid, Madrid, Spain; [3] Institute of Nursing and Health Research, School of Health Sciences, Ulster University, Northern Ireland, UK; [4] Department of Physical Education, Faculty of Education Sciences, University of Cádiz, Cádiz, Spain; [5] PROmoting FITness and Health Through Physical Activity (PROFITH) Research Group, Department of Physical Education and Sport, Faculty of Sport Sciences, University of Granada, Granada, Spain; [6] Sport and Exercise Science Research Institute, Ulster University, Northern Ireland, UK; [7] Department of Physical Education and Sport, University of Seville, Seville, Spain; [8] Physiotherapy, School of Clinical Therapies, College of Medicine and Health, University College Cork, Cork, Ireland.

* Equally contribution.

Correspondence to Fernando Estévez-López, PhD, Postdoctoral Researcher, Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Tel: (+31) 30 253 4700; E-mail: fer@estevez-lopez.com

Effectiveness of exercise on fatigue and sleep quality in fibromyalgia: a systematic review and meta-analysis of randomised trials

Abstract

Objectives: To determine the effects of exercise on fatigue and sleep quality in fibromyalgia (primary aim) and to identify which type of exercise is the most effective in achieving these outcomes (secondary aim).

Data sources: PubMed and Web of Science were searched from inception until October 18th, 2018.

Study selection: Eligible studies contained information on population (fibromyalgia), intervention (exercise) and outcomes (fatigue or sleep). Randomised controlled trials (RCTs) testing the effectiveness of exercise in comparison to usual care and randomised trials (RTs) comparing the effectiveness of two different exercise interventions were included for the primary and secondary aims of the present review, respectively. Two independent researchers performed the search, screening and final eligibility of the articles. From 696 identified studies, 17 RCTs (n=1003) were included for fatigue and 12 RCTs (n=731) for sleep. Furthermore, 21 RTs compared the effectiveness of different exercise interventions (n=1254).

Data extraction: Two independent researchers extracted the key information from each eligible study.

Data synthesis: Separate random-effect meta-analyses were performed to examine the effects from RCTs and from RTs (primary and secondary aims). Standardised mean differences (SMD) effect sizes were calculated using Hedges' adjusted g. Effect sizes of 0.2, 0.4 and 0.8 were considered small, moderate and large. In comparison to usual care, exercise had moderate effects on fatigue and a small effect on sleep quality; SMD (95% confidence interval) = -0.47 (-0.67 to -0.27, $p < 0.001$) and -0.17 (-0.32 to -0.01, $p = 0.04$). RTs in which fatigue was the primary outcome were the most beneficial for lowering fatigue. Additionally, meditative exercise programs were the most effective for improving sleep quality.

Conclusions: Exercise is moderately effective for lowering fatigue and has small effects on enhancing sleep quality in fibromyalgia. Meditative exercise programs may be considered for improving sleep quality in fibromyalgia.

PROSPERO registration number CRD42018118005

Keywords: Chronic pain; Sleeplessness; Management; Physical exercise; Rehabilitation; Training; Vitality.

INTRODUCTION

More than 80% of people with fibromyalgia experience severe fatigue [1] or poor sleep quality [2], both of which are identified by people with fibromyalgia and healthcare providers as priority targets for treatment. Increased fatigue and poor sleep quality are therefore acknowledged as core symptoms in the updated fibromyalgia diagnostic criteria [3]. Despite the importance of fatigue and sleep quality, most of the research to date has traditionally focused on pain-related outcomes. For instance, the European League Against Rheumatism (EULAR), highlights that exercise is the only therapy supported by 'strong' evidence for the management of fibromyalgia [4]. However, the recommendations were based on previous reviews that provided evidence of the benefits of exercise for pain but unclear for other symptoms [5–9]

In their earliest works, Busch and colleagues performed comprehensive reviews including all types of exercise (e.g., aerobic, resistance and flexibility training) [5,9]. These reviews concluded that it was unknown the effects of exercise on fatigue or sleep due to the paucity of research by that time. A number of subsequent systematic reviews focused on specific types of exercise have been published (i.e., flexibility [10], aerobic [6], resistance [8] and vibration [7] training) which have explored the effects of exercise in fatigue and sleep quality among other outcomes. Although the contribution of these reviews to the evidence base is acknowledged, the decision to narrow the scope of each review resulted in the inclusion of a restricted number of studies. For instance, for fatigue, only four and two studies were included in the reviews by Bidonde et al. [6] and by Busch et al. [8], respectively. Consequently, it is difficult to make robust conclusions about the effects of exercise interventions on fatigue and sleep in fibromyalgia. In comparison to previous reviews, a recent systematic review has focused on mixed exercise training; i.e., where two or more types of exercise are combined [11]. This review included a larger number of studies (i.e., 11 studies conducted in a total sample of 493 adults with fibromyalgia) and concluded that the effect of mixed exercise resulted in improvements in fatigue, while omitted the study of sleep quality [11]. To date no review has summarised all relevant literature on the effectiveness of exercise interventions (of any type) on fatigue and sleep quality in fibromyalgia, in doing so the current review will include a large sample size and accurately estimate, for the first time, the effects of physical exercise on these two outcomes.

The aims of this systematic review were: (i) to determine the effectiveness of exercise for reducing fatigue and improving sleep quality in people with fibromyalgia (primary aim), and (ii) to identify which type of exercise interventions might be the most effective in achieving these outcomes (secondary aim).

METHODS

A multidisciplinary and international task force was set up to conduct this review. The PRISMA guidelines were used to guide this systematic review and meta-analysis [12]. The protocol of the present review was specified in advance and registered in the PROSPERO database (registration number, CRD42018118005).

Data Sources and Searches

PubMed and Web of Science were searched from inception until October 18th, 2018. Search terms used in PubMed were: "Fibromyalgia"[MeSH] AND ("Exercise"[MeSH]) OR “Training”[All Fields]) OR "Yoga"[MeSH]) OR "Tai Ji"[MeSH]) OR "Qigong"[MeSH]) OR “Hydrotherapy"[MeSH]) OR "body awareness"[Title/Abstract]) OR danc*[Title/Abstract]). In Web of Science, the search terms were: TI=(fibromyalgia) AND TI=("exercise" OR "training" OR "yoga" OR "tai chi" OR "tai ji" OR qigong OR hydrotherapy OR "physical activity" OR "body awareness" OR danc*).

Study Selection

Two independent researchers (FEL and CMC), performed the search, screened the titles and abstracts of all retrieved articles and examined the final eligibility of the full-text articles. When a paper did not report data on fatigue or sleep quality but the study used questionnaires including these outcomes, the authors were contacted for further information. No restrictions were applied for language. This review followed the PICOS framework.

Population: adults with fibromyalgia, diagnosed using one of the recognised American College of Rheumatology (ACR) criteria: 1990, 2010, 2011, or 2016.

Intervention (exposure): based on exercise. Mixed interventions that consisted of exercise combined with other interventions (i.e., co-interventions) were considered, so long as exercise comprised at least, 50% of the intervention.

Comparison**:** studies should have either (i) an intervention group with exercise and a non-intervention control group (e.g., treatment as usual); or (ii) two exercise groups. Therefore, the primary sub-set of studies included randomised controlled trials (RCT) and the second sub-set of studies included randomised trials (RT).

Outcome: fatigue and sleep quality. When a study included more than one assessment per outcome, all the figures were extracted but only the most common assessment among the included studies was meta-analysed.

Study design: RCTs and RTs were included for the primary and secondary aims, respectively.

Data Extraction and Quality Assessment

Two independent researchers (CCM and DR) extracted the key information from each eligible study. When the information to be extracted was unavailable, authors were contacted. Disagreements were solved in a consensus meeting between the independent reviewers with a third reviewer (CMH).

The GRADE framework was used to assess the quality of the evidence across studies for fatigue and sleep quality separately. ***Risk of bias of individual studies*** was assessed using the Cochrane Risk of Bias tool. Studies with a score of, at least, five points were considered as having high risk of bias. ***Inconsistency*** across studies was considered serious when heterogeneity was high ($I^2 \geq 50\%$). ***Indirectness*** was considered serious when interventions included both exercise and additional components (i.e., co-interventions). ***Imprecision*** was considered serious when the 95% confidence interval (CI) was wide and crossed the line of no effect, and as such the interpretation of the data would be different if the true effect were at one end of the CI or the other. Finally, ***publication bias*** was assessed via funnel plots.

Two researchers independently assessed risk of bias (ICAG and MRA) and the inconsistency, indirectness, imprecision, and publication bias of the included (FEL and JGMcV) of each eligible study. Disagreements on these assessments were solved in a consensus meeting between the independent reviewers with a third reviewer (CMH).

Data Synthesis and Analysis

For the primary aim, quantitative synthesis of RCTs (i.e., meta-analyses) were performed using Review Manager V.5.3. (Cochrane Collaboration, Copenhagen, Denmark). Statistical significance was set at $p<0.05$. Standardised mean differences (SMD) between the exercise and control groups were computed for both outcomes separately. When a control group was used as a comparator twice in the same study, we halved the sample size of the control group. Weighted mean differences were calculated using a random effects model. Heterogeneity was measured using the I^2 statistic (the percentage of total variability attributed to between-study heterogeneity). When heterogeneity was high ($I^2 \geq 50\%$), further explorations based on subgroups analyses were computed. SMD effect sizes were calculated using Hedges' adjusted g (similar to Cohen's d). Effect sizes of 0.2, 0.4 and 0.8 were considered small, moderate and large, respectively.

For the secondary aim, a narrative synthesis structured around each outcome was conducted. When at least three of the included studies presented similar comparisons, we performed meta-analyses using the same methods that have been described for the primary aim.

RESULTS

Study selection and characteristics

Thirty-seven unique studies were included in this review [13–49]. Of them, four studies included 3-arms (i.e., control group and two exercise intervention groups, each one with a different exercise training such as aerobic in a group and flexibility in the another group) and therefore they were included for addressing both aims of the present review [40–43]. Thus, a total 20 RCTs [13–20,25,36,40–49] and 21 RTs [21–24,26–35,37–43] were included in the review. Figure 1 displays a PRISMA diagram.

From the 20 included RCTs that compared the effectiveness of exercise vs. usual care, 9 included both outcomes of interest [14,16,17,19,20,41,43,44,46], 8 included only fatigue [13,15,18,25,36,40,42,50] and 3 included only sleep quality [45,47,49]. From the 21 included RTs that compared the effectiveness of different exercise interventions, 12 included both outcomes of interest [21–23,27,28,30,31,33,37,39,41,43], 7 included only fatigue [24,26,29,32,35,40,42], and 2 included only sleep quality [34,38]. A summary of the characteristics of the RCTs and RTs included in the present review is presented in Supplementary Tables S1 and S2, respectively.

A moderate risk of bias was present in most of the included RCTs and RTs (see Electronic Supplementary Figures S1 and S2 for overall summaries and Electronic Supplementary Figures S3 and S4 for specific information on each individual included work per study design). No study reported having conflicts of interests.

Synthesis of the data

Figure 2 presents a meta-analysis conducted in 1003 people with fibromyalgia (61% randomly allocated into exercise interventions). In comparison to usual care, exercise interventions were effective for reducing fatigue in fibromyalgia; pooled SMD (95% CI) = -0.47 (-0.67 to -0.27). This finding was robust across two sensitivity analyses, as showed in Supplementary Figures S5 and S6: (i) when a study with high risk of bias [40] was not included in the meta-analysis; pooled SMD (95% CI) = -0.49 (-0.71 to -0.27), (ii) when fixed effects model were computed; pooled SMD (95% CI) = -0.40 (-

0.53 to -0.26). Supplementary Figure S7 presents the funnel plot, which did not indicate publication bias.

Due to the high heterogeneity (i.e., I^2=51%) observed across RCTs testing the effects of exercise on fatigue, we explored several post hoc analyses. Most of them were not significant as the effects on fatigue were similar between (i) levels of adherence: studies in which participants had to attend to, at least, 80% of the training sessions (i.e, adherence) to be included in the analyses and those studies with a lower or no adherence criterion, (ii) gender of participants: studies in which only women participated vs those in which both genders were included, (iii) type of intervention: only exercise vs co-interventions, (iv) type of exercise: meditative exercise programs (i.e., Tai Chi, Yoga and Quigong) vs others (i.e., aerobic, muscular resistance and flexibility), (v) sample size, those with at least 20 participants in each group vs the others, (vi) type of setting in which exercise was performed: land-based vs water-based, (vii) training intensity: low-to-moderate vs moderate-to-high. Supplementary figures S8 to S14 show all these non-significant findings. Interestingly, those studies in which fatigue was the primary outcome (Figure 3) and employed a shorter (less than 24 weeks) non-aerobic exercise intervention resulted in greater impact on fatigue (greater effect sizes) than comparative studies; Supplementary Figures S15 and S16.

Figure 4 presents a meta-analysis conducted in 731 people with fibromyalgia (59% randomly allocated into exercise interventions). In comparison to usual care, exercise interventions had a small effect on enhancing sleep quality in fibromyalgia; pooled

SMD (95% CI) = -0.17 (-0.32 to -0.01). This finding was robust across two sensitivity analyses, as showed in Supplementary Figures S17 and S18: (i) when a study with high risk of bias [45] was not included in the meta-analysis; pooled SMD (95% CI) = -0.19 (-0.35 to -0.02), (ii) when fixed effects model were computed; pooled SMD (95% CI) = -0.17 (-0.32 to -0.02). Due to the small heterogeneity (i.e., I^2= 5%), post hoc analyses were not needed. Supplementary Figure S19 presents the funnel plot, which did not indicate publication bias.

Table 1 shows that when comparing exercise vs. usual care, there was ‘low to moderate’ quality evidence for the beneficial effects of exercise on fatigue, while the evidence was ‘moderate’ for benefits on sleep quality.

In the 21 RTs included in the present review, a wide range of exercise interventions were implemented and compared in a total of 1254 people with fibromyalgia (all randomly allocated into different interventions). Thus, it was difficult to perform robust comparisons. However, we were able to quantify one comparison for sleep quality and three for fatigue. First, when comparing different types of exercise, meditative exercise programs were more effective for improving sleep quality but not for lowering fatigue; Figure 5, pooled SMD (95% CI) = -0.80 (-1.57 to -0.02) and Supplementary Figure S20, pooled SMD (95% CI) = -0.39 (-0.88 to 0.11), respectively. Second, the effectiveness of resistance vs flexibility training was similar for fatigue; Supplementary Figure S21, pooled SMD (95% CI) = -1.64 (-4.31 to 1.02). Third, the effectiveness of water vs land-based exercise was also similar for fatigue; Supplementary Figure S22, pooled SMD (95% CI) = 0.00 (-0.42 to 0.43).

DISCUSSION

This systematic review set out to determine the effectiveness of exercise on fatigue and sleep quality in those with fibromyalgia and to identify which type of exercise interventions might be the most effective in achieving these outcomes. In the current review we have found that, in comparison to usual care, exercise has moderate effects for lowering fatigue and small effects for improving sleep quality. We have also observed that, in comparison of other types of exercise, meditative exercise programs were more effective for improving sleep quality but not for lowering fatigue. In interpreting the findings of this review a number of factors must to be noted. First, most of the studies were based on aerobic exercise. Thus, the effect sizes of the present meta-analyses may reflect more accurately the effectiveness of aerobic training on fatigue and sleep quality than the effects of other types of exercise. Indeed, we observed that those exercise interventions that did not include aerobic exercise seemed to be more effective at reducing fatigue. Second, the effects of exercise on fatigue were highly variable across studies (I^2 = 51%) and remarkably higher when fatigue was the main outcome. Third, there is a lack of high quality studies in the field and consequently the quality of evidence provided in the present review is low to moderate for the effectiveness of exercise in reducing fatigue (the evidence is in favour of exercise but the effect size is unclear, likely to be moderate) and moderate for small effects of exercise (of any type) on enhancing sleep quality.

Effectiveness of exercise for reducing fatigue in fibromyalgia

Due to the limited number of studies included in previous meta-analyses, their findings were inconclusive and inconsistent. For example, Busch and colleagues meta-analysed two resistance training studies (n=81) showing significant pooled reductions on fatigue ($p<0.001$) [8]. However, Bidonde and colleagues have recently meta-analysed four aerobic exercise studies (n=286) [6] in which the p-value of the pooled effects for the exercise group was 0.06. Recently, Bidonde and colleagues have meta-analysed a sample size of 493 adults with fibromyalgia estimating that the effects of mixed exercise training (i.e., where two or more types of exercise are combined) on fatigue were significant ($p<0.001$) [11]. Using similar statistical methods to previous meta-analyses but in a larger sample size (n=1003), our pooled estimation showed that exercise produces significant and probably a meaningful (moderate effect size) reduction in fatigue in fibromyalgia. Thus, the comprehensive approach followed in the present meta-analysis allowed us to robustly determine, for the first time, the overall effects of exercise on fatigue in large sample of people with fibromyalgia.

Effectiveness of exercise for improving sleep quality in fibromyalgia

Previous meta-analyses were unable to determine the effectiveness of exercise on sleep quality in fibromyalgia due to the paucity of research. Indeed, most of them failed to find RCTs on this topic. Given the extent of sleep dysfunction in those with fibromyalgia, it is important to determine the effectiveness of exercise for improving sleep quality in this population. The most comprehensive review to date included only two studies examining sleep and reported moderate effects of exercise for enhancing sleep quality (n= 104) [51]. The number of included studies in the present work was considerably higher (13 RCTs, n=806) leading us to better estimate the effectiveness of exercise for improving sleep quality in fibromyalgia. The effectiveness of exercise (of any type) in enhancing sleep quality in fibromyalgia was limited (small effect), however, meditative exercise programs (i.e., Tai Chi, Yoga and Quigong) may offer a promising approach. Although there are potential mechanisms which can provide a rationale to support the effectiveness of meditative exercise on improving sleep quality (see, next section), our finding is based on an imprecise estimation (SMD (95% CI) = -0.80 (-1.57 to -0.02)), from a relatively small sample size (141 participants in meditative exercise vs. 177 participants in other types of exercise). Thus, further high quality experimental research is required to confirm or refute our findings.

Exercise mechanisms for fatigue and sleep quality in fibromyalgia

Aberrations in the central nervous system are well-known in fibromyalgia [52–54]. For example, in comparison to non-fibromyalgia controls, abnormal levels of metabolites (e.g., reductions in the ratio of N-acetylaspartate to creatine) have been observed in the hippocampus of people with fibromyalgia [52] as well as structural abnormalities (e.g., lower volume) [53] and functional changes (e.g., increased activation) [54]. Another system that might be altered in fibromyalgia is the hypothalamic-pituitary-adrenal axis (HPA) as well as a sympathetic hyperactivity mediated by a dysfunction in the autonomic nervous system (ANS) [55]. These alterations may in turn be related to increased levels of fatigue [52]. Interestingly, exercise may revert these aberrations by regulating the levels of metabolites as well as promoting angiogenesis, neurogenesis and connectivity of the hippocampus [56,57].

In the present meta-analyses demonstrated that exercise had a small beneficial effect on sleep quality in fibromyalgia. In this disease, hyperactivity of the sympathetic nervous system is well-documented and, thus, stress levels are considerably high [58,59]. Physiological responses to exercise often include a decrease in this sympathetic tone and a shift toward parasympathetic activity, which in turn may be related to muscular and nervous relaxation, leading to reductions in stress levels and, finally, to improvement in sleep quality [60–62]. In this respect, our review showed that meditative exercise programs were more effective in improving sleep quality than other types of exercise. Although meditative exercise is safe in fibromyalgia, little is known about their mechanisms of action. It is likely that this type of exercise is able to enhance the parasympathetic activity and reduce sympathetic tone by decreasing activation of HPA axis. Moreover, meditative exercise may facilitate enhanced rapid eye movement (REM) sleep by increasing central nervous system inhibitory c-aminobutyric acid (GABA) and serotonin levels [63].

Clinical applications

The recent European League Against Rheumatism (EULAR) recommendations for the management of fibromyalgia, highlight exercise as the only therapy with a strong level of evidence [4]. These recommendations were based on the findings provided by systematic review of previous reviews. As we have discussed, while previous systematic reviews showed reliable findings for pain management, they have provided limited evidence on the effectiveness of exercise for reducing fatigue and increasing sleep quality in fibromyalgia [6,8]. The present meta-analyses suggest that the effectiveness of exercise may differ for different outcomes. This means that it cannot be assumed that the benefits of exercise on pain automatically extend to other symptoms of the condition. An interesting finding for healthcare providers has emerged from our review in that fatigue reductions were higher when the main outcome of the study was fatigue. Therefore, instead of designing a 'fix-all' exercise protocol for fibromyalgia, exercise programmes should be designed as outcome-specific, by considering how fibromyalgia manifests in the person who is going to engage in the programme. For example, meditative exercise programs (e.g., Tai Chi or Qigong) may be more advisable for people with fibromyalgia who experience difficulty sleeping.

The studies included in the present systematic review investigated a wide range of exercise programs, including different types of exercise, intensities, frequencies, and program duration. Although we explored several post-hoc analyses, we were unable to determine the most effective exercise intervention for reducing fatigue. From our approach to subgrouping the effects of different exercise interventions, in comparison to usual care, we observed that the ideal intervention for lowering fatigue in fibromyalgia seems to be specifically designed for such outcome, lasts less than 24 weeks and does not involve aerobic exercise. Collectively, the high heterogeneity that emerged from the effects of exercise on fatigue limits the establishment of evidence-based guidelines. Although the American College of Sports Medicine (ACSM) has launched specific recommendations to consider when conducting exercise interventions in fibromyalgia [64], a recent review has reported poor therapeutic validity of studies that accomplish these ACSM exercise recommendations [65].

Standard exercise interventions for the ‘average’ or ‘most common profile’ of people with fibromyalgia seems misjudged as people with fibromyalgia are heterogeneous [66]. Thus, personalised exercise programs are warranted. In this context, some people with fibromyalgia may experience fears related to engaging in exercise [67] or a discordance of being more capable to engage in exercise than is self-perceived [68]. People with these characteristics may be more likely to experience exercise as stressful. Therefore, exercise interventions should not only be tailored to how fibromyalgia manifests in each person but also to (more) general characteristics of the person.

Implications for research agenda

Findings of the present study provide evidence indicating that exercise is effective for reducing fatigue in fibromyalgia. However, future research is needed to determine what type of exercise is most beneficial for people with fibromyalgia, which intensity is best, the optimal length of the training, as well as the most beneficial delivery method.

Our findings indicate that exercise seems to promote only small benefits on sleep quality in fibromyalgia, and while relaxation is a potential mechanism by which exercise might improve sleep quality, not all types of exercise promote relaxation. Thus, meditative exercise programs that do suppose a lower physical load than other types of exercises could be more effective for enhancing sleep quality in fibromyalgia. Therefore, future large experimental studies of high quality, testing the effectiveness of very gentle exercise specifically designed for enhancing sleep quality in fibromyalgia are warranted. Additionally, further research testing the effectiveness of exercise in objectively measured fatigue or sleep quality is warranted.

Limitations and strengths of the present study

The most common limitations among the included studies were: (i) the long-term effects of the interventions were not reported, (ii) results were not stratified by sex and most of the participants were women. Moreover, we did not include conference proceedings and other types of grey literature due to the often low quality of reporting in conference abstracts.

Conclusions

We provided low-to-moderate quality evidence that exercise is moderately effective for lowering fatigue and that there is moderate evidence of small effects of exercise for enhancing sleep quality in fibromyalgia. Although speculative, meditative exercise programs may be a promising approach for improving sleep quality in fibromyalgia. As most of the interventions involved aerobic exercise, research using other types of exercise is warranted. Instead of designing ‘fix-all’ and ‘one size fits all’ protocols, exercise programmes, in order to be as effective as possible, should be specifically designed for the outcome that is targeted and tailored to the characteristics of the person who is going to engage in the exercise.

Funding statement. This work was supported by the Health and Social Care Public Health Agency, Northern Ireland [STL/5268/16 to CH and JGMcV]. FE-L has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 707404. The funders of the present study did not have any role in the study design, data collection and analyses, decision to publish, or preparation of the manuscript. FEL is the guarantor of the review.

Conflicts of interest statement: The authors declare no conflicts of interest to report.

Tables (Legends)

Table 1. Level of quality of the evidence for the effectiveness of exercise for reducing fatigue and enhancing sleep quality in fibromyalgia.

CI, Confidence interval; SMD, Standardised mean difference

Figures (Legends)

Figure 1. Flow chart showing the results of the selection process.

* Four studies included a usual care (control) group and two different exercise interventions. Thus, they were included in the analyses related to the primary and second aims of the present review.

Figure 2. Pooled effects of randomised controlled trials analysing the effectiveness of exercise in reducing fatigue in people with fibromyalgia.

Analyses were conducted using a random effects model. CI, Confidence Interval; df, degrees of freedom; Std, Standardised; SD, Standard Deviation; IV, Inverse Variance; A, Aerobic exercise; Co-,
Co-intervention (Edu, education; Photo, phototherapy); F, Flexibility exercise; L- and W-B, land- and water-based exercise, respectively; M, Meditative exercise; R, Resistance exercise; TC, Tai Chi; Y, Yoga.

Figure 3. Post hoc analysis showing the pooled effects of randomised trials analysing the effectiveness of studies in which fatigue was the primary outcome vs. the remaining studies for lowering fatigue in people with fibromyalgia

Analyses were conducted using a random effects model. CI, Confidence Interval; df, degrees of freedom; Std, Standardised; SD, Standard Deviation; IV, Inverse Variance; A, Aerobic exercise; Co-, Co-intervention (Edu, education; Photo, phototherapy); F, Flexibility exercise; L- and W-B, land- and water-based exercise, respectively; M, Meditative exercise; R, Resistance exercise; TC, Tai Chi; Y, Yoga.

Figure 4. Pooled effects of randomised controlled trials analysing the effectiveness of exercise in enhancing sleep quality in people with fibromyalgia.

Analyses were conducted using a random effects model. CI, Confidence Interval; df, degrees of freedom; Std, Standardised; SD, Standard Deviation; IV, Inverse Variance; A, Aerobic exercise; Co-, Co-intervention (CBT, Cognitive Behaviour Therapy; Edu, education; Photo, phototherapy); F, Flexibility exercise; L- and W-B, land- and water-based exercise, respectively; M, Meditative exercise; R, Resistance exercise; TC, Tai Chi; QG, Qigong; Y, Yoga.

Figure 5. Post hoc analysis showing the pooled effects of randomised trials analysing the effectiveness of meditative exercise vs. the remaining types of exercise for enhancing sleep quality in people with fibromyalgia.

Analyses were conducted using a random effects model. CI, Confidence Interval; df, degrees of freedom; Std, Standardised; SD, Standard Deviation; IV, Inverse Variance; A, Aerobic exercise; AC, Ai Chi; AqBD, Aquatic Biodanza; BA, Body Awareness; F, flexibility exercise; TC, Tai Chi.

REFERENCES

[1] Overman CL, Kool MB, Da Silva J a. P, Geenen R. The prevalence of severe fatigue in rheumatic diseases: an international study. Clin Rheumatol 2016;35:409–15. doi:10.1007/s10067-015-3035-6.

[2] Bigatti SM, Hernandez AM, Cronan TA, Rand KL. Sleep disturbances in fibromyalgia syndrome: relationship to pain and depression. Arthritis Rheum 2008;59:961–7. doi:10.1002/art.23828.

[3] Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10. doi:10.1002/acr.20140.

[4] Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis 2017;76:318–28. doi:10.1136/annrheumdis-2016-209724.

[5] Busch AJ, Schachter CL, Overend TJ, Peloso PM, Barber KAR. Exercise for fibromyalgia: a systematic review. J Rheumatol 2008;35:1130–44.

[6] Bidonde J, Busch AJ, Schachter CL, Overend TJ, Kim SY, Góes SM, et al. Aerobic exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2017;2017. doi:10.1002/14651858.CD012700.

[7] Bidonde J, Busch AJ, van der Spuy I, Tupper S, Kim SY, Boden C. Whole body

vibration exercise training for fibromyalgia. Cochrane Database Syst Rev 2017;9:CD011755. doi:10.1002/14651858.CD011755.pub2.

[8] Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. Cochrane Database Syst Rev 2013;12:CD010884. doi:10.1002/14651858.CD010884.

[9] Busch A, Schachter CL, Peloso PM, Bombardier C. Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev 2002:CD003786. doi:10.1002/14651858.CD003786.

[10] Kim SY, Busch AJ, Overend TJ, Schachter CL, van der Spuy I, Boden C, et al. Flexibility exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2019;9:CD013419. doi:10.1002/14651858.CD013419.

[11] Bidonde J, Busch AJ, Schachter CL, Webber SC, Musselman KE, Overend TJ, et al. Mixed exercise training for adults with fibromyalgia. Cochrane Database Syst Rev 2019. doi:10.1002/14651858.CD013340.

[12] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535. doi:10.1016/j.jcms.2010.11.001.

[13] Alentorn-Geli E, Padilla J, Moras G, Lazaro Haro C, Fernandez-Sola J. Six Weeks of Whole-Body Vibration Exercise Improves Pain and Fatigue in Women with Fibromyalgia. J Altern Complement Med 2008;14:975–81. doi:10.1089/acm.2008.0050.

[14] Carson JW, Carson KM, Jones KD, Bennett RM, Wright CL, Mist SD. A pilot randomized controlled trial of the Yoga of Awareness program in the management of fibromyalgia. Pain 2010;151:530–9.

doi:10.1016/j.pain.2010.08.020.

[15] Tomas-Carus P, Gusi N, Hakkinen A, Hakkinen K, Leal A, Ortega-Alonso A. 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. doi:10.2340/16501977-0168.

[16] Tomas-Carus P, Hakkinen A, Gusi N, Leal A, Hakkinen K, Ortega-Alonso A. Aquatic training and detraining on fitness and quality of life in fibromyalgia. Med Sci Sports Exerc 2007;39:1044–50. doi:10.1249/mss.0b0138059aec4.

[17] Valkeinen H, Alen M, Hakkinen A, Hannonen P, Kukkonen-Harjula K, Hakkinen K. Effects of concurrent strength and endurance training on physical fitness and symptoms in postmenopausal women with fibromyalgia: A randomized controlled trial 2008;89:1660–6. doi:10.1016/j.apmr.2008.01.022.

[18] van Santen M, Bolwijn P, Verstappen F, Bakker C, Hidding A, Houben H, et al. A randomized clinical trial comparing fitness and biofeedback training versus basic treatment in patients with fibromyalgia. J Rheumatol 2002;29:575–81.

[19] Wigers SH, Stiles TC, Vogel PA. Effects of aerobic exercise versus stress management treatment in fibromyalgia. A 4.5 year prospective study. Scand J Rheumatol 1996;25:77–86. doi:10.3109/03009749609069212.

[20] Wong A, Figueroa A, Sanchez-Gonzalez MA, Son W-M, Chernykh O, Park S-Y. Effectiveness of Tai Chi on Cardiac Autonomic Function and Symptomatology in Women With Fibromyalgia: A Randomized Controlled Trial. J Aging Phys Act 2018;26:214–21. doi:10.1123/japa.2017-0038.

[21] Bircan C, Karasel SA, Akgun B, El O, Alper S. Effects of muscle strengthening versus aerobic exercise program in fibromyalgia. Rheumatol Int 2008;28:527–32.

doi:10.1007/s00296-007-0484-5.

[22] Calandre EP, Rodriguez-Claro ML, Rico-Villademoros F, Vilchez JS, Hidalgo J, Delgado-Rodriguez A. Effects of pool-based exercise in fibromyalgia symptomatology and sleep quality: a prospective randomised comparison between stretching and Ai Chi. Clin Exp Rheumatol 2009;27:S21–8.

[23] Demir-Gocmen D, Altan L, Korkmaz N, Arabaci R. Effect of supervised exercise program including balance exercises on the balance status and clinical signs in patients with fibromyalgia. Rheumatol Int 2013;33:743–50. doi:10.1007/s00296-012-2444-y.

[24] Fernandes G, Jennings F, Nery Cabral MV, Pirozzi Buosi AL, Natour J. Swimming Improves Pain and Functional Capacity of Patients With Fibromyalgia: A Randomized Controlled Trial. Arch Phys Med Rehabil 2016;97:1269–75. doi:10.1016/j.apmr.2016.01.026.

[25] Collado-Mateo D, Dominguez-Munoz FJ, Adsuar JC, Garcia-Gordillo MA, Gusi N. Effects of Exergames on Quality of Life, Pain, and Disease Effect in Women With Fibromyalgia: A Randomized Controlled Trial. Arch Phys Med Rehabil 2017;98:1725–31. doi:10.1016/j.apmr.2017.02.011.

[26] Gavi MBRO, Vassalo DV, Amaral FT, Macedo DCF, Gava PL, Dantas EM, et al. Strengthening exercises improve symptoms and quality of life but do not change autonomic modulation in fibromyalgia: a randomized clinical trial. PLoS One 2014;9:e90767. doi:10.1371/journal.pone.0090767.

[27] Jentoft ES, Kvalvik AG, Mengshoel M. Effects of pool-based and land-based aerobic exercise on women with fibromyalgia/chronic widespread muscle pain. ARTHRITIS Rheum CARE Res 2001;45:42–7.

[28] Jones KD, Burckhardt CS, Clark SR, Bennett RM, Potempa KM. A Randomized controlled trial of muscle strengthening versus flexibility training in fibromyalgia. J Rheumatol 2002;29:1041–8.

[29] Kendall SA, Brolin-Magnusson K, Soren B, Gerdle B, Henriksson KG. A pilot study of body awareness programs in the treatment of fibromyalgia syndrome. Arthritis Care Res (Hoboken) 2000;13:304–11.

[30] López-Pousa S, Bassets Pagès G, Monserrat-Vila S, de Gracia Blanco M, Hidalgo Colomé J, Garre-Olmo J. Sense of Well-Being in Patients with Fibromyalgia: Aerobic Exercise Program in a Mature Forest-A Pilot Study. Evid Based Complement Alternat Med 2015;2015:614783. doi:10.1155/2015/614783.

[31] Lopez-Rodriguez MM, Fernandez-Martinez M, Mataran-Penarrocha GA, Rodriguez-Ferrer ME, Granados Gamez G, Aguilar Ferrandiz E. [Effectiveness of aquatic biodance on sleep quality, anxiety and other symptoms in patients with fibromyalgia]. Med Clin (Barc) 2013;141:471–8. doi:10.1016/j.medcli.2012.09.036.

[32] Mannerkorpi K, Nordeman L, Cider Å, Jonsson G. Does moderate-to-high intensity Nordic walking improve functional capacity and pain in fibromyalgia? A prospective randomized controlled trial. Arthritis Res Ther 2010;12:R189. doi:10.1186/ar3159.

[33] Norregaard J, Lykkegaard JJ, Mehlsen J, DanneskioldSamsoe B. Exercise training in treatment of fibromyalgia. J Musculoskelet Pain 1997;5:71–9. doi:10.1300/J094v05n01_05.

[34] van Santen M, Bolwijn P, Landewe R, Verstappen F, Bakker C, Hidding A, et al. High or low intensity aerobic fitness training in fibromyalgia: Does it matter? J

Rheumatol 2002;29:582–7.

[35] Vitorino DFD, Carvalho LBC, do Prado GF. Hydrotherapy and conventional physiotherapy improve total sleep time and quality of life of fibromyalgia patients: Randomized clinical trial. SLEEP Med 2006;7:293–6. doi:10.1016/j.sleep.2005.09.002.

[36] Etnier JL, Karper WB, Gapin JI, Barella LA, Chang YK, Murphy KJ. Exercise, Fibromyalgia, and Fibrofog: A Pilot Study. J Phys Act Health 2009;6:239–46. doi:10.1123/jpah.6.2.239.

[37] Wang C, Schmid CH, Rones R, Kalish R, Yinh J, Goldenberg DL, et al. A Randomized Trial of Tai Chi for Fibromyalgia. N Engl J Med 2010;363:743–54. doi:10.1056/NEJMoa0912611.

[38] Wang C, Schmid CH, Fielding RA, Harvey WF, Reid KF, Price LL, et al. Effect of tai chi versus aerobic exercise for fibromyalgia: comparative effectiveness randomized controlled trial. BMJ 2018;360:k851. doi:10.1136/bmj.k851.

[39] Genc A, Tur BS, Aytur YK, Oztuna D, Erdogan MF. Does aerobic exercise affect the hypothalamic-pituitary-adrenal hormonal response in patients with fibromyalgia syndrome? J Phys Ther Sci 2015;27:2225–31. doi:10.1589/jpts.27.2225.

[40] Assumpcao A, Matsutani LA, Yuan SL, Santo AS, Sauer J, Mango P, et al. Muscle stretching exercises and resistance training in fibromyalgia: which is better? A three-arm randomized controlled trial. Eur J Phys Rehabil Med 2018;54:663–70. doi:10.23736/S1973-9087.17.04876-6.

[41] McBeth J, Prescott G, Scotland G, Lovell K, Keeley P, Hannaford P, et al. Cognitive behavior therapy, exercise, or both for treating chronic widespread

pain. Arch Intern Med 2012;172:48–57. doi:10.1001/archinternmed.2011.555.

[42] Schachter CL, Busch AJ, Peloso PM, Sheppard MS. Effects of short versus long bouts of aerobic exercise in sedentary women with. Fibromyalgia: A randomized controlled trial. Phys Ther 2003;83:340–58.

[43] da Silva MM, Albertini R, de Tarso Camillo de Carvalho P, Leal-Junior ECP, Bussadori SK, Vieira SS, et al. Randomized, blinded, controlled trial on effectiveness of photobiomodulation therapy and exercise training in the fibromyalgia treatment. Lasers Med Sci 2018;33:343–51. doi:10.1007/s10103-017-2388-2.

[44] Giannotti E, Koutsikos K, Pigatto M, Rampudda ME, Doria A, Masiero S. Medium-/long-term effects of a specific exercise protocol combined with patient education on spine mobility, chronic fatigue, pain, aerobic fitness and level of disability in fibromyalgia. Biomed Res Int 2014;2014:474029. doi:10.1155/2014/474029.

[45] Haak T, Scott B. The effect of Qigong on Fibromyalgia (FMS): A controlled randomized study. Disabil Rehabil 2008;30:625–33. doi:10.1080/09638280701400540.

[46] Hakkinen A, Hakkinen K, Hannonen P, Alen M. Strength training induced adaptations in neuromuscular function of premenopausal women with fibromyalgia: comparison with healthy women. Ann Rheum Dis 2001;60:21–6. doi:10.1136/ard.60.1.21.

[47] Lynch M, Sawynok J, Hiew C, Marcon D. A randomized controlled trial of qigong for fibromyalgia. Arthritis Res Ther 2012;14:R178. doi:10.1186/ar3931.

[48] Mannerkorpi K, Nyberg B, Ahlmen M, Ekdahl C. Pool exercise combined with

an education program for patients with fibromyalgia syndrome. A prospective, randomized study. J Rheumatol 2000;27:2473–81.

[49] Sanudo B, Carrasco L, de Hoyo M, Figueroa A, Saxton JM. Vagal modulation and symptomatology following a 6-month aerobic exercise program for women with fibromyalgia. Clin Exp Rheumatol 2015;33:S41-5.

[50] Palstam A, Mannerkorpi K. Work ability in fibromyalgia: an update in the 21st century. Curr Rheumatol Rev 2017. doi:10.2174/1573397113666170502152955.

[51] Bidonde J, AJ B, SC W, CL S, Danyliw A, TJ O, et al. Aquatic exercise training for fibromyalgia. Cochrane Database Syst Rev 2014;10:CD011336. doi:10.1002/14651858.CD011336.

[52] Wood PB, Ledbetter CR, Glabus MF, Broadwell LK, Patterson JC. Hippocampal metabolite abnormalities in fibromyalgia: correlation with clinical features. J Pain 2009;10:47–52. doi:10.1016/j.jpain.2008.07.003.

[53] McCrae C, O'Shea A, Boissoneault J, Vatthauer K, Robinson M, Staud R, et al. Fibromyalgia patients have reduced hippocampal volume compared with healthy controls. J Pain Res 2015;8:47. doi:10.2147/JPR.S71959.

[54] González-Roldán AM, Bomba IC, Diesch E, Montoya P, Flor H, Kamping S. Controllability and hippocampal activation during pain expectation in fibromyalgia syndrome. Biol Psychol 2016;121:39–48. doi:10.1016/j.biopsycho.2016.09.007.

[55] Martinez-Lavin M. Biology and therapy of fibromyalgia. Stress, the stress response system, and fibromyalgia. Arthritis Res Ther 2007;9:216. doi:10.1186/ar2146.

[56] Valim V, Natour J, Xiao Y, Pereira AFA, Lopes BB da C, Pollak DF, et al.

Effects of physical exercise on serum levels of serotonin and its metabolite in fibromyalgia: a randomized pilot study. Rev Bras Reumatol n.d.;53:538–41. doi:10.1016/j.rbr.2013.02.001.

[57] Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosci 2008;9:58–65. doi:10.1038/nrn2298.

[58] Petzke F, Clauw DJ. Sympathetic nervous system function in fibromyalgia. Curr Rheumatol Rep 2000;2:116–23.

[59] SARZI-PUTTINI P, Atzeni F, Diana A, Doria A, Furlan R. Increased Neural Sympathetic Activation in Fibromyalgia Syndrome. Ann N Y Acad Sci 2006;1069:109–17. doi:10.1196/annals.1351.009.

[60] O'Connor PJ, Youngstedt SD. Influence of exercise on human sleep. Exerc Sport Sci Rev 1995;23:105–34.

[61] Kubitz KA, Landers DM, Petruzzello SJ, Han M. The effects of acute and chronic exercise on sleep. A meta-analytic review. Sports Med 1996;21:277–91. doi:10.2165/00007256-199621040-00004.

[62] Stults-Kolehmainen MA, Sinha R. The Effects of Stress on Physical Activity and Exercise. Sport Med 2014;44:81–121. doi:10.1007/s40279-013-0090-5.

[63] Karimi M, Yazdani Noori A. Serotonin and Mood State Changes in Response to a Period of Yoga Training in Well-Trained Wrestlers. Int J Wrestl Sci 2015;5:89–92. doi:10.1080/21615667.2015.1116647.

[64] Pescatello LS, American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. 9th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health; 2014.

[65] Álvarez-Gallardo IC, Bidonde J, Busch A, Westby M, Kenny GP, Delgado-Fernández M, et al. Therapeutic validity of exercise interventions in the management of fibromyalgia. J Sports Med Phys Fitness 2018. doi:10.23736/S0022-4707.18.08897-7.

[66] Estévez-López F, Segura-Jiménez V, Álvarez-Gallardo IC, Borges-Cosic M, Pulido-Martos M, Carbonell-Baeza A, et al. Adaptation profiles comprising objective and subjective measures in fibromyalgia: the al-Ándalus project. Rheumatology 2017;56:2015–24. doi:10.1093/rheumatology/kex302.

[67] Russell D, Álvarez Gallardo IC, Wilson I, Hughes CM, Davison GW, Sañudo B, et al. “Exercise to me is a scary word”: perceptions of fatigue, sleep dysfunction, and exercise in people with fibromyalgia syndrome-a focus group study. Rheumatol Int 2018;38:507–15. doi:10.1007/s00296-018-3932-5.

[68] Estévez-López F, Álvarez-Gallardo IC, Segura-Jiménez V, Soriano-Maldonado A, Borges-Cosic M, Pulido-Martos M, et al. The discordance between subjectively and objectively measured physical function in women with fibromyalgia: association with catastrophizing and self-efficacy cognitions. The al-Ándalus project. Disabil Rehabil 2018;40:329–37. doi:10.1080/09638288.2016.1258737.

<u>Electronic Supplementary Material</u>: Effectiveness of exercise in fatigue and sleep quality in fibromyalgia: a systematic review and meta-analysis of randomised trials by Estévez-López et al.

List of Electronic Supplementary Material tables

Electronic Supplementary Material Table S1. Summary of the randomised controlled trials testing the effectiveness of exercise interventions in comparison to usual care.

ACR, American College Rheumatology; CPG, Chronic Pain Grade Questionnaire; GHQ, General Health Questionnaire; FIQ, Fibromyalgia Impact Questionnaire; FIQR, Fibromyalgia Impact Questionnaire Revised; FM; Fibromyalgia; FSS, Fatigue Severity Scale; HAQ, Stanford Health Assessment Questionnaire; HRmax, Maximum Heart Rate; PSQI, Pittsburgh Sleep Quality Index; min, minutes; RM, Repetition Maximum; USA, United States of America; VNS, daily self-recordings of a 15-item Visual Numerological Scale; VAS, Visual Analogue Scale.

* Primary outcome of the study.

Electronic Supplementary Material Table S2. Summary of the randomised trials comparing the effectiveness of different exercise interventions.

ACR, American College Rheumatology; CPG, Chronic Pain Grade Questionnaire; GHQ, General Health Questionnaire; FIQ, Fibromyalgia Impact Questionnaire; FIQR, Fibromyalgia Impact Questionnaire Revised; FM; Fibromyalgia; FSS, Fatigue Severity Scale; HAQ, Stanford Health Assessment Questionnaire; MFI-20, Multidimensional Fatigue Inventory ; HRmax, Maximum Heart Rate; PSQI, Pittsburgh Sleep Quality Index; min, minutes; RM, Maximum Repetition; VNS, daily self-recordings of a 15-item Visual Numerological Scale; VAS, Visual Analogue Scale *Primary outcome

List of Electronic Supplementary Material figures

Electronic Supplementary Material Figure S1. Summary of risk of bias of the randomised controlled trials testing the effectiveness of exercise interventions in comparison to usual care.

Electronic Supplementary Material Figure S2. Summary of risk of bias of the randomised trials comparing the effectiveness of different exercise interventions.

Electronic Supplementary Material Figure S3. Risk of bias of each randomised controlled trial testing the effectiveness of exercise interventions in comparison to usual care.

Electronic Supplementary Material Figure S4. Risk of bias of each randomised trial comparing the effectiveness of different exercise interventions.

Electronic Supplementary Material Figure S5. Pooled effects of the randomised controlled trials analysing the effectiveness of exercise in reducing fatigue in people with fibromyalgia: sensitivity analyses excluding studies with high risk of bias.

Electronic Supplementary Material Figure S6. Pooled effects of the randomised controlled trials analysing the effectiveness of exercise in reducing fatigue in people with fibromyalgia: sensitivity analyses using fixed effects model.

Electronic Supplementary Material Figure S7. Funnel plot of the randomised controlled trials analysing the effectiveness of exercise in reducing fatigue in people with fibromyalgia.

Electronic Supplementary Material Figure S8. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to levels of adherence (80% adherence at minimum vs. lower rates or none adherence criterion).

Electronic Supplementary Material Figure S9. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: gender of the participants (only women vs both genders).

Electronic Supplementary Material Figure S10. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to type of interventions (only exercise vs exercise + co-intervention).

Electronic Supplementary Material Figure S11. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to type of exercise (meditative exercise vs other types).

Electronic Supplementary Material Figure S12. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to sample size (20 participants per group at minimum vs lower sample sizes).

Electronic Supplementary Material Figure S13. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to settings (land-based vs water-based exercise).

Electronic Supplementary Material Figure S14. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to intensity (low-to-moderate vs moderate-to-high).

Electronic Supplementary Material Figure S15. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to aerobic exercise (aerobic exercise in insolation or combined with other exercises vs other exercises that did not include aerobic training).

Electronic Supplementary Material Figure S16. Post hoc analyses of the randomised controlled trials testing the effectiveness of exercise interventions in fatigue in comparison to usual care: subgroups according to length of the intervention (24 weeks at minimum vs shorter interventions).

Electronic Supplementary Material Figure S17. Pooled effects of the randomised controlled trials analysing the effectiveness of exercise in enhancing sleep quality in people with fibromyalgia: sensitivity analyses excluding studies with high risk of bias.

Electronic Supplementary Material Figure S18. Pooled effects of the randomised controlled trials analysing the effectiveness of exercise in enhancing sleep quality in people with fibromyalgia: sensitivity analyses using fixed effects model.

Electronic Supplementary Material Figure S19. Funnel plot of the randomised controlled trials analysing the effectiveness of exercise in enhancing sleep quality in people with fibromyalgia.

Electronic Supplementary Material Figure S20. Pooled effects of the randomised trials comparing the effectiveness of meditative exercise and other types of exercise in fatigue in people with fibromyalgia.

Electronic Supplementary Material Figure S21. Pooled effects of the randomised trials comparing the effectiveness of resistance and flexibility exercise in fatigue in people with fibromyalgia.

Electronic Supplementary Material Figure S22. Pooled effects of the randomised trials comparing the effectiveness of land-based and water-based exercise in fatigue in people with fibromyalgia.

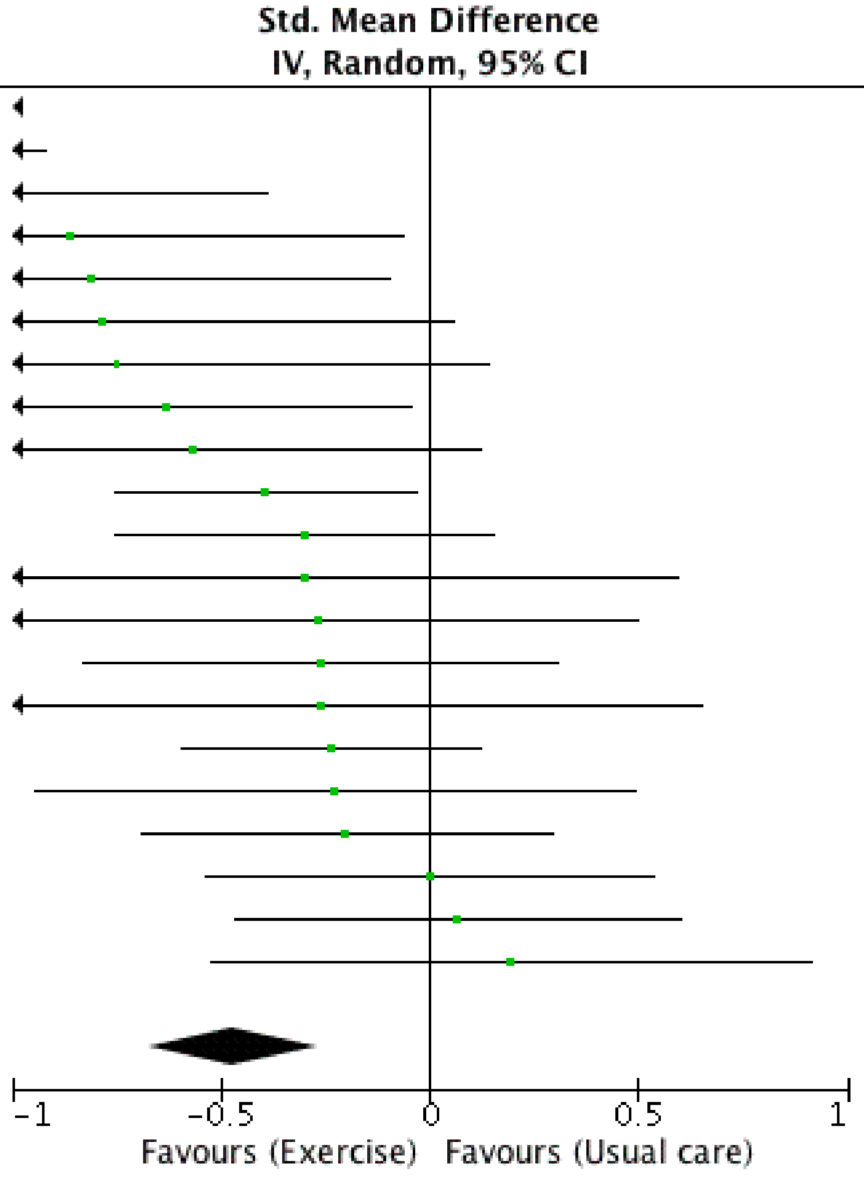
Table 1. Level of quality evidence for the effectiveness of exercise for reducing fatigue and enhancing sleep quality in fibromyalgia.

Certainty assessment							№ of participants		Effect		Certainty	Direction
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Exercise**	**Usual care**	**SMD (95% CI)**	**Size**		
Outcome = Fatigue												
17	Randomised controlled trials	Not serious	Serious	Not serious	Unclear	Not serious	612/1003 (62%)	391/1003 (39%)	**-0.47** (-0.67 to -0.27)	**Moderate**	⊕⊕◯◯ to ⊕⊕⊕◯ Low to moderate	**In favour of exercise**
Outcome = Sleep quality												
12	Randomised controlled trials	Not serious	Not serious	Not serious	Serious	Not serious	431/731 (59%)	300/731 (41%)	**-0.17** (-0.32 to -0.01)	**Small**	⊕⊕⊕◯ Moderate	**In favour of exercise**

CI, Confidence interval; SMD, Standardised mean differences

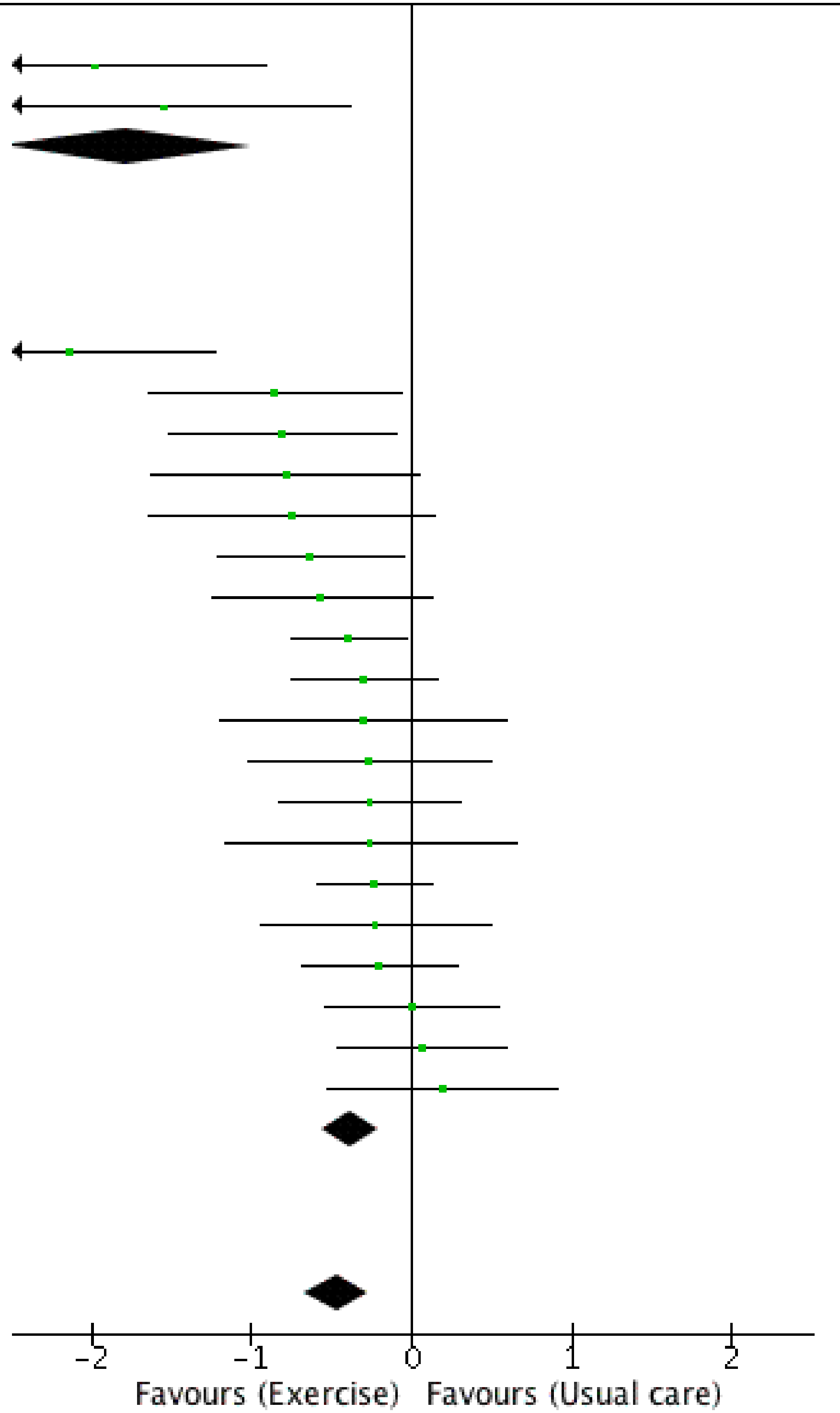
Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**

Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$
Test for overall effect: Z = 4.55 (P < 0.00001)

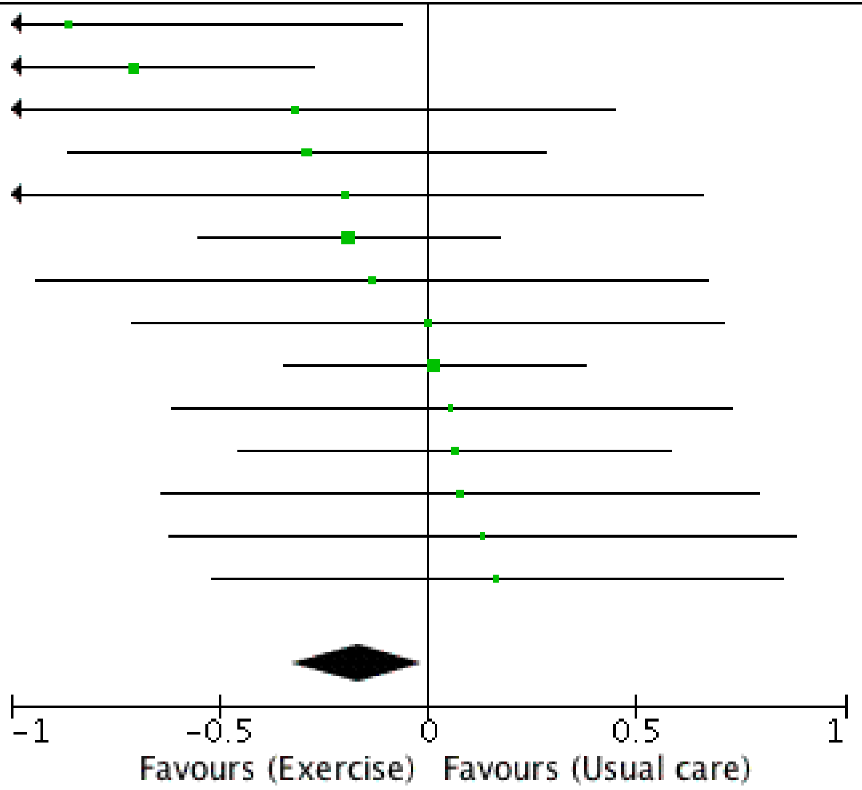


Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 The primary outcome was fatigue								
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Subtotal (95% CI)			**20**			**18**	**5.0%**	**-1.78 [-2.56, -1.00]**
Heterogeneity: Tau² = 0.00; Chi² = 0.29, df = 1 (P = 0.59); I² = 0%								
Test for overall effect: Z = 4.47 (P < 0.00001)								
1.1.2 The primary outcome was not fatigue								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**592**			**373**	**95.0%**	**-0.39 [-0.56, -0.21]**
Heterogeneity: Tau² = 0.05; Chi² = 28.43, df = 18 (P = 0.06); I² = 37%								
Test for overall effect: Z = 4.24 (P < 0.0001)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: Tau² = 0.10; Chi² = 41.16, df = 20 (P = 0.004); I² = 51%								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: Chi² = 11.66, df = 1 (P = 0.0006), I² = 91.4%								

Std. Mean Difference IV, Random, 95% CI



Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	3.8%	-0.86 [-1.65, -0.07]	
Lynch et al., 2012 (M-QG; L-B)	-3.04	3.75	44	-0.62	3.02	45	12.2%	-0.71 [-1.13, -0.28]	
Da Silva et al., 2017 (A & F; L-B)	-0.6	1.46	20	0	2.46	10	4.1%	-0.32 [-1.08, 0.45]	
Carson et al., 2010 (M-Y; L-B)	-1.44	3.89	22	0.28	7.08	26	7.1%	-0.29 [-0.86, 0.28]	
Hakkinen et al., 2001 (R; L-B)	-10	28.66	11	-3	39.02	10	3.2%	-0.20 [-1.06, 0.66]	
McBeth et al., 2012 (A; L-B; Co-CBT)	-1.3	7.94	94	0.3	9.4	44	16.8%	-0.19 [-0.55, 0.17]	
Valkeinen et al., 2008 (A & R; L-B)	-4.23	22.25	13	-1.18	22.25	11	3.7%	-0.13 [-0.94, 0.67]	
Wong et al., 2018 (M-TC; L-B)	-0.2	2	17	-0.2	2.1	14	4.7%	0.00 [-0.71, 0.71]	
McBeth et al., 2012 (A; L-B)	0.4	6.07	92	0.3	7.75	44	16.7%	0.01 [-0.34, 0.37]	
Tomas-Carus et al., 2007 (A; W-B)	0.43	1.57	17	0.34	1.57	17	5.2%	0.06 [-0.62, 0.73]	
Haak et al., 2007 (M-QG; L-B)	0.43	1.18	29	0.34	1.54	28	8.5%	0.06 [-0.45, 0.58]	
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-0.5	4.07	20	-0.84	4.6	12	4.6%	0.08 [-0.64, 0.79]	
Sañudo et al., 2015 (A; L-B)	0.2	3.24	16	-0.3	4.21	12	4.2%	0.13 [-0.62, 0.88]	
Wigers et al., 1996 (A; L-B)	10	44.37	16	2	50	17	5.0%	0.16 [-0.52, 0.85]	
Total (95% CI)			**431**			**300**	**100.0%**	**-0.17 [-0.32, -0.01]**	

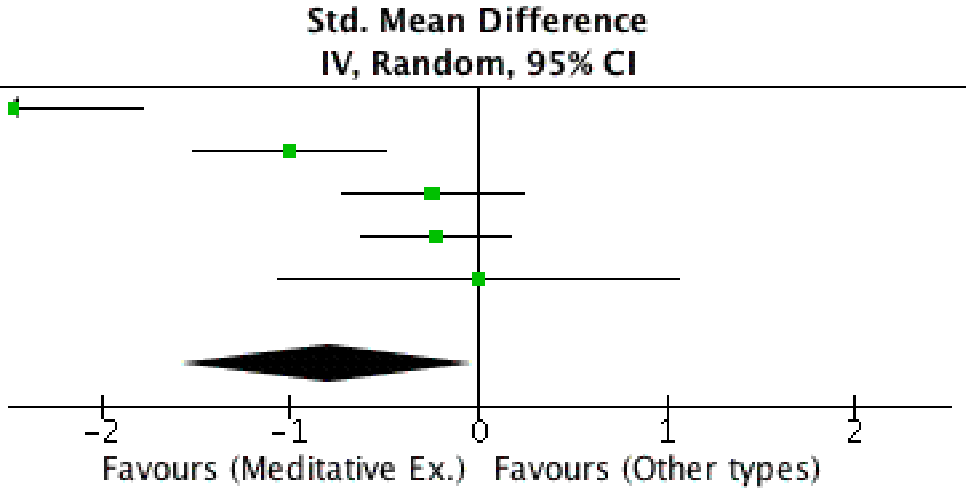


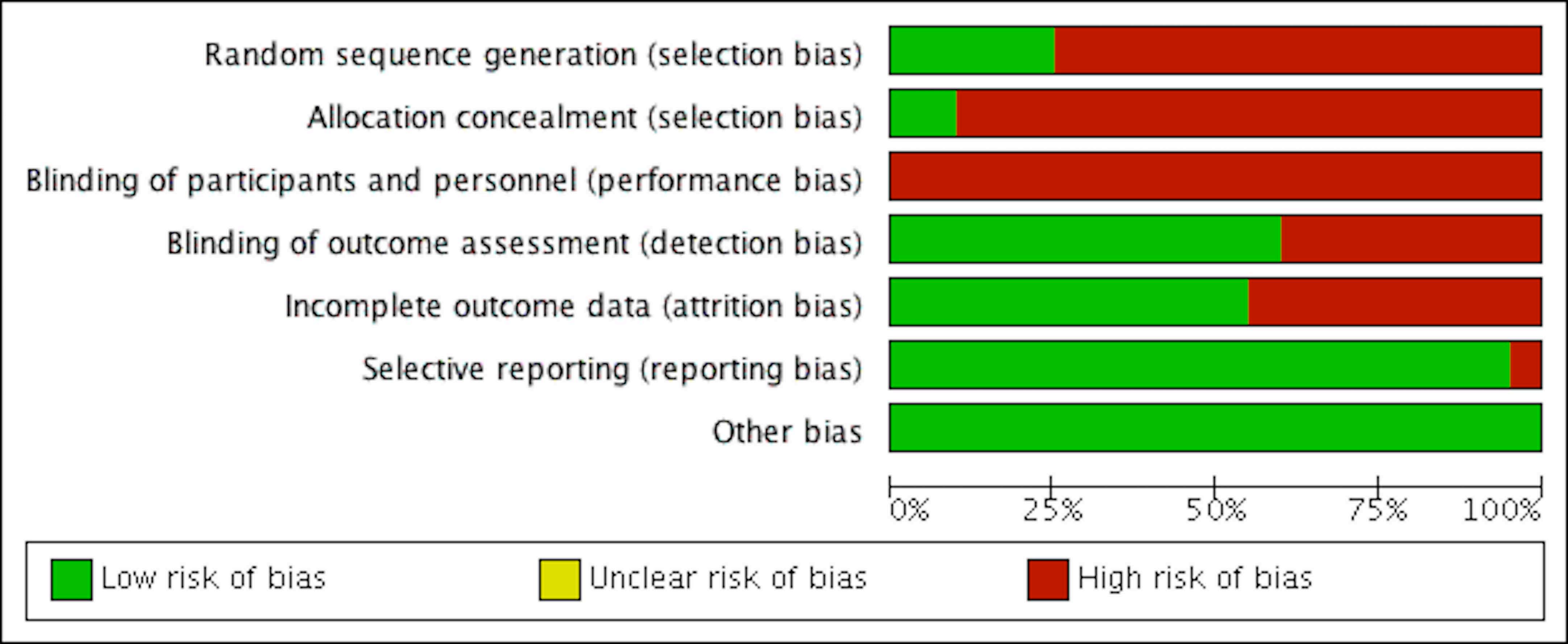
Heterogeneity: Tau² = 0.00; Chi² = 13.68, df = 13 (P = 0.40); I² = 5%
Test for overall effect: Z = 2.09 (P = 0.04)

Study or Subgroup	Meditative exercise Mean	SD	Total	Other types of exercise Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Lopez-Rodriguez et al., 2013 (AqBD vs F)	-7.51	2.38	29	-0.3	3.28	30	19.6%	-2.48 [-3.16, -1.79]
Wang et al., 2010 (TC vs F)	-3.6	3.5	33	-0.7	2	33	21.1%	-1.01 [-1.52, -0.49]
Calandre et al., 2009 (AC vs F)	-0.91	2.62	32	-0.29	2.43	34	21.3%	-0.24 [-0.73, 0.24]
Wang et al., 2018 (TC vs A)	-2.1	5.1	36	-1.1	4.09	75	21.9%	-0.22 [-0.62, 0.17]
Norregaard et al., 1997 (BA vs A)	0	2.3	11	0	2	5	16.2%	0.00 [-1.06, 1.06]
Total (95% CI)			**141**			**177**	**100.0%**	**-0.80 [-1.57, -0.02]**

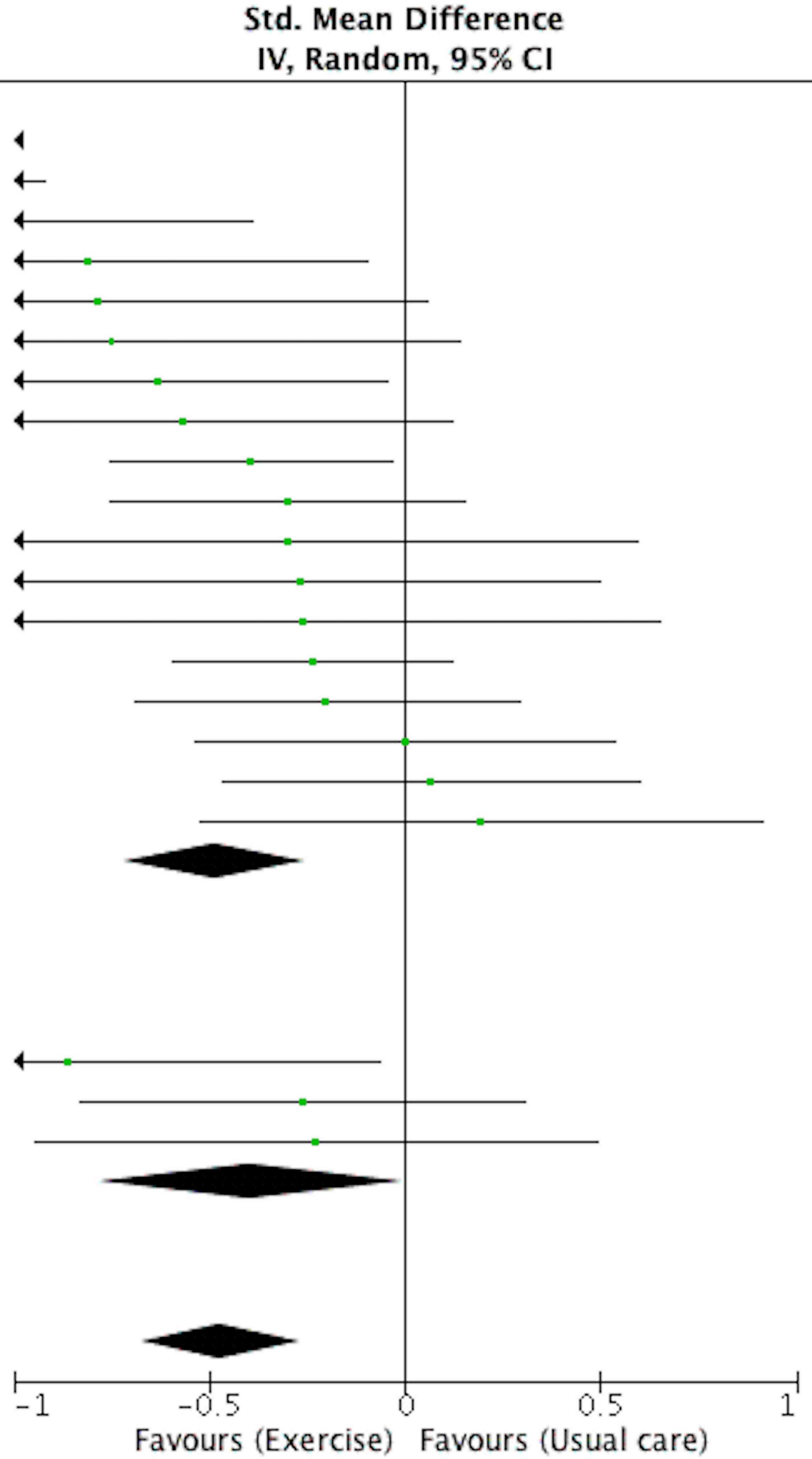
Heterogeneity: $Tau^2 = 0.67$; $Chi^2 = 37.39$, df = 4 ($P < 0.00001$); $I^2 = 89\%$

Test for overall effect: $Z = 2.02$ ($P = 0.04$)

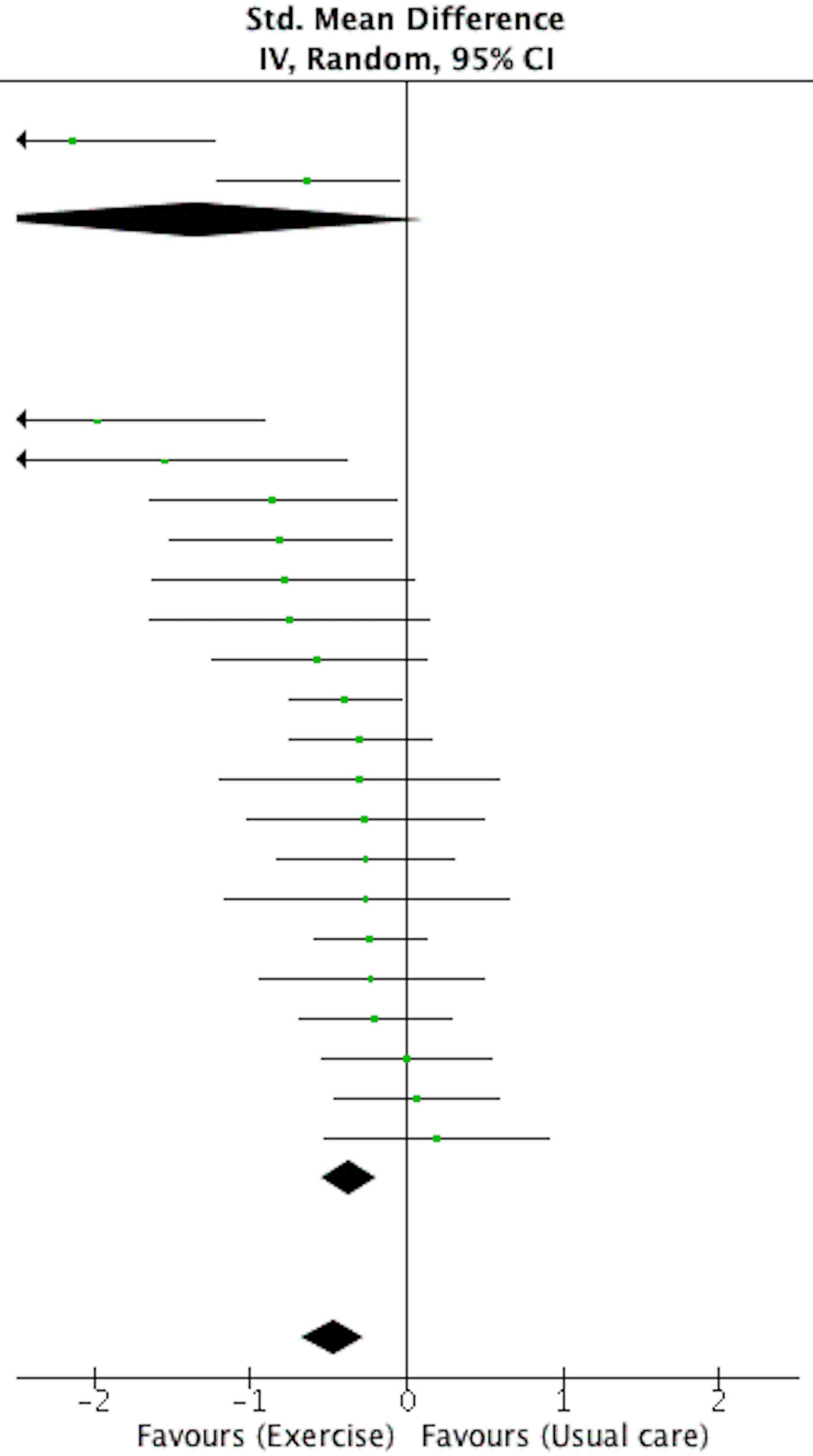



Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Other bias
0%
25%
50%
75%
100%
Low risk of bias
Unclear risk of bias
High risk of bias

Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Type of intervention (only exercise)								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**552**			**339**	**85.9%**	**-0.49 [-0.72, -0.26]**
Heterogeneity: $Tau^2 = 0.13$; $Chi^2 = 39.41$, df = 17 (P = 0.002); $I^2 = 57\%$								
Test for overall effect: Z = 4.16 (P < 0.0001)								
1.1.2 Remaining studies (mixed interventions)								
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Subtotal (95% CI)			**60**			**52**	**14.1%**	**-0.39 [-0.78, -0.01]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.74$, df = 2 (P = 0.42); $I^2 = 0\%$								
Test for overall effect: Z = 1.99 (P = 0.05)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 0.17$, df = 1 (P = 0.68), $I^2 = 0\%$								

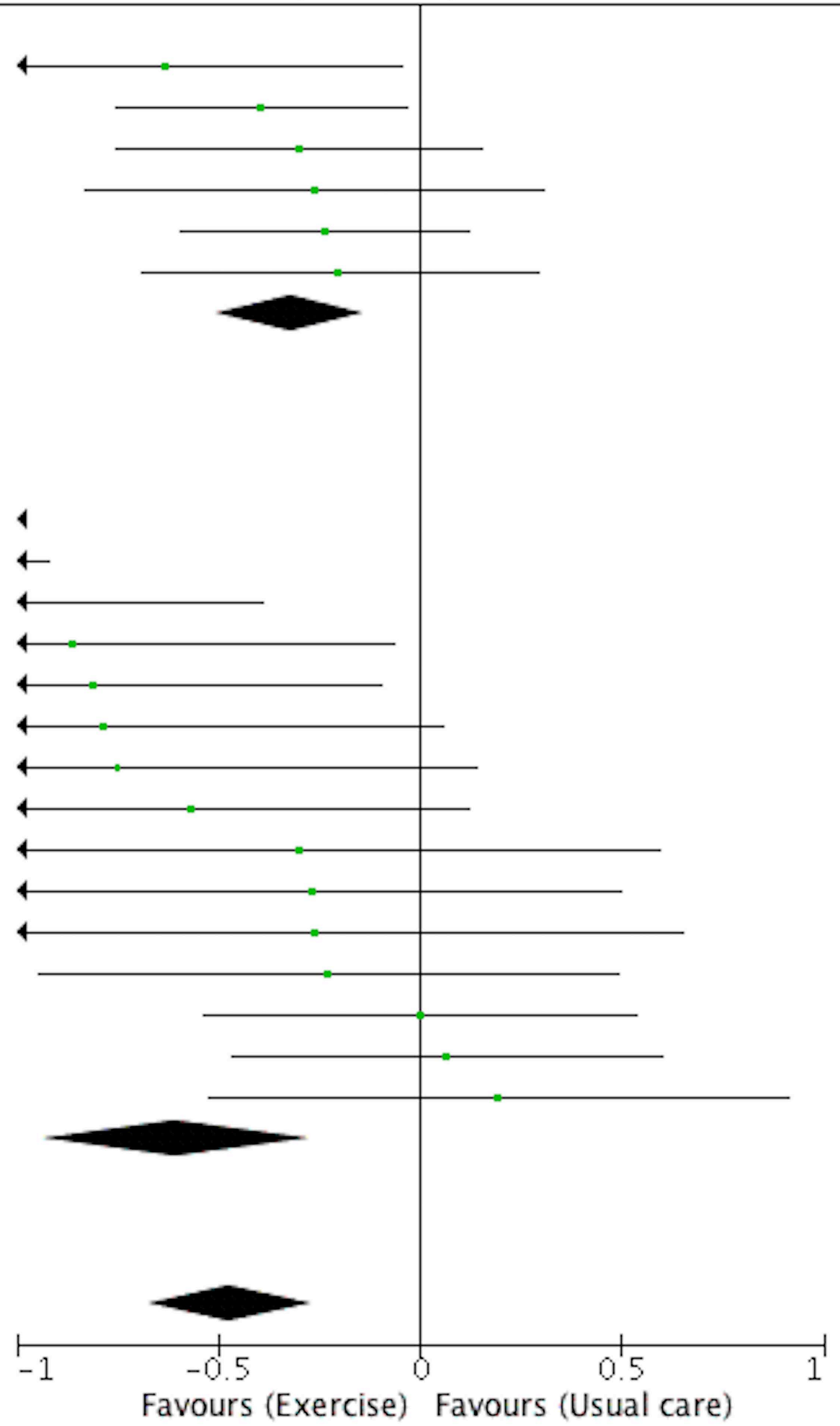


Study or Subgroup	Favours (Exercise) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Type of exercise (meditative exercise programs))								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Subtotal (95% CI)			**39**			**40**	**8.9%**	**-1.34 [-2.81, 0.13]**
Heterogeneity: $Tau^2 = 0.98$; $Chi^2 = 7.49$, df = 1 (P = 0.006); $I^2 = 87\%$								
Test for overall effect: Z = 1.78 (P = 0.07)								
1.1.2 Remaining studies (i.e., did not involve meditative exercise)								
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**573**			**351**	**91.1%**	**-0.37 [-0.55, -0.20]**
Heterogeneity: $Tau^2 = 0.04$; $Chi^2 = 25.88$, df = 18 (P = 0.10); $I^2 = 30\%$								
Test for overall effect: Z = 4.19 (P < 0.0001)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 1.63$, df = 1 (P = 0.20), $I^2 = 38.7\%$								



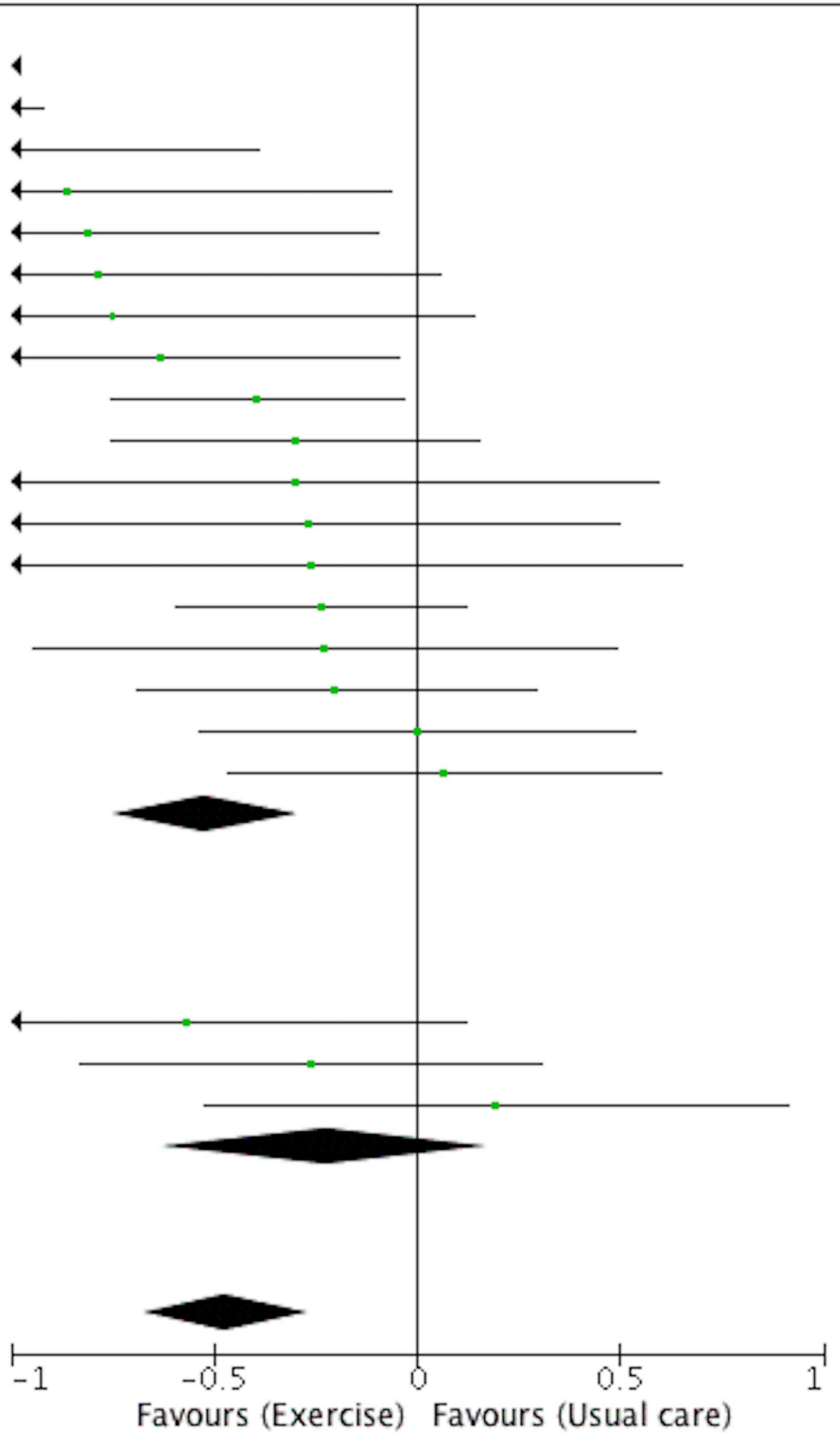
Study or Subgroup	Favours (Exercise) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Sample size (at least, 20 participants/group)								
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Subtotal (95% CI)			**306**			**207**	**39.6%**	**-0.32 [-0.50, -0.14]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.74$, df = 5 (P = 0.88); $I^2 = 0\%$								
Test for overall effect: Z = 3.47 (P = 0.0005)								
1.1.2 Remaining studies (group(s) with fewer than 20 participants)								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**306**			**184**	**60.4%**	**-0.60 [-0.93, -0.27]**
Heterogeneity: $Tau^2 = 0.26$; $Chi^2 = 38.00$, df = 14 (P = 0.0005); $I^2 = 63\%$								
Test for overall effect: Z = 3.59 (P = 0.0003)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 2.17$, df = 1 (P = 0.14), $I^2 = 54.0\%$								

Std. Mean Difference
IV, Random, 95% CI

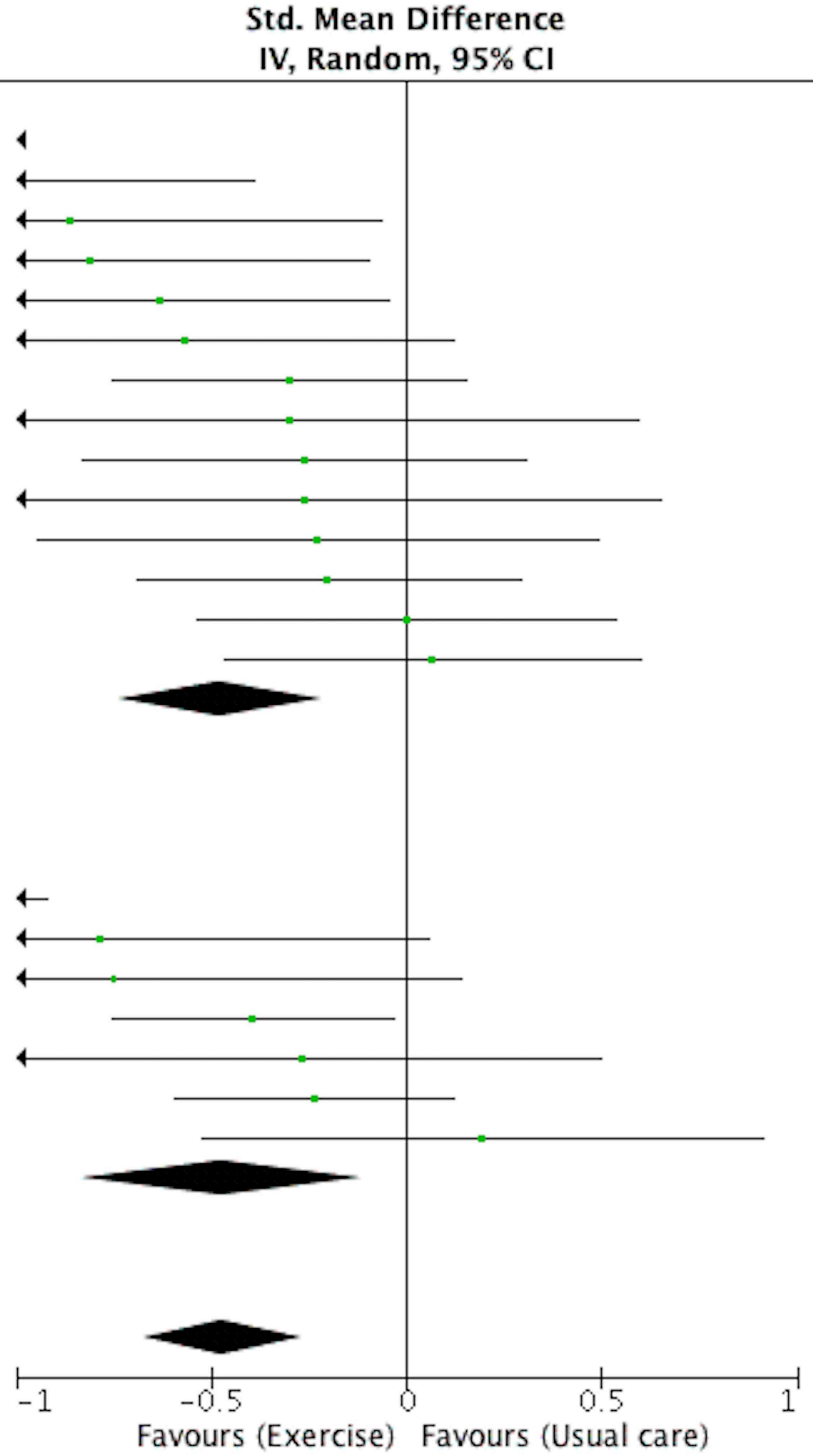


	Favours (Exercise)			Usual care (Control)				Std. Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
1.1.1 Land-based exercise								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Subtotal (95% CI)			**560**			**329**	**85.2%**	**-0.52 [-0.75, -0.29]**
Heterogeneity: $Tau^2 = 0.12$; $Chi^2 = 38.00$, df = 17 (P = 0.002); $I^2 = 55\%$								
Test for overall effect: Z = 4.48 (P < 0.00001)								
1.1.2 Remaining studies (i.e., water-based exercise)								
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**52**			**62**	**14.8%**	**-0.23 [-0.62, 0.17]**
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 2.26$, df = 2 (P = 0.32); $I^2 = 12\%$								
Test for overall effect: Z = 1.11 (P = 0.27)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 1.59$, df = 1 (P = 0.21), $I^2 = 37.1\%$								

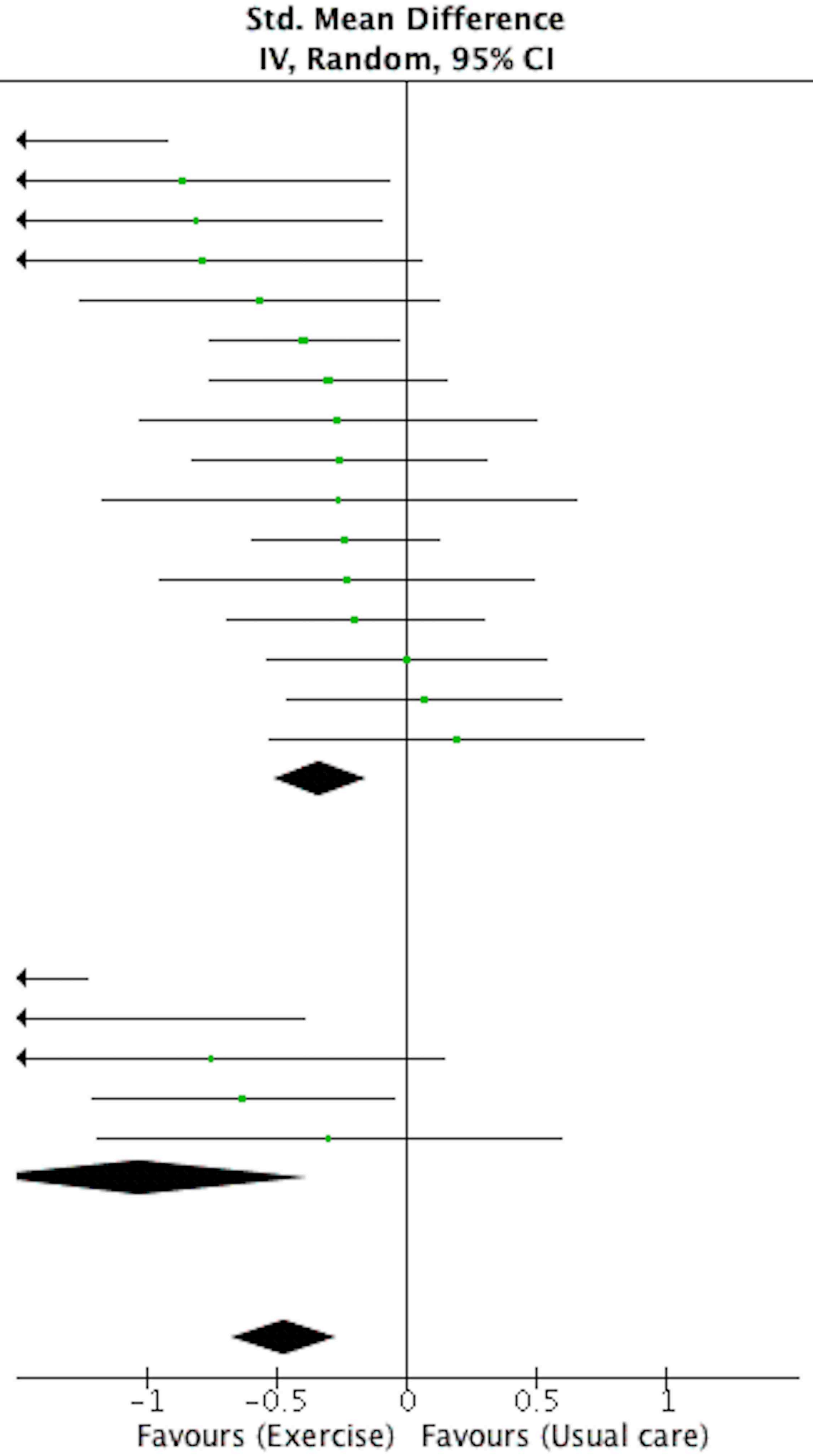
Std. Mean Difference
IV, Random, 95% CI



Study or Subgroup	Favours (Exercise) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Training intensity (from low to moderate)								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Subtotal (95% CI)			**355**			**247**	**66.2%**	**-0.48 [-0.74, -0.22]**
Heterogeneity: $Tau^2 = 0.12$; $Chi^2 = 27.69$, df = 13 (P = 0.01); $I^2 = 53\%$								
Test for overall effect: Z = 3.61 (P = 0.0003)								
1.1.2 Training intensity (from moderate to high)								
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**257**			**144**	**33.8%**	**-0.47 [-0.83, -0.12]**
Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 13.46$, df = 6 (P = 0.04); $I^2 = 55\%$								
Test for overall effect: Z = 2.60 (P = 0.009)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 0.00$, df = 1 (P = 0.99), $I^2 = 0\%$								

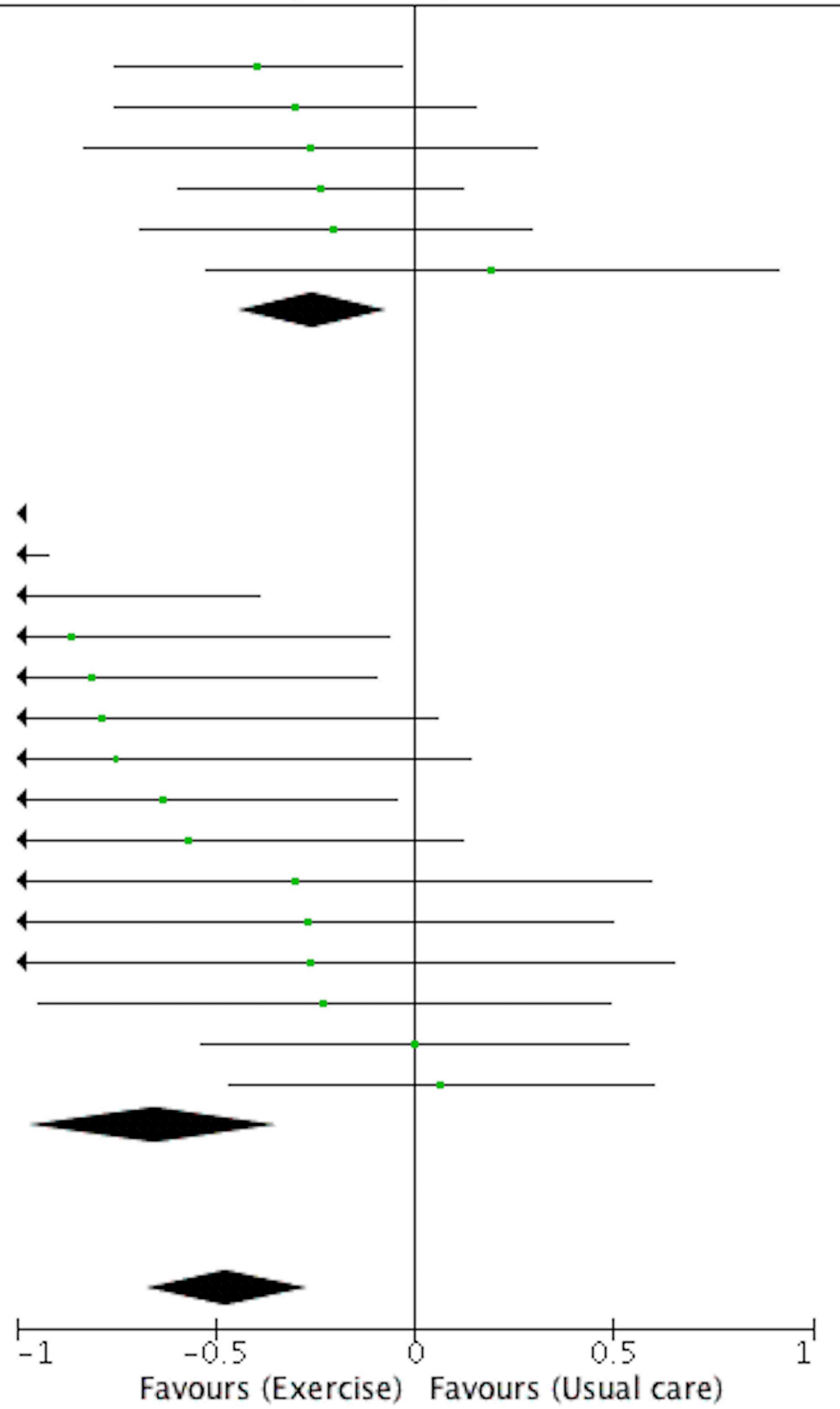


Study or Subgroup	Favours (Exercise) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Type of exercise (involved aerobic)								
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**538**			**326**	**82.0%**	**-0.33 [-0.51, -0.16]**
Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 20.76$, df = 15 (P = 0.14); $I^2 = 28\%$								
Test for overall effect: Z = 3.72 (P = 0.0002)								
1.1.2 Remaining studies (i.e., did not involve aerobic exercise)								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Subtotal (95% CI)			**74**			**65**	**18.0%**	**-1.02 [-1.66, -0.39]**
Heterogeneity: $Tau^2 = 0.33$; $Chi^2 = 10.96$, df = 4 (P = 0.03); $I^2 = 64\%$								
Test for overall effect: Z = 3.15 (P = 0.002)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 4.22$, df = 1 (P = 0.04), $I^2 = 76.3\%$								



Study or Subgroup	Favours (Exercise) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Length of the intervention (at least, 24 weeks)								
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**299**			**196**	**38.6%**	**-0.26 [-0.44, -0.07]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 2.16$, df = 5 (P = 0.83); $I^2 = 0\%$								
Test for overall effect: Z = 2.72 (P = 0.006)								
1.1.2 Remaining studies (i.e., interventions lasted less than 24 weeks)								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Subtotal (95% CI)			**313**			**195**	**61.4%**	**-0.65 [-0.96, -0.35]**
Heterogeneity: $Tau^2 = 0.21$; $Chi^2 = 34.40$, df = 14 (P = 0.002); $I^2 = 59\%$								
Test for overall effect: Z = 4.15 (P < 0.0001)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 4.71$, df = 1 (P = 0.03), $I^2 = 78.8\%$								

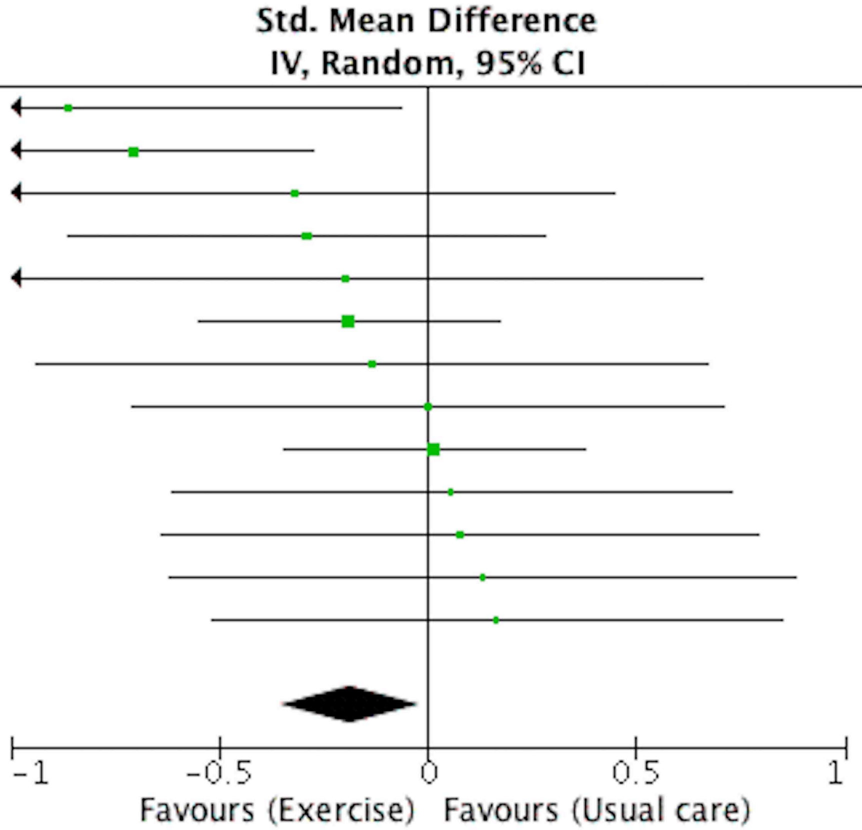
Std. Mean Difference
IV, Random, 95% CI



Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.2%	-0.86 [-1.65, -0.07]
Lynch et al., 2012 (M-QG; L-B)	-3.04	3.75	44	-0.62	3.02	45	13.2%	-0.71 [-1.13, -0.28]
Da Silva et al., 2017 (A & F; L-B)	-0.6	1.46	20	0	2.46	10	4.5%	-0.32 [-1.08, 0.45]
Carson et al., 2010 (M-Y; L-B)	-1.44	3.89	22	0.28	7.08	26	7.9%	-0.29 [-0.86, 0.28]
Hakkinen et al., 2001 (R; L-B)	-10	28.66	11	-3	39.02	10	3.6%	-0.20 [-1.06, 0.66]
McBeth et al., 2012 (A; L-B; Co-CBT)	-1.3	7.94	94	0.3	9.4	44	18.0%	-0.19 [-0.55, 0.17]
Valkeinen et al., 2008 (A & R; L-B)	-4.23	22.25	13	-1.18	22.25	11	4.1%	-0.13 [-0.94, 0.67]
Wong et al., 2018 (M-TC; L-B)	-0.2	2	17	-0.2	2.1	14	5.2%	0.00 [-0.71, 0.71]
McBeth et al., 2012 (A; L-B)	0.4	6.07	92	0.3	7.75	44	18.0%	0.01 [-0.34, 0.37]
Tomas-Carus et al., 2007 (A; W-B)	0.43	1.57	17	0.34	1.57	17	5.8%	0.06 [-0.62, 0.73]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-0.5	4.07	20	-0.84	4.6	12	5.1%	0.08 [-0.64, 0.79]
Sañudo et al., 2015 (A; L-B)	0.2	3.24	16	-0.3	4.21	12	4.7%	0.13 [-0.62, 0.88]
Wigers et al., 1996 (A; L-B)	10	44.37	16	2	50	17	5.6%	0.16 [-0.52, 0.85]
Total (95% CI)			**402**			**272**	**100.0%**	**-0.19 [-0.35, -0.02]**

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 12.84$, df = 12 (P = 0.38); $I^2 = 7\%$

Test for overall effect: Z = 2.22 (P = 0.03)

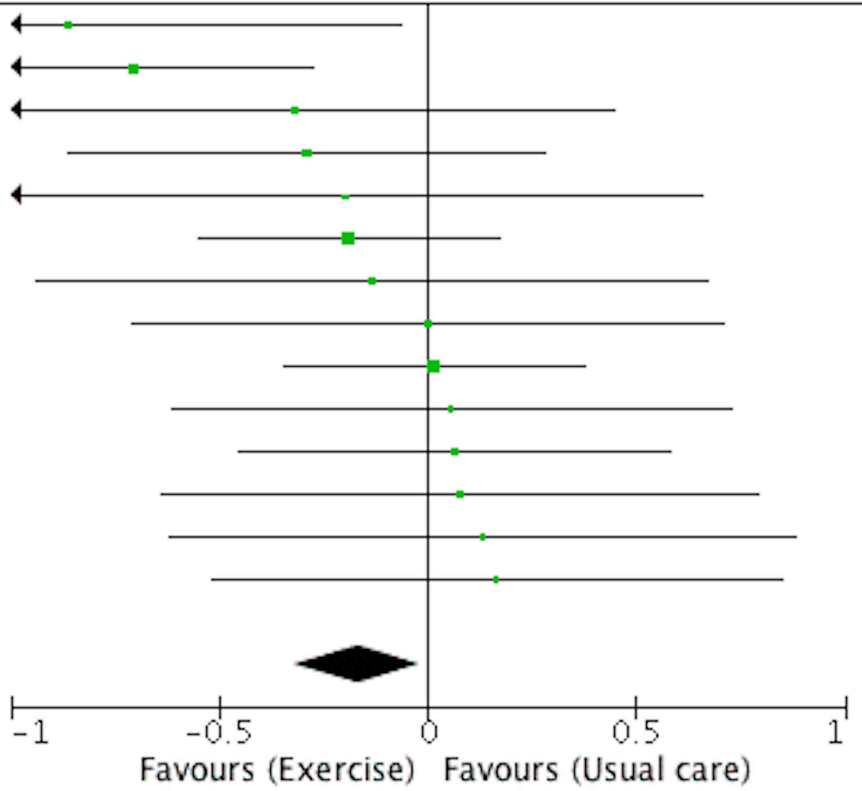


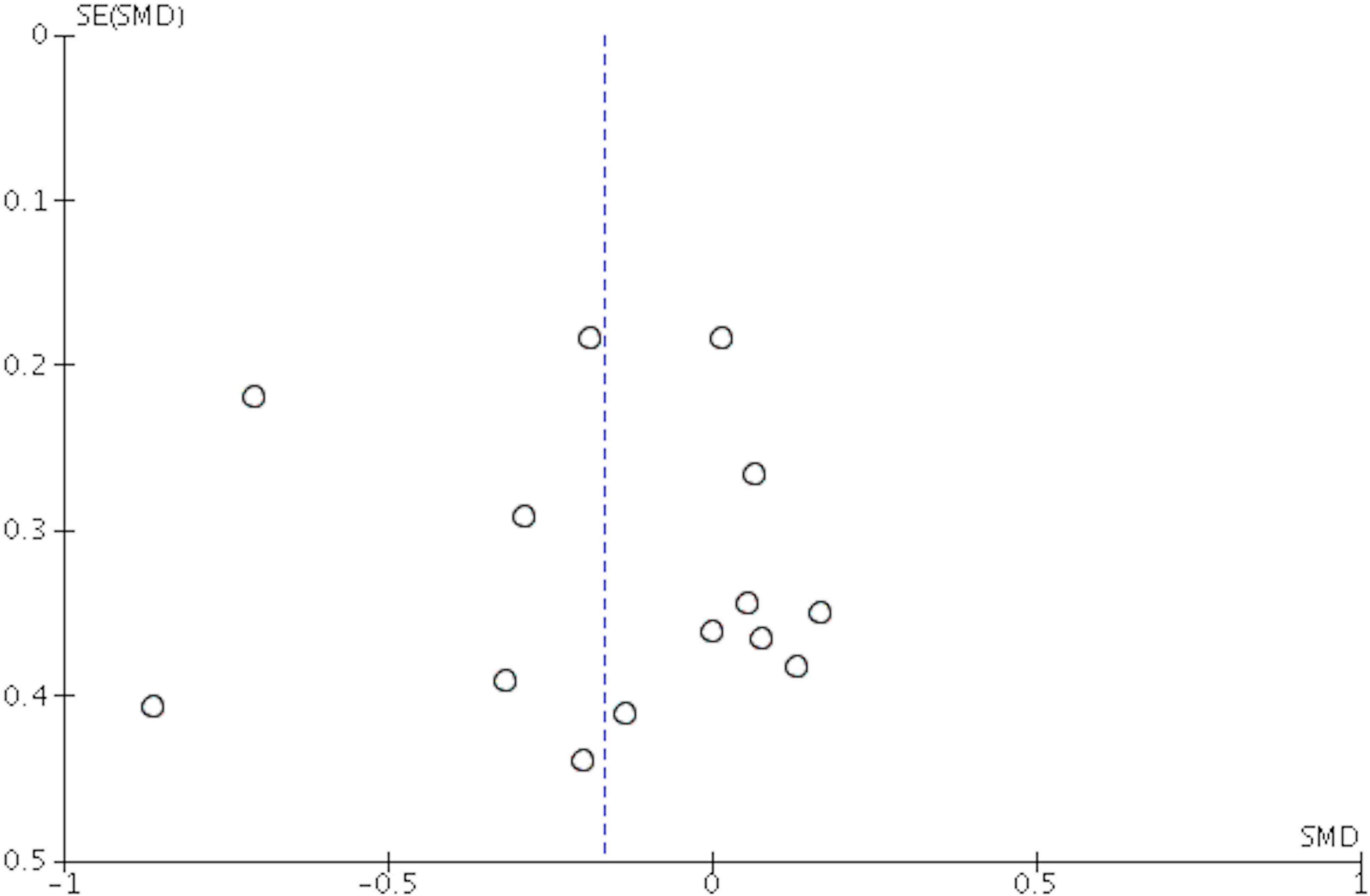
Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	3.6%	-0.86 [-1.65, -0.07]
Lynch et al., 2012 (M-QG; L-B)	-3.04	3.75	44	-0.62	3.02	45	12.4%	-0.71 [-1.13, -0.28]
Da Silva et al., 2017 (A & F; L-B)	-0.6	1.46	20	0	2.46	10	3.9%	-0.32 [-1.08, 0.45]
Carson et al., 2010 (M-Y; L-B)	-1.44	3.89	22	0.28	7.08	26	7.0%	-0.29 [-0.86, 0.28]
Hakkinen et al., 2001 (R; L-B)	-10	28.66	11	-3	39.02	10	3.1%	-0.20 [-1.06, 0.66]
McBeth et al., 2012 (A; L-B; Co-CBT)	-1.3	7.94	94	0.3	9.4	44	17.7%	-0.19 [-0.55, 0.17]
Valkeinen et al., 2008 (A & R; L-B)	-4.23	22.25	13	-1.18	22.25	11	3.5%	-0.13 [-0.94, 0.67]
Wong et al., 2018 (M-TC; L-B)	-0.2	2	17	-0.2	2.1	14	4.5%	0.00 [-0.71, 0.71]
McBeth et al., 2012 (A; L-B)	0.4	6.07	92	0.3	7.75	44	17.6%	0.01 [-0.34, 0.37]
Tomas-Carus et al., 2007 (A; W-B)	0.43	1.57	17	0.34	1.57	17	5.0%	0.06 [-0.62, 0.73]
Haak et al., 2007 (M-QG; L-B)	0.43	1.18	29	0.34	1.54	28	8.4%	0.06 [-0.45, 0.58]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-0.5	4.07	20	-0.84	4.6	12	4.4%	0.08 [-0.64, 0.79]
Sañudo et al., 2015 (A; L-B)	0.2	3.24	16	-0.3	4.21	12	4.0%	0.13 [-0.62, 0.88]
Wigers et al., 1996 (A; L-B)	10	44.37	16	2	50	17	4.9%	0.16 [-0.52, 0.85]
Total (95% CI)			**431**			**300**	**100.0%**	**-0.17 [-0.32, -0.02]**

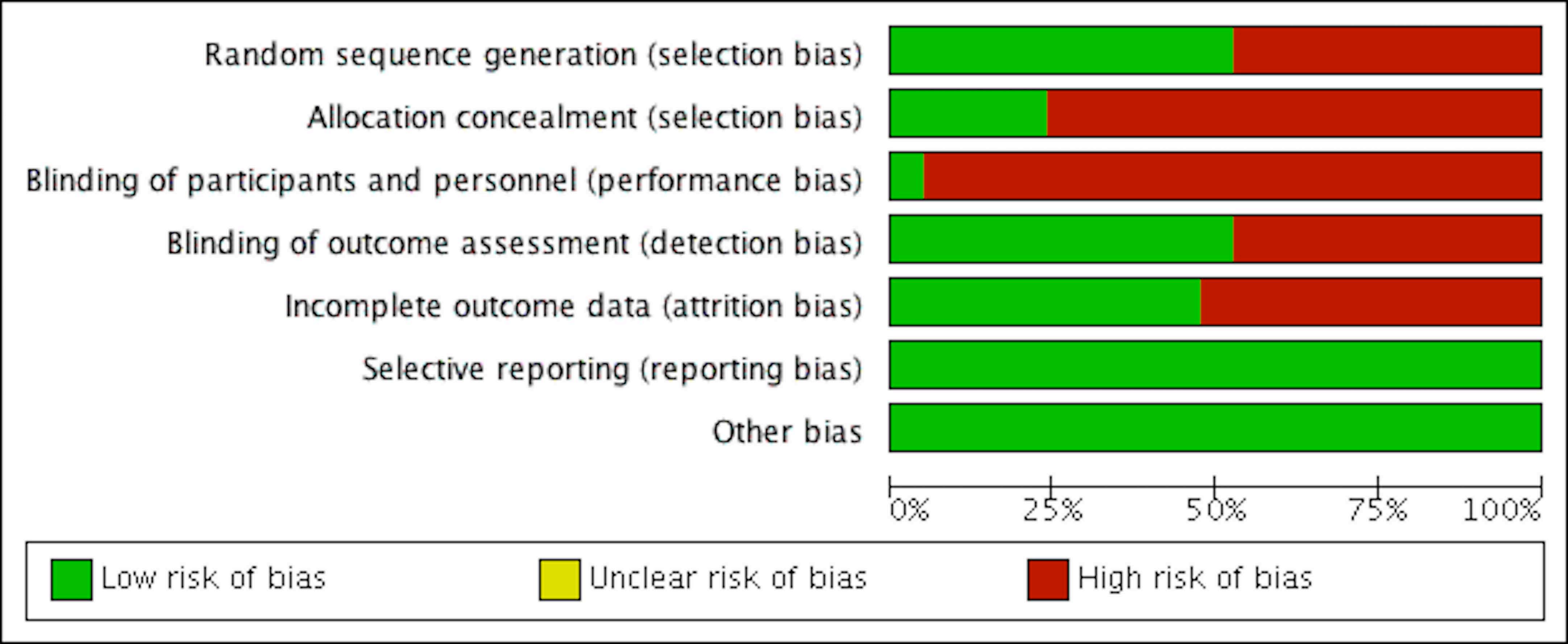
Heterogeneity: $Chi^2 = 13.68$, df = 13 (P = 0.40); $I^2 = 5\%$

Test for overall effect: Z = 2.17 (P = 0.03)

Std. Mean Difference IV, Fixed, 95% CI

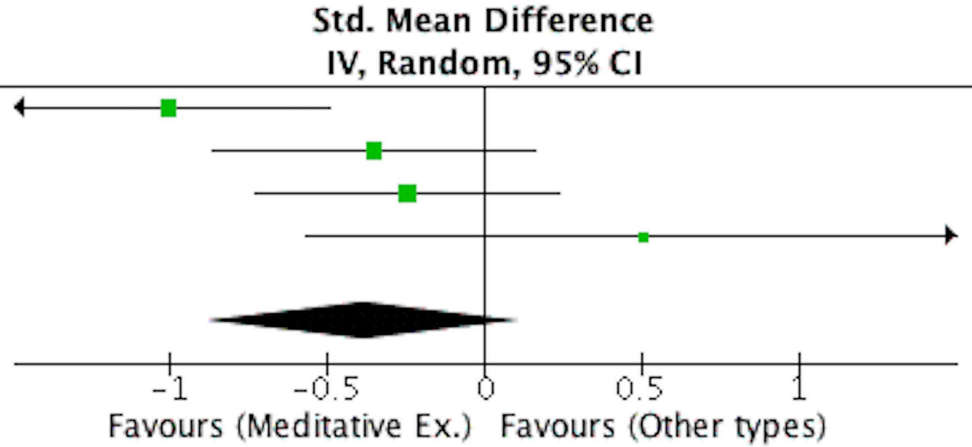



SE(SMD)
0
0.1
0.2
0.3
0.4
0.5
SMD
-1
-0.5
0
0.5
1


Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Other bias
0%
25%
50%
75%
100%
Low risk of bias
Unclear risk of bias
High risk of bias

Study or Subgroup	Meditative exercise Mean	Meditative exercise SD	Meditative exercise Total	Other types of exercise Mean	Other types of exercise SD	Other types of exercise Total	Weight	Std. Mean Difference IV, Random, 95% CI
Wang et al., 2010 (TC vs F)	-3.6	3.5	33	-0.7	2	33	28.4%	-1.01 [-1.52, -0.49]
Lopez-Rodriguez et al., 2013 (AqBD vs F)	-0.93	2.36	29	-0.2	1.7	30	28.3%	-0.35 [-0.87, 0.16]
Calandre et al., 2009 (AC vs F)	-0.91	2.62	32	-0.29	2.43	34	29.4%	-0.24 [-0.73, 0.24]
Norregaard et al., 1997 (BA vs A)	0	1.4	11	-1	2.7	5	13.9%	0.51 [-0.57, 1.58]
Total (95% CI)			**105**			**102**	**100.0%**	**-0.39 [-0.88, 0.11]**

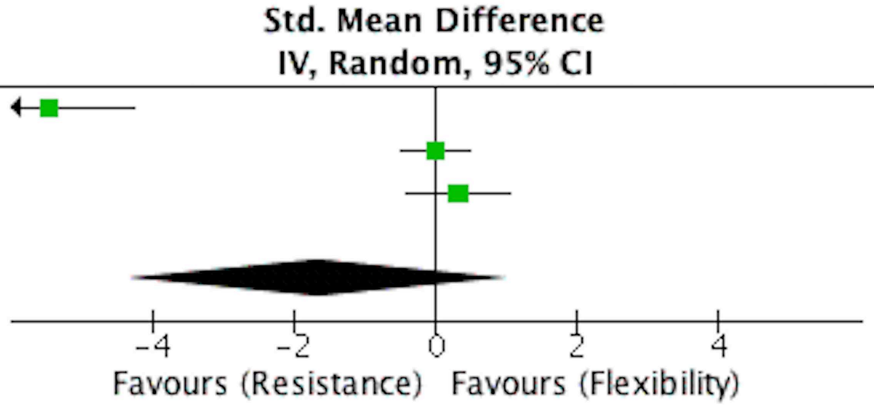
Heterogeneity: $Tau^2 = 0.16$; $Chi^2 = 8.37$, df = 3 (P = 0.04); $I^2 = 64\%$
Test for overall effect: Z = 1.53 (P = 0.13)



Study or Subgroup	Resistance Mean	Resistance SD	Resistance Total	Flexibility Mean	Flexibility SD	Flexibility Total	Weight	Std. Mean Difference IV, Random, 95% CI
Jones et al., 2002	−2.43	0.49	28	0.68	0.63	28	32.3%	−5.43 [−6.60, −4.27]
Gavi et al., 2014	−16	27.58	35	−16.13	29.06	31	34.1%	0.00 [−0.48, 0.49]
Assumpçao et al., 2018	−1.93	3.92	16	−3.1	2.91	14	33.6%	0.33 [−0.40, 1.05]
Total (95% CI)			**79**			**73**	**100.0%**	**−1.64 [−4.31, 1.02]**

Heterogeneity: $Tau^2 = 5.36$; $Chi^2 = 77.77$, df = 2 ($P < 0.00001$); $I^2 = 97\%$

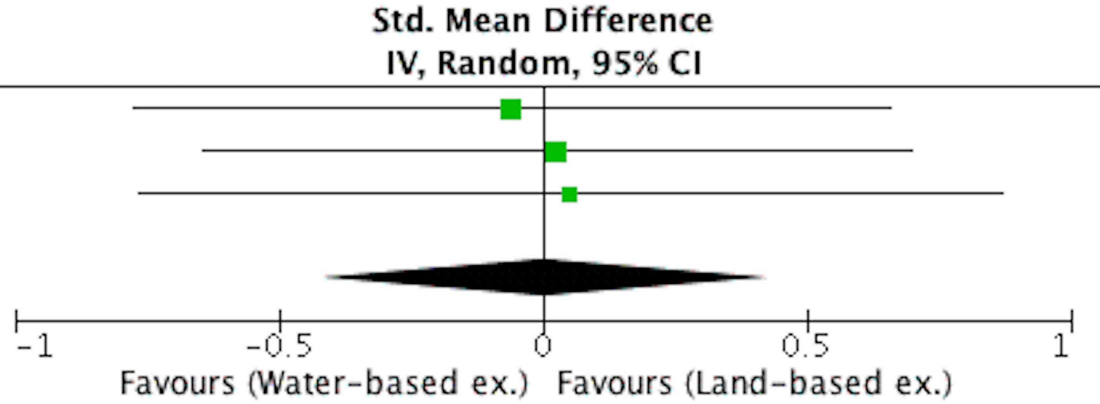
Test for overall effect: $Z = 1.21$ ($P = 0.23$)



	Water-based exercise			Land-based exercise				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Fernandes et al., 2016	-19	49.12	14	-16	50	16	34.4%	-0.06 [-0.78, 0.66]
Jentoft et al., 2001	-1.4	3.39	18	-1.5	3.71	16	39.1%	0.03 [-0.65, 0.70]
Vitorino et al., 2006	-25.8	26.96	11	-27.2	23.51	12	26.5%	0.05 [-0.76, 0.87]
Total (95% CI)			**43**			**44**	**100.0%**	**0.00 [-0.42, 0.43]**

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.05$, df = 2 (P = 0.98); $I^2 = 0\%$

Test for overall effect: Z = 0.02 (P = 0.98)



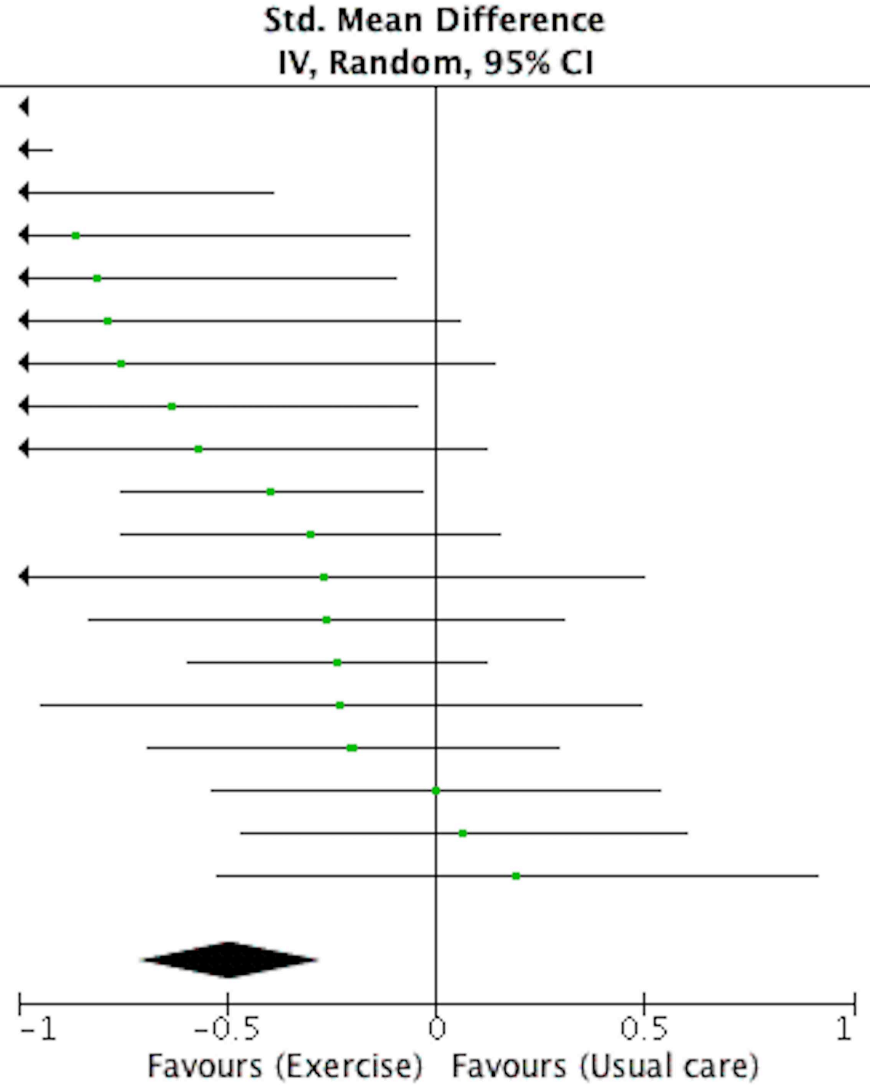
Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Alentorn-Geli et al., 2008	−	−	−	+	+	−	+
Assumpção et al., 2018	−	−	−	−	−	+	+
Carson et al., 2010	+	+	−	+	−	+	+
Collado-Mateo et al., 2016	+	−	−	+	−	+	+
Da Silva et al., 2017	−	−	−	+	+	+	+
Etnier et al., 2009	−	−	−	−	+	+	+
Gianotti et al., 2014	+	−	−	−	−	+	+
Haak et al., 2007	−	−	−	−	−	+	+
Hakkinen et al., 2001	−	−	−	−	+	+	+
Lynch et al., 2012	−	+	−	−	−	+	+
Mannerkorpi et al., 2000	−	−	−	−	+	+	+
McBeth et al., 2012	−	−	−	+	−	+	+
Sañudo et al., 2015	−	−	−	+	+	+	+
Schachter et al., 2003	−	−	−	+	+	+	+
Tomas-Carus et al., 2007	−	−	−	+	+	+	+
Tomas Carus et al., 2008	−	−	−	+	+	+	+
Valkeinen et al., 2008	−	−	−	−	+	+	+
Van Santen et al., 2002	−	−	−	+	−	+	+
Wigers et al., 1996	+	−	−	+	+	+	+
Wong et al., 2018	+	−	−	+	−	+	+

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Assumpçao et al., 2018	−	−	−	−	−	+	+
Bircan et al., 2008	−	−	−	−	+	+	+
Calandre et al., 2009	+	−	−	−	−	+	+
da Silva et al., 2018	−	−	−	+	+	+	+
Demir-Gocmen et al., 2013	−	−	−	−	+	+	+
Fernandes et al., 2016	+	+	−	+	+	+	+
Gavi et al., 2014	−	−	−	+	−	+	+
Genc et al., 2015	+	−	−	−	+	+	+
Jentoft et al., 2001	−	−	−	−	+	+	+
Jones et al., 2002	+	−	−	+	+	+	+
Kendall et al., 2000	+	−	−	−	−	+	+
Lopez-Pousa et al., 2015	+	−	+	−	−	+	+
Lopez-Rodriguez et al, 2013	+	−	−	+	+	+	+
Mannerkorpi et al., 2010	+	+	−	+	−	+	+
McBeth et al., 2012	−	−	−	+	−	+	+
Norregaard et al., 1997	+	−	−	−	−	+	+
Schachter et al., 2003	−	−	−	+	+	+	+
van Santen et al., 2002	−	−	−	−	+	+	+
Vitorino et al., 2006	−	+	−	+	−	+	+
Wang et al., 2010	+	+	−	+	−	+	+
Wang et al., 2018	+	+	−	+	−	+	+

Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.7%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	3.0%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.6%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.3%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.9%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	4.1%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.8%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.9%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	5.1%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	8.0%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	7.1%	-0.30 [-0.75, 0.15]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.5%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	6.0%	-0.26 [-0.83, 0.31]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	8.0%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.8%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.7%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	6.3%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.4%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.8%	0.19 [-0.52, 0.91]
Total (95% CI)			**582**			**377**	**100.0%**	**-0.49 [-0.71, -0.27]**

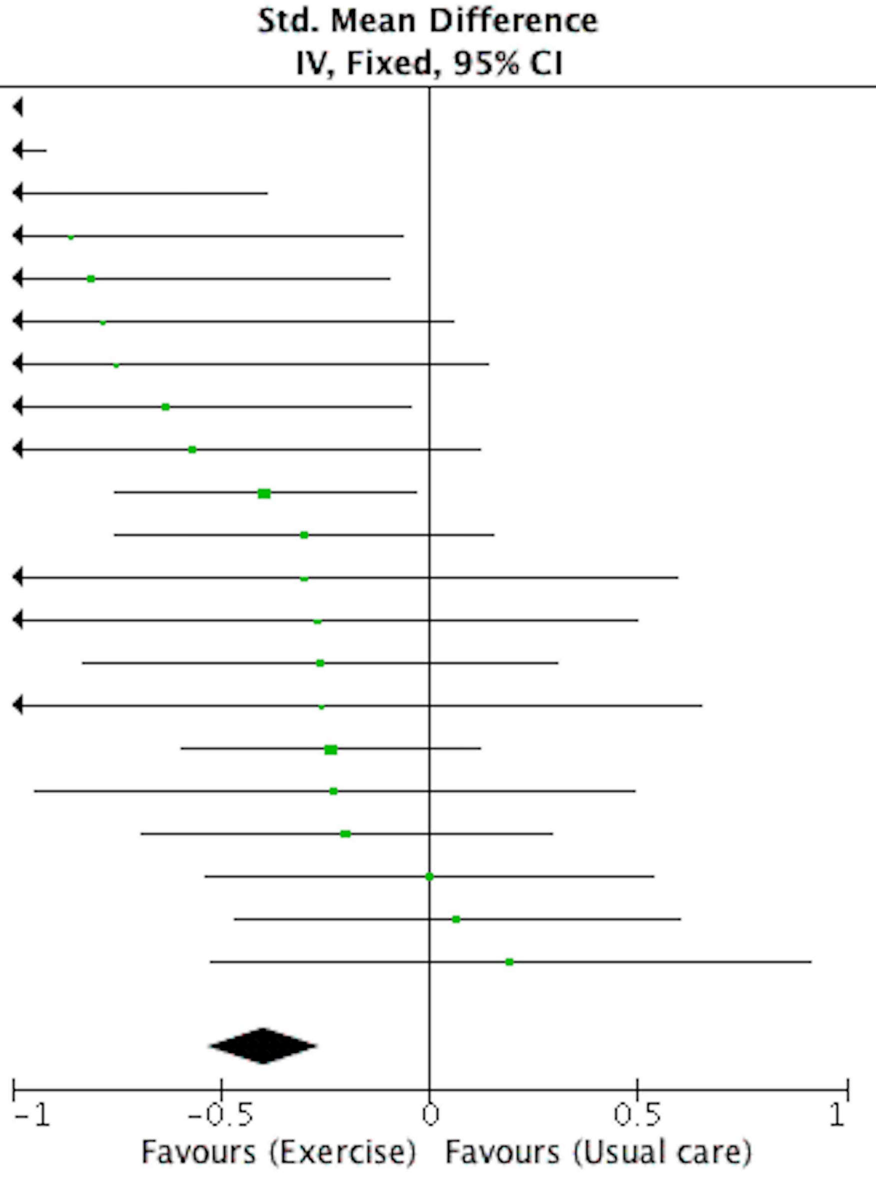
Heterogeneity: $Tau^2 = 0.12$; $Chi^2 = 41.02$, $df = 18$ ($P = 0.002$); $I^2 = 56\%$

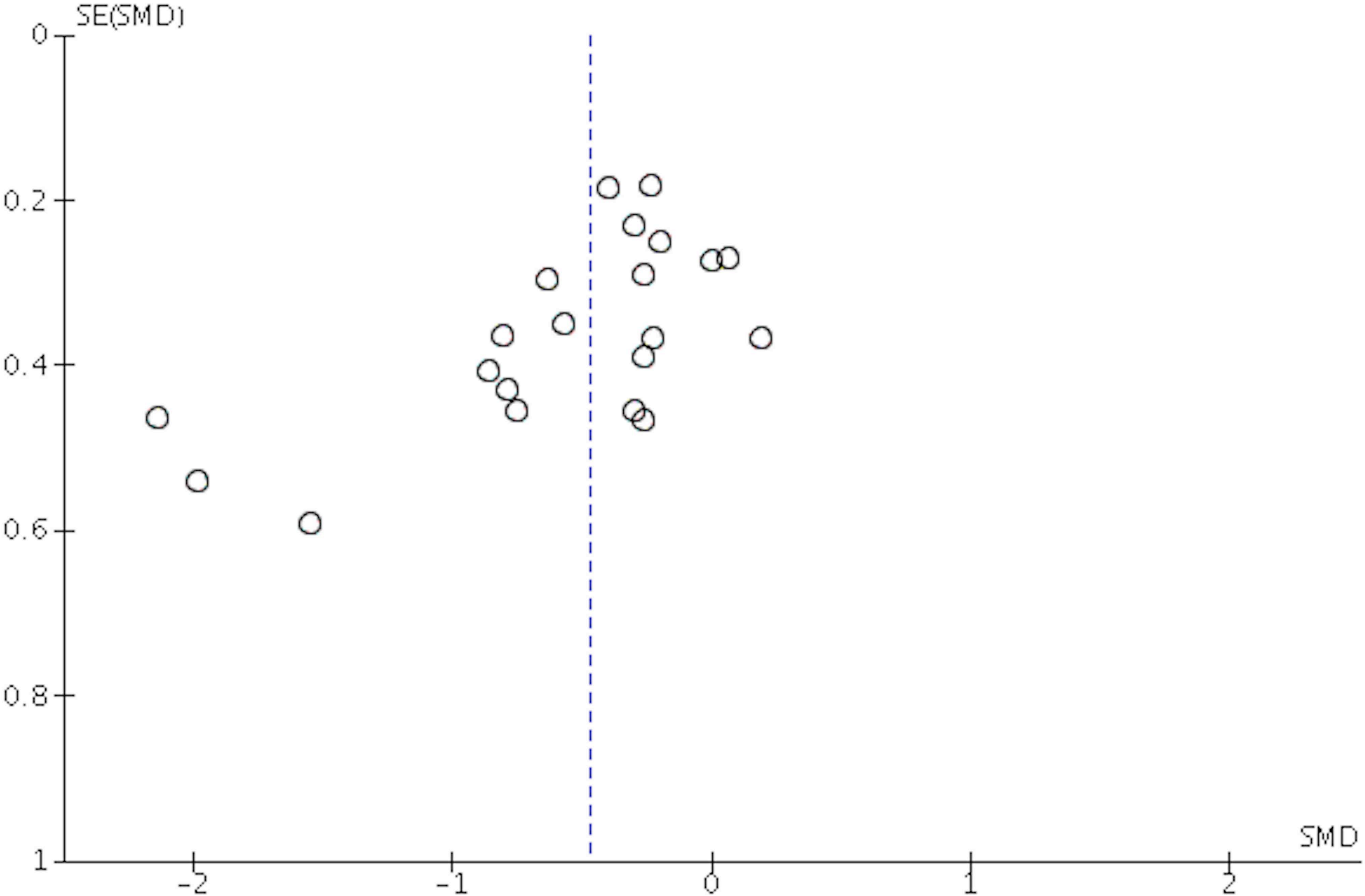
Test for overall effect: $Z = 4.42$ ($P < 0.00001$)



Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	2.1%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	1.6%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	1.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	2.8%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	3.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	2.5%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	2.2%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.2%	-0.63 [-1.21, -0.05]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	3.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	13.4%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	8.6%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	2.2%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	3.0%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.5%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	2.1%	-0.26 [-1.17, 0.65]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	13.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	3.4%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	7.3%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	6.1%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.3%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	3.4%	0.19 [-0.52, 0.91]
Total (95% CI)			**612**			**391**	**100.0%**	**-0.40 [-0.53, -0.26]**

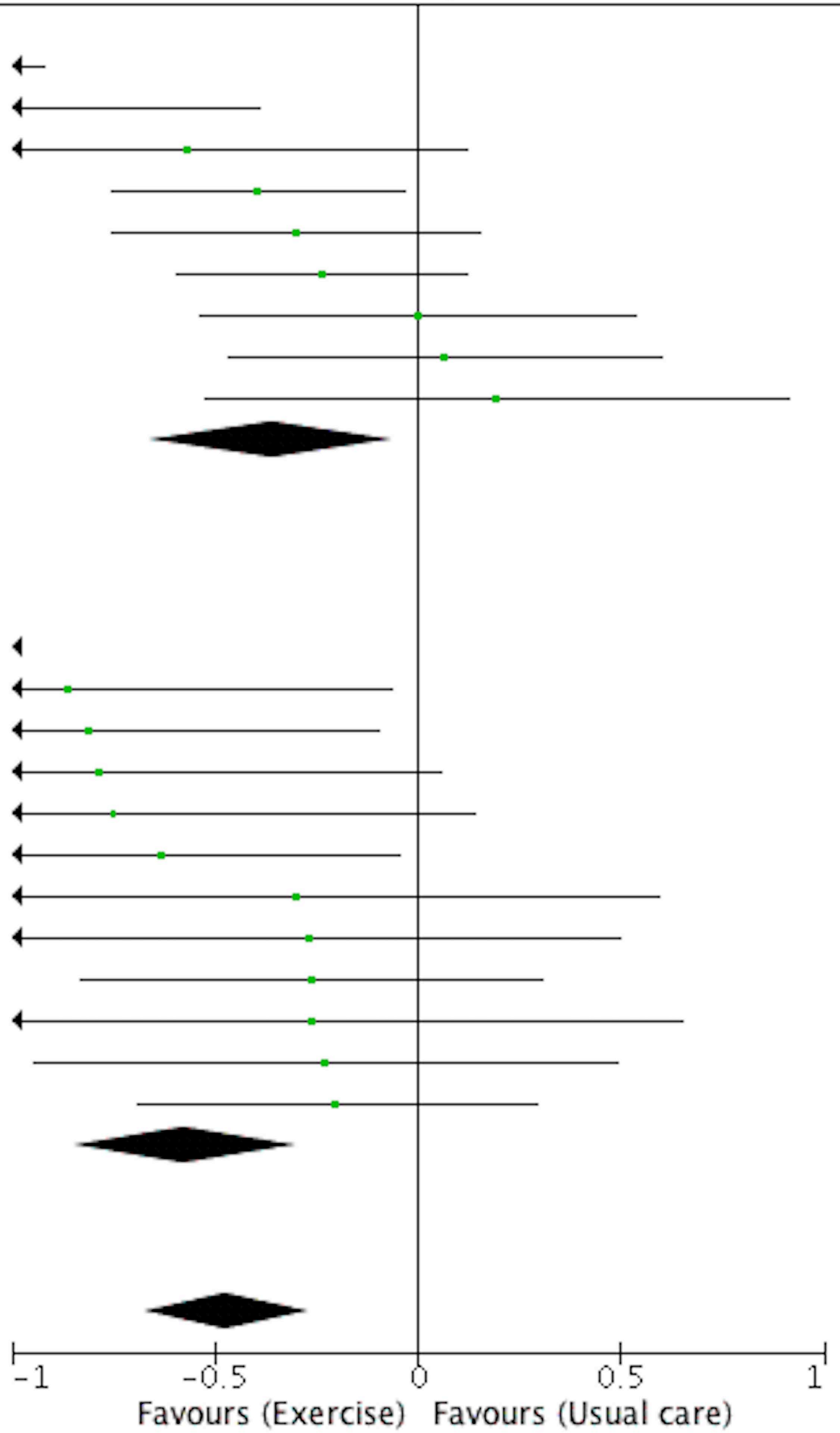
Heterogeneity: $Chi^2 = 41.16, df = 20 (P = 0.004); I^2 = 51\%$
Test for overall effect: $Z = 5.84 (P < 0.00001)$



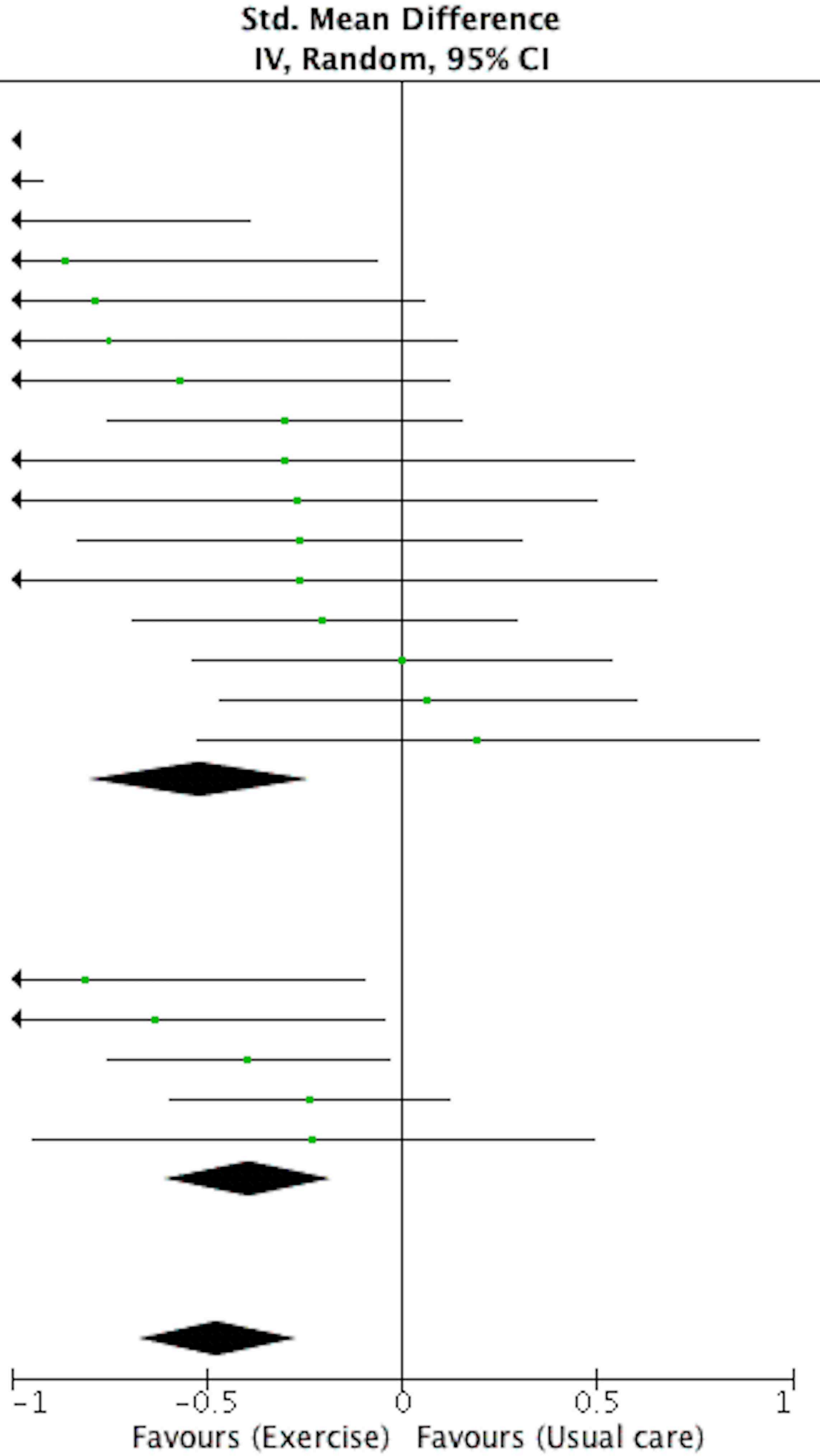

SE(SMD)
0
0.2
0.4
0.6
0.8
1
SMD
-2
-1
0
1
2

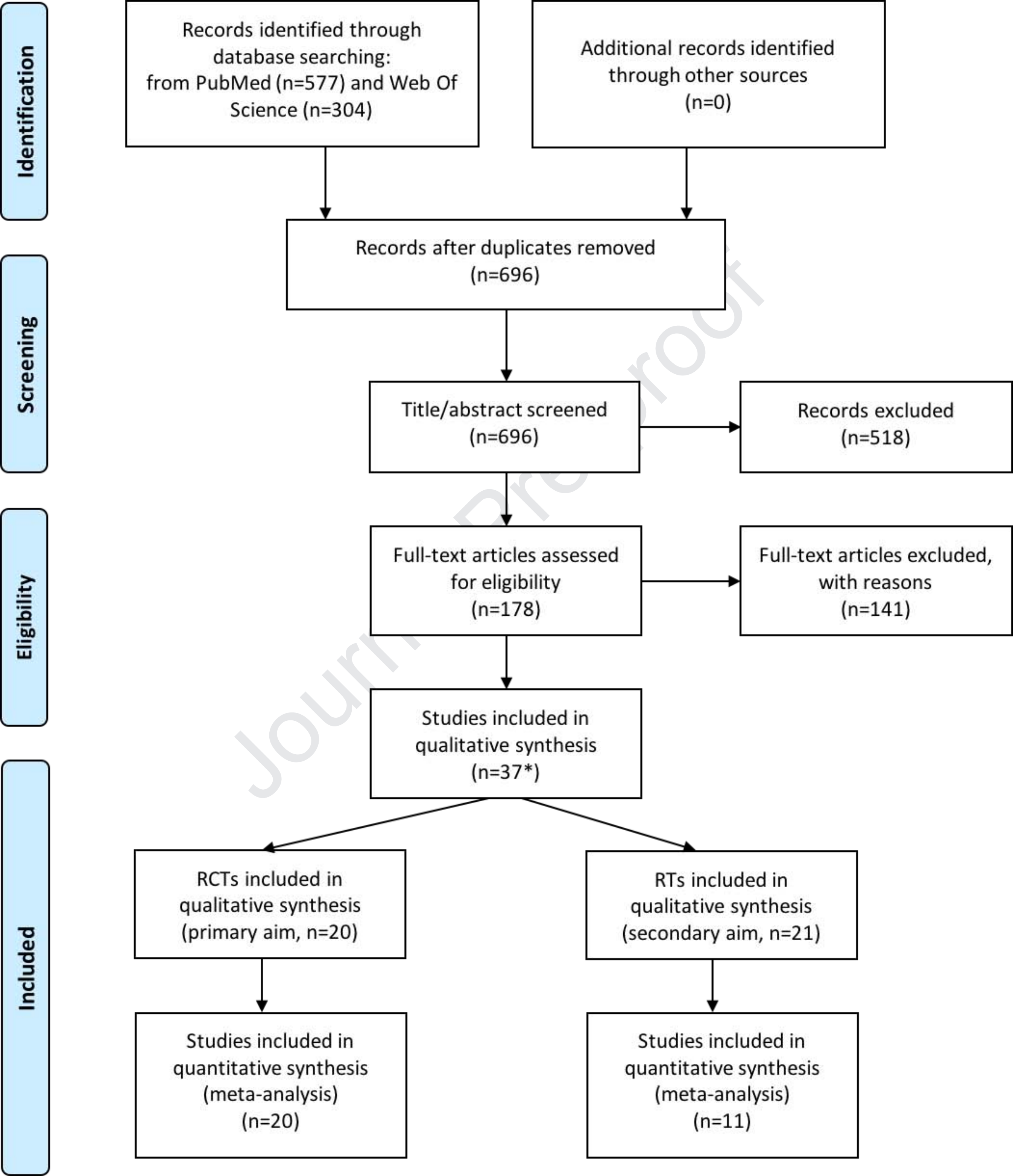
Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Adherence criteria (attendace to, at least, 80% of the training sessions)								
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**386**			**209**	**48.2%**	**-0.36 [-0.66, -0.06]**
Heterogeneity: Tau² = 0.11; Chi² = 19.90, df = 8 (P = 0.01); I² = 60%								
Test for overall effect: Z = 2.37 (P = 0.02)								
1.1.2 Remaining studies								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Subtotal (95% CI)			**226**			**182**	**51.8%**	**-0.57 [-0.84, -0.30]**
Heterogeneity: Tau² = 0.09; Chi² = 18.28, df = 11 (P = 0.08); I² = 40%								
Test for overall effect: Z = 4.13 (P < 0.0001)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: Tau² = 0.10; Chi² = 41.16, df = 20 (P = 0.004); I² = 51%								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: Chi² = 1.07, df = 1 (P = 0.30), I² = 6.6%								

Std. Mean Difference
IV, Random, 95% CI



Study or Subgroup	Exercise (Experimental) Mean	SD	Total	Usual care (Control) Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 All the participants were women								
Wong et al., 2018 (M-TC; L-B)	-2.4	0.74	17	0.02	1.43	14	3.3%	-2.13 [-3.04, -1.23]
Alentorn-Geli et al., 2008 (A & F; L-B)	43	13.3	12	76	18.9	10	2.7%	-1.98 [-3.03, -0.92]
Etnier et al., 2009 (F & R; L-B)	59.6	17.1	8	87.6	17.1	8	2.3%	-1.55 [-2.71, -0.39]
Da Silva et al., 2017 (A & F; L-B; Co-photo)	-1.2	1.13	20	-0.1	1.46	10	4.0%	-0.86 [-1.65, -0.07]
Valkeinen et al., 2008 (A & R; L-B)	-19.08	35.09	13	8.27	32.03	11	3.7%	-0.78 [-1.62, 0.06]
Hakkinen et al., 2001 (R; L-B)	-19	35.02	11	1	4.35	10	3.4%	-0.75 [-1.64, 0.14]
Tomas-Carus et al., 2007 (A; W-B)	-1.5	3.12	17	0.3	3.09	17	4.7%	-0.57 [-1.25, 0.12]
Collado-Mateo et al., 2016 (A & R; L-B)	-0.64	2.99	41	0.22	2.66	35	6.7%	-0.30 [-0.75, 0.15]
Assumpção et al., 2018 (R; L-B)	-1.93	4.2	16	-0.27	7.5	7	3.4%	-0.30 [-1.19, 0.59]
Da Silva et al., 2017 (A & F; L-B)	-0.5	1.46	20	-0.1	1.46	10	4.2%	-0.27 [-1.03, 0.50]
Mannerkorpi et al., 2000 (A & F; W-B; Co-edu)	-0.9	2.35	20	-0.1	3.41	30	5.6%	-0.26 [-0.83, 0.31]
Assumpção et al., 2018 (F; L-B)	-1.58	2.91	14	-0.27	7.5	7	3.3%	-0.26 [-1.17, 0.65]
Van Santen et al., 2002 (A, B, F & R; L-B)	-5.1	16.14	37	-1.9	15.39	28	6.4%	-0.20 [-0.69, 0.29]
Schachter et al., 2003 (A, longer bouts; L-B)	-0.5	2.41	51	-0.5	3.45	18	5.9%	0.00 [-0.54, 0.54]
Schachter et al., 2003 (A, shorter bouts; L-B)	-0.3	2.79	56	-0.5	3.45	18	6.0%	0.07 [-0.46, 0.60]
Tomas Carus et al., 2008 (A; W-B)	-0.6	3.07	15	-1.2	2.98	15	4.5%	0.19 [-0.52, 0.91]
Subtotal (95% CI)			**368**			**248**	**70.1%**	**-0.52 [-0.80, -0.23]**
Heterogeneity: $Tau^2 = 0.19$; $Chi^2 = 38.27$, df = 15 (P = 0.0008); $I^2 = 61\%$								
Test for overall effect: Z = 3.57 (P = 0.0004)								
1.1.2 Remaining studies (both women and men participated)								
Wigers et al., 1996 (A; L-B)	-27	32.48	16	3	39.35	17	4.5%	-0.81 [-1.52, -0.10]
Carson et al., 2010 (M-Y; L-B)	-1.6	3.3	22	0.32	2.73	26	5.5%	-0.63 [-1.21, -0.05]
McBeth et al., 2012 (A; L-B)	-3.6	6.24	92	-0.9	7.88	44	7.7%	-0.39 [-0.76, -0.03]
McBeth et al., 2012 (A; L-B; Co-CBT)	-2.7	7.5	94	-0.9	7.88	44	7.7%	-0.23 [-0.59, 0.12]
Gianotti et al., 2014 (A, F & R; L-B; Co-edu)	-4.8	16.39	20	-1.42	10.43	12	4.5%	-0.23 [-0.95, 0.49]
Subtotal (95% CI)			**244**			**143**	**29.9%**	**-0.39 [-0.60, -0.18]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 2.88$, df = 4 (P = 0.58); $I^2 = 0\%$								
Test for overall effect: Z = 3.62 (P = 0.0003)								
Total (95% CI)			**612**			**391**	**100.0%**	**-0.47 [-0.67, -0.27]**
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 41.16$, df = 20 (P = 0.004); $I^2 = 51\%$								
Test for overall effect: Z = 4.55 (P < 0.00001)								
Test for subgroup differences: $Chi^2 = 0.48$, df = 1 (P = 0.49), $I^2 = 0\%$								




Identification
Screening
Eligibility
Included
Records identified through database searching: from PubMed (n=577) and Web Of Science (n=304)
Additional records identified through other sources (n=0)
Records after duplicates removed (n=696)
Title/abstract screened (n=696)
Records excluded (n=518)
Full-text articles assessed for eligibility (n=178)
Full-text articles excluded, with reasons (n=141)
Studies included in qualitative synthesis (n=37*)
RCTs included in qualitative synthesis (primary aim, n=20)
RTs included in qualitative synthesis (secondary aim, n=21)
Studies included in quantitative synthesis (meta-analysis) (n=20)
Studies included in quantitative synthesis (meta-analysis) (n=11)