Scand J Med Sci Sports 2014: 24: 259–272
doi: 10.1111/sms.12050

SCANDINAVIAN JOURNAL OF
MEDICINE & SCIENCE
IN SPORTS

Review

Physical exercise intervention in depressive disorders: Meta-analysis and systematic review

T. Josefsson[1,2], M. Lindwall[2,3], T. Archer[2,4]

[1]School of Social and Health Sciences, Halmstad University, Halmstad, Sweden, [2]Department of Psychology, University of Gothenburg, Gothenburg, Sweden, [3]Department of Food and Nutrition, and Sport Science, University of Gothenburg, Gothenburg, Sweden, [4]School of Education, Psychology and Sport Science, Linnaeus University, Kalmar, Sweden
Corresponding author: Torbjörn Josefsson, School of Social and Health Sciences, Halmstad University, Box 823, S-301 18 Halmstad, Sweden. Tel: +46 35 16 71 46, Fax: +46 35 18 61 92, E-mail: Torbjorn.Josefsson@hh.se

Accepted for publication 23 November 2012

Previous meta-analyses investigating the effect of exercise on depression have included trials where the control condition has been categorized as placebo despite the fact that this particular placebo intervention (e.g., meditation, relaxation) has been recognized as having an antidepressant effect. Because meditation and mindfulness-based interventions are associated with depression reduction, it is impossible to separate the effect of the physical exercise from the meditation-related parts. The present study determined the efficacy of exercise in reducing symptoms of depression compared with no treatment, placebo conditions or usual care among clinically defined depressed adults. Of 89 retrieved studies, 15 passed the inclusion criteria of which 13 studies presented sufficient information for calculating effect sizes. The main result showed a significant large overall effect favoring exercise intervention. The effect size was even larger when only trials that had used no treatment or placebo conditions were analyzed. Nevertheless, effect size was reduced to a moderate level when only studies with high methodological quality were included in the analysis. Exercise may be recommended for people with mild and moderate depression who are willing, motivated, and physically healthy enough to engage in such a program.

Unipolar major depression, as measured by the integrated Disability-Adjusted-Life-Years (DALYs) instrument, defined as "the sum years lost due to premature mortality and years lived with disability adjusted for severity" (Murray & Lopez, 1997; p. 1436), has become a serious threat to public health worldwide, and it is rated as the third leading cause of burden of disease in high-income countries (Lopez et al., 2006). The disorder, a common condition compared with other medical diagnoses (Ebmeier et al., 2006), is estimated to increase, affecting nearly 340 million people worldwide, and 18 million people in the United States at any one time; it is expected to be the second highest cause of burden of disease (DALYs) by 2020 (Murray and Lopez, 1997) with chronic lifelong risk for recurrent relapse (Segal, Williams & Teasdale, 2002). It is associated also with high morbidity, co-morbidity and mortality (Cassano & Fava, 2002) and people with chronic diseases such as diabetes or arthritis show an increasing risk of developing major depression compared with the general population (Moussavi et al., 2007).

The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) categorizes major depressive disorder as a clinical syndrome characterized by depressive mood or loss of interest in activities for two weeks or more as key elements, accompanied by at least four additional features of emotional, physiological, or cognitive symptoms (e.g., sleeping disturbances, changes in eating habits, fatigue, suicidal thoughts, reduced ability to concentrate) (DSM-IV, American Psychiatric Association, 1994).

Depression is usually treated with medication or psychotherapy or a combination of both. However, many depressed people do not seek any help, and depression is therefore generally considered to be vastly undertreated (Cassano & Fava, 2002; Segal et al., 2002) with approximately only 12% seeking professional help, partly because of the stigma associated with depression (Segal et al., 2002). Of those who do seek help at mental health services, it is frequently reported that many do not get any treatment at all or that they receive inadequate treatment (Cassano & Fava, 2002). Only 18–25% of the treated depressed patients in the United States actually received adequate treatment (Ebmeier et al., 2006). Psychological treatments such as cognitive behavioral therapy (CBT) are superior to placebo and equivalent to medication. Other psychotherapies (e.g., interpersonal therapy) have shown similar results and no single therapy seems to be more efficacious than others in treating depression (see Ebmeier et al., 2006 for an overview). A recently conducted meta-analysis (Kirsch et al., 2008) on the effect of antidepressants found that it has no clinically significant effect on mild, moderate, or severe depression compared with placebo. A small significant effect was detected only for those patients who scored very high on the Hamilton Rating Scale of Depression, and the authors finally concluded that "there seems little evidence to support the prescription of antidepressant medication to any but the most severely depressed patients, unless alternative treatments have failed to provide benefit" (p. 266). In addition, antidepressants are associated with side effects such as withdrawal symptoms (Ebmeier et al., 2006), weight gain (Berken et al., 1984), dry mouth, nausea, insomnia, constipation, anxiety (Trindade et al., 1998), sexual dysfunction (Hsu & Shen, 1995), and not having a long-term effect (Segal et al., 2002). Another disadvantage with antidepressants is low adherence; 50% stop taking the prescribed antidepressant within the first month (Cassano & Fava, 2002).

Several studies indicate that many depressed people (as well as people with anxiety disorders) in Western societies tend to use complementary and alternative medicine (CAM) instead of traditional treatment (Unutzer et al., 2000; van der Watt et al., 2008). In fact, CAM is more frequently used than conventional therapies among severely depressed people in the United States, and 53.6% of the respondents in a depressed sample reported to have used CAM during the last 12 months (Kessler et al., 2001). There is a wide range of different complementary and alternative therapies/interventions available in the treatment of depression (e.g., herbal interventions, aromatherapy, acupuncture, light therapy; van der Watt et al., 2008). The specific types of therapies most commonly used by the respondents in the aforementioned study (Kessler et al., 2001) were relaxation training and spiritual healing, and to a lesser extent, self-help group and imagery. The efficacy of these various complementary and alternative interventions is not sufficiently investigated, and the evidence so far is – for the vast majority of these interventions – poor (van der Watt et al., 2008). Nevertheless, this frequent use of CAM does reflect the attraction these nonconventional interventions seem to have on depressed people. In contrast to what one may first easily assume, it does not seem to be dissatisfaction with conventional treatment primarily that generates this great interest in seeking alternative remedies for depressive symptoms. The main reason why people prefer alternative therapies is rather "because they find these health care alternatives to be more congruent with their own values, beliefs, and philosophical orientations toward health and life" (Astin, 1998; p. 1548).

The widespread aspect of depression demands efficacious treatment interventions as well as better preventive methods (O'Neal et al., 2000; Segal et al., 2002; Landers & Arent, 2007). The present account examines the influence of physical exercise as a suitable intervention for the disorder.

Exercise and depression

Physical activity is strongly related to several physical health benefits and regular exercise has been successfully included in primary prevention, treatment, and rehabilitation for many chronic diseases (e.g., cardiovascular disease, diabetes, cancer) as well as for premature mortality (Warburton et al., 2006; Haskell et al., 2007). Physical activity has also become increasingly and firmly associated with improvements in mental health and psychological well-being (Mutrie, 2000; Landers & Arent, 2007). In particular, exercise is believed to be effective in preventing depression and also to significantly reduce depressive symptoms in clinical as well as in nonclinical populations (O'Neal et al., 2000; Landers & Arent, 2007). Several correlational studies show that exercise is negatively related to depressive symptoms (e.g., Galper et al., 2006; Hassmén et al., 2000). Moreover, a considerably large number of intervention studies have by now investigated the effect of various exercise programs on depression and the vast majority of them indicate that exercise significantly reduces depression (e.g., Blumenthal et al., 2007; Martinsen et al., 1985; Singh et al., 1997).

An early attempt to systematically evaluate the effect of exercise on depression was done by North et al. (1990) who analyzed a collection of 80 studies and found a moderate overall effect size ($d = -0.53$), indicating that exercise reduces depression scores by one-half a standard deviation compared with the control groups. However, this meta-analysis included also studies with non-depressed populations, which limit the conclusions that may be drawn from these results. Craft and Landers (1998) therefore decided to include trials only where the population was defined as depressed in their meta-analysis. Thirty studies were analyzed, resulting in an overall effect size of $d = -0.72$, again indicating a larger reduction in depression scores for the exercise group in comparison with the control group. Further, Craft and Landers (1998) found no significant differences when exercise was compared with other treatments such as individual psychotherapy. However, these two meta-analyses (North et al., 1990; Craft & Landers, 1998) also included studies with methodological weaknesses such as uncontrolled nonrandomized trials with small samples. In an attempt to produce an analysis with better quality, Lawlor and Hopker (2001) further narrowed the inclusion criteria in their meta-analysis to only contain randomized controlled trials, resulting in a collection of 14 studies. Compared with no treatment, an overall effect size, $d = -1.1$ [95% confidence interval (CI), –1.5, –0.6], in favor of exercise, was reported. The standardized mean difference in depression scores between the exercise group and the control group was –7.3 (95% CI, –10.0, –4.6). Similar to Craft and Landers's (1998) results, Lawlor and Hopker (2001) found no differences between exercise and cognitive therapy in the reduction of depressive symptoms. Despite this rather large effect size, the authors concluded that "the effectiveness of exercise in reducing symptoms of depression cannot be determined because of a lack of good quality research on clinical populations with adequate follow up" (Lawlor & Hopker, 2001; p. 1). The methodological weaknesses highlighted by Lawlor and Hopker (2001) were mainly of experimental nature, for instance inadequate allocation of concealment, the use of self-report scales to diagnose depression instead of clinical interviews, no follow-ups, lack of intention-to-treat analyses, and blinded outcome assessments.

Lawlor and Hopker (2001) were taken to task for interpreting their results far too negatively implying that such a large effect size could rather be regarded as "extremely compelling evidence" (Mutrie, 2002; p. 412) for the antidepressive effect of exercise. Biddle and Mutrie (2008) argued that the effect size found in Lawlor and Hopker (2001) is similar to the effect of CBT on depression (see Butler et al., 2006 for an overview) and the mean difference score (–7.3) may well be clinically significant.

Lawlor (2001) replied to some of this criticism by stating that poor methodological quality (inadequate allocation concealment, no intention-to-treat analyses, and lack of blinding) is estimated to exaggerate the treatment effect by 20–40%. Methodological problems such as inadequately concealed treatment allocation have been estimated to result in an exaggeration of treatment effects by 41%, and unclear concealed trials by 30%. Studies with a non double-blind design have been associated with 17% larger effect sizes (Schultz et al., 1995). Intervention studies with inadequate allocation concealment and lack of blinding were especially associated with exaggerated treatment effects when the outcome was subjectively assessed (Wood et al., 2008). Callaghan (2004) as well as Landers and Arent (2007) acknowledge the possibility that the large effect size found in Lawlor and Hopker (2001) could be overestimated, but also argue that even with a reduction by 41% of the effect size, it would still be a moderate to large effect size (–0.69). Thus, the evidence for the antidepressant effect on exercise may be weakened because of methodological problems, but should definitely not be dismissed nor invalidated (Callaghan, 2004; Landers & Arent, 2007).

Since the highly influential article by Lawlor and Hopker (2001), four meta-analyses have been published investigating the effect of exercise on depression. Stathopoulou et al. (2006) analyzed 11 randomized controlled trials, including four studies that were published after the Lawlor and Hopker (2001) review. A large effect size of $g = -1.39$ for the advantage of exercise compared with control condition was found (Stathopoulou et al., 2006). However, methodological problems such as the ones previously described were present in this analysis as well, suggesting that this effect size may also be exaggerated to a certain extent.

In an ambitious effort to address the aforementioned methodological problems highlighted by Lawlor and Hopker (2001), Rethorst et al. (2009) performed a rigorous search strategy resulting in an inclusion of 58 randomized trials. The main purpose was to provide Level 1, Grade A evidence for the effect of exercise on depression, something that previously had not been done before. Based on the guidelines developed by Guyatt et al. (2006) for grading the strength of recommendations and quality of evidence in clinical research, Level 1, Grade A reflects a strong recommendation and the highest quality of evidence. This can be provided by randomized controlled trials with large sample sizes (which is possible to obtain by a meta-analysis). Rethorst et al. (2009) found an overall large effect size ($d = -0.80$, 95% CI, –0.92, 0.67), indicating a significant reduction in depression for exercise treatment compared with control condition. The results further showed that the reduction in depression scores was significantly larger for the clinical population ($d = -1.03$) than for the nonclinical population ($d = -0.59$), thus suggesting that clinically depressed patients benefit more from exercise than nonclinical persons. Furthermore, Rethorst et al. (2009) surprisingly found that adequate allocation concealment and intention-to-treat analysis were associated with larger effect sizes, and that clinical interviews did not reveal any differences in effects sizes compared with self-report assessments of depression. This may suggest that the methodological problems that have been widely debated over may have had less impact on the results than expected. The authors conclude that the inclusion of only randomized trials with a cumulative large sample size (almost 3000 participants) could be classified as Level 1, Grade A evidence. The large effect sizes found – especially for the clinically depressed sample – indicate that exercise is an effective treatment for major depression (Rethorst et al., 2009).

Mead et al. (2009), in a study of 28 trials fulfilling the inclusion criteria, obtained a meta-analytic result with a large overall effect size ($d = -0.82$, 95% CI –1.12, –0.51) that indicated a significant reduction of depression scores for the exercise condition compared with the control condition. When analyzing only those trials (Mather et al., 2002; Dunn et al., 2005; Blumenthal et al., 2007), fulfilling all three methodological quality criteria – adequate allocation concealment, intention-to-treat analysis, and blinded outcome assessment – the effect size was substantially reduced from large to moderate ($d = -0.42$), and not significant. Mead et al. (2009) concluded that exercise could be recommended to depressed people, but in line with Lawlor and Hopker's (2001) conclusions, they also stated, somewhat contradictorily, that "outstanding uncertainties remain about how effective exercise is for depression, mainly because of methodological considerations" (p. 13).

The most recently published meta-analysis was conducted by Krogh et al. (2011) who aimed at determining if exercise should be provided by healthcare services for clinically depressed adults. Thirteen trials were finally included in the main analysis resulting in an overall effect size of $g = –0.40$ (95% CI –0.66, –0.14). An analysis of those trials (Dunn et al., 2005; Blumenthal et al., 2007; Krogh et al., 2009) that were considered to be of methodologically robust quality (adequate allocation concealment, intention-to-treat analysis and blinded outcome assessment) showed a considerable reduction in the effect of exercise ($g = –0.19$, 95% CI –0.70, 0.31). The conclusions drawn by Krogh et al. (2011) were that exercise may have a small short-term effect on depression, but it cannot yet be recommended as a treatment for clinical depression.

In sum, the main results from seven meta-analyses so far show that exercise has an antidepressant effect compared with control conditions that ranges from slightly moderate ($g = –0.40$; Krogh et al., 2011) to very large ($g = –1.39$; Stathopoulou et al., 2006). However, the majority of the included studies in all these meta-analyses suffer more or less from serious methodological problems (e.g., small samples, inadequate allocation concealment, lack of intention-to-treat analysis and blinding, and lack of clinical interviews to diagnose depression) that may have biased the results in favor of exercise. Krogh et al. (2011) as well as Mead et al. (2009) showed that the effect size was substantially reduced when only robust trials were analyzed ($g = –0.19$ and $d = –0.42$, respectively). In addition, it should be noted that the subanalyses of robust trials in Krogh et al. (2011) and Mead et al. (2009) only contained three studies each of which limits the interpretations of these results. The methodological concern has lead some researchers to be cautious and call for further research before any prescriptions of exercise for depression can be made (e.g., Krogh et al., 2011) while others argue that there are enough evidence to recommend exercise as a treatment for depression even if they also acknowledge that additional research in some areas are necessary (Biddle & Mutrie, 2008; Rethorst et al., 2009). Mead et al. (2009) believe that there is enough evidence of the antidepressant effect of exercise to recommend it to a depressed population, but they also emphasize that it is not yet possible to determine exactly how effective exercise interventions are.

Hence, despite numerous studies as well as several meta-analyses and reviews, there still remain uncertainties concerning the effectiveness of exercise as a treatment for depression. Although all four meta-analyses that have been published since Lawlor and Hopker (2001) are impressively ambitious and carefully conducted, there are a few minor, but nevertheless essential, improvements that need to be made in future meta-analyses in order to establish the effect of exercise on depression. First, it is of major importance that the exercise condition is compared with either control condition; (a) no treatment; (b) a placebo treatment; or (c) the usual care in- or outpatients receive, to establish if there really is an antidepressant effect of exercise to begin with.

The four latest published meta-analyses (Stathopoulou et al., 2006; Mead et al., 2009; Rethorst et al., 2009; Krogh et al., 2011) have all included studies where the control group has received some type of meditation practice and/or relaxation training (e.g., Klein et al., 1985; Krogh et al., 2009). Although these practices have been categorized as placebo treatments, both have, in fact, been widely recognized for their ability to reduce anxiety and depression symptoms (Reynolds & Coats, 1986; Murphy et al., 1995; Tloczynski & Tantriella, 1998; Fortney & Taylor, 2010). Similarly, Rethorst et al. (2009) included a study where exercise was compared with another alternative treatment; bright light therapy (Pinchasov et al., 2000). Second, Mead et al. (2009) have included studies where the exercise group received eastern meditative practices such as Tai Chi or Qigong (Chou et al., 2004; Tsang et al., 2006). These practices involve several different components: meditation, mind tranquility, balance, deep breathing, relaxation, and bodily movements (Lan et al., 2002; Tsang et al., 2006, 2008; Wang et al., 2004). The physical exercise intensity for Tai Chi has been classified as "low-speed and low-impact exercise" (Wang et al., 2004; p. 217) whereas Qigong involves physical movements that are similarly "executed at very low energy expenditure levels" (Tsang et al., 2008; p. 305). Both Tai Chi and Qigong are associated with reductions in depression symptoms (Tsang et al., 2008). However, it is impossible to separate the effect of the physical exercise from the effect of the meditation-related components. Thus, trials where the exercise group practice Tai Chi or Qigong should not be included in a meta-analysis examining the effect of exercise on depression. Third, Rethorst et al. (2009) included several studies where the participants had a main diagnosis other than depression (e.g., schizophrenia in Jorgensen, 1986; multiple sclerosis in Petajan et al., 1996). In order to establish the effect of exercise on depression it would be preferable to use a fairly homogenous population of clinically or nonclinically depressed people without multiple disorders or other main medical diagnoses.

Since the meta-analysis by Krogh et al. (2011), several trials investigating the effect of exercise on depression have been published, of which the ambitious TREAD study by Chalder et al. (2012) is the absolutely largest ($n = 288$) published trial on the subject thus far. Such a large study may indeed have a substantial impact on the overall effect size in a meta-analysis. Thus, there is already a need to update previous meta-analytic results.

The objective of the current study was to present the results of a meta-analysis of exercise intervention in clinical depression (using only pure control groups).

Method

Data collection

Several search strategies were conducted based on title and abstract: (a) The following databases and websites were searched: Medline, PsycINFO, Cochrane Controlled Trials Register, the Cochrane Database of Systematic Reviews, the Social Science Citation Index, Google Scholar, ERIC, BioMed Central, TRoPHI, DoPHER, Centre for Reviews and Dissemination, Thomson Reuters, SIGLE, Ongoing Reviews Database, DARE, the Campbell Collaboration (C 2 Social, Psychological, Education), http://www.controlled-trials.com, and http://www.clinicaltrials.com. The word terms used in the searches were *exercise*, *physical activity*, *depression*, *depressive disorders*, *aerobic*, *physical fitness*. (b) The content of lists in the following journals (2008–2010) were hand-searched: Ment Health Phys Act, British Journal of Sports Medicine, the American Journal of Sports Medicine, Journal of Consulting and Clinical Psychology, and Journal of Exercise and Sport Psychology. (c) The reference lists of previous meta-analyses, systematic, and nonsystematic reviews were also hand-searched. (d) An expert in the research area was contacted for information of ongoing studies and unpublished studies. The majority of searches were performed during November and December 2010. Additional searches were done in March and April 2011, and April 2012. Only randomized controlled studies were included. A trial was defined as being a randomized controlled trial if the allocation of participants to experiment group and control group was described by the authors as randomized (*randomly*, *random*, and *randomization*).

Studies that compared aerobic exercise or nonaerobic exercise with no treatment (e.g., waitlist) or with a placebo group were included. Studies that had included placebo treatments, which themselves have been associated with reductions in depressive symptoms, were excluded (e.g., meditation, relaxation training). For the same reason, studies where the exercise intervention consisted of eastern meditative practices such as yoga, Tai Chi, or Qigong were also excluded. Studies comparing exercise with another type of intervention (e.g., medication, psychotherapy) were also excluded. Studies where exercise was combined with another intervention were also excluded (e.g., exercise and psychotherapy vs psychotherapy, or exercise and medication vs medication). However, studies where the control group received what was categorized as their standard care (e.g., in- or outpatients receiving their regular care) were included. Likewise, studies in which both the exercise group and the control group received their usual care were included. Studies where the control group received a placebo intervention consisting of a low dose of exercise (e.g., stretching) were also included.

Only studies that included adult participants over 18 years old and above, defined by the authors as having depression or depressive symptoms (by any severity; mild, moderate or severe depression), were included. Thus, studies where the participants were children or adolescents were excluded. Studies including participants with multiple diagnoses were excluded (e.g., depressive symptoms and neurotic disorders). Similarly, studies with a heterogeneous sample consisting of depressed participants as well as participants with other diagnoses were excluded (e.g., anxiety disorders, neurotic disorders). Studies where the participants' main diagnosis was physical (e.g., fibromyalgia, cancer, multiple sclerosis) and not depression were also excluded.

Coding system

A coding manual was used to rate participant details, intervention characteristics, and methodological quality. Type of population was coded as clinical if the participants were either in- or outpatients recruited from psychiatric health service or general practice. If the participants were volunteers, for instance recruited by media, they were classified as nonclinical. Any means of assessing depression were accepted; clinical interviews, diagnoses according to a diagnostic system (e.g., DSM-IV or International Classification of Diseases, Tenth Revision) or by a cut-off point on a self-report depression scale (e.g., Beck Depression Inventory, HRDS). Assessment of depression outcome was coded. Because several studies used more than one measure of depression, the measure defined by the authors as their main outcome measure was used in this meta-analysis. If a study did not specify which the main outcome measure was, the measure first reported in the abstract was used.

Several intervention characteristics were coded: type of exercise (aerobic or nonaerobic), exercise frequency (training sessions/week), length of exercise session (minutes/session), exercise dose/intensity, if the exercise intervention was supervised or not, control group characteristics (e.g., intervention), duration of exercise intervention (weeks), and adherence (%).

Methodological quality was coded by three criteria: allocation concealment, blinded outcome assessment, and intention-to-treat analysis. In accordance with recommendations in Pildal et al. (2005), allocation concealment was considered to be adequate if any of these methods were used: "central randomisation; numbered coded vehicles; opaque, sealed, and sequentially numbered envelopes; and other methods containing convincing means of concealment" (p. 2). Other methods or unclear methods were defined as inadequate. A trial was coded as having used intention-to-treat analysis if all participants that were randomly assigned to the exercise group and the control group were included in the analyses (Hollis & Campbell, 1999). Thus, studies that analyzed only those participants that completed the intervention were not defined as using intention-to-treat analysis. A study was coded as blinded outcome assessment if the assessor was unaware of the treatment allocation (Forder et al., 2005).

Data analysis

The effect sizes for each trial were calculated using the overall pooled standardized mean difference between the exercise group and the control group in depression scores. Because several of the included studies had small sample sizes, Hedges's g was used to calculate the standardized mean differences (Borenstein, 2009). The obtained overall effect sizes were interpreted according to the guiding principles in Cohen (1988); 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect. Because the included studies differ in several ways (e.g., type of exercise, intervention duration, intensity/dose, and population) a random effect model was used to calculate the pooled effect size because this is generally recommended when heterogeneity between the studies is assumed (Borenstein et al., 2009). Homogeneity of variance was tested by the Q-value (a significant P-value indicates heterogeneity among studies; Shadish & Haddock, 2009), and the I^2 index, which "describes the percentage of total variation across studies that is due to heterogeneity rather than chance" (Higgins et al., 2003; p. 558). The I^2 ranges from 0 to 100% where 0 represents no heterogeneity and 100% maximum heterogeneity (Higgins et al., 2003).

A sensitivity analysis was conducted investigating the impact of methodological quality on effect size. Thus, only those studies that were considered to be of high methodological quality (adequate allocation concealment, intention-to-treat analysis and blinded outcome assessment blinding) were analyzed. However, only two studies (Dunn et al., 2005; Blumenthal et al., 2007) fulfilled these criteria. In addition, we also made a sensitivity analysis including only trials that compared exercise with no treatment or placebo treatment.

For studies that included multiple intervention groups (e.g., exercise, psychotherapy, and control), data were extracted from the exercise group and the control group only. For trials using several exercise groups, the exercise intervention with the highest effect was used in the present meta-analysis.

The Comprehensive Meta Analysis program version 2.0 was used for all analyses (Borenstein et al., 2005).

Results

A total of 89 potential studies were retrieved. Fifteen studies passed the inclusion criteria of which 10 reported sufficient information for calculation of effect sizes (see Tables 2, 3). The authors of the five studies with missing information were contacted by e-mail. All the five authors replied, but only three of them were able to provide the necessary additional information. Hence, 74 studies were excluded (see Table 1), 15 trials were included of which 13 were finally used for statistical analyses (exercise group: $n = 366$, control group: $n = 354$) (see Table 2).

Four of the included trials (Chalder et al., 2012, de Zeeuw et al., 2010; Foley et al., 2008; Legrand & Heuze, 2007) have not been included in any of the previously published meta-analyses.

Population characteristics

Ten studies recruited nonclinical participants (Hess-Homeier, 1981; Setaro, 1985; Epstein, 1986; Doyne et al., 1987; McNeil et al., 1991; Dunn et al., 2005; Singh et al., 2005; Blumenthal et al., 2007; Legrand & Heuze, 2007; de Zeeuw et al., 2010), two studies (Mather et al., 2002; Foley et al., 2008) included both clinical and nonclinical participants, while the remaining three studies (Mutrie, 1988; Veale et al., 1992; Chalder et al., 2012) recruited participants solely from clinical populations (see Table 3).

Table 1. Characteristics of excluded studies

Total number of excluded studies	74
Control condition that was not placebo or “no treatment”	19
Population of multiple or mixed diagnoses	9
Reviews, meta-analyses or commentaries	31
Nonrandomized studies	5
Non depressed population	2
Children or adolescent population	5
Other	3

Assessment of depression diagnosis

A clinical interview for diagnosing depression was used in seven studies (Hess-Homeier, 1981; Doyne et al., 1987; Veale et al., 1992; Mather et al., 2002; Singh et al., 2005; Blumenthal et al., 2007; Foley et al., 2008). In one study (Mutrie, 1988), depression was diagnosed by a general practitioner. The remaining five studies used a diagnostic manual system or a cut-off point on a depression scale to determine depression (Setaro, 1985; Epstein, 1986; McNeil et al., 1991; Dunn et al., 2005; Legrand & Heuze, 2007; de Zeeuw et al., 2010; Chalder et al., 2012). See Table 3.

Exercise intervention characteristics

All studies except Singh et al. (2005) and Mather et al. (2002) provided aerobic exercise. The median of exercise frequency was three sessions per week. The mean length of exercise session was 36.4 min/session. The mean of exercise duration was 9.4 weeks. All studies used supervised exercise interventions. The mean of adherence was 81.3%. The Chalder et al. (2012) study used an individually tailored Physical Activity Facilitator (PAF) program. Thus, type of exercise, frequency of exercise sessions and dose/intensity were not reported (see Table 3).

Table 2. Results on the unstandardized pre mean scores and post mean scores on depression

Study	Publication type	Exercise		Control		n
		Pre mean (SD)	Post mean (SD)	Pre mean (SD)	Post mean (SD)	
Hess-Homeier (1981)	D	24.4 (DNA)	9.8 (6.93)	23.5 (DNA)	16.2 (7.4)	11
Setaro (1985)	D	68.92 (5.27)	62.00 (6.51)	68.56 (3.87)	69.88 (3.46)	50
Epstein (1986)	J	25.29 (6.52)	9.00 (10.94)	22.00 (7.16)	16.30 (7.44)	17
Mutrie (1988)	D	22.4 (6.82)	9.46 (4.28)	23.0 (5.80)	21.4 (5.26)	16
Doyne et al. (1987)	J	19.27 (5.61)	8.18 (5.27)	16.06 (5.12)	15.25 (6.30)	19
McNeil et al. (1991)	J	16.6 (3.1)	11.1 (3.0)	15.2 (2.4)	14.7 (3.7)	20
Veale et al. (1992)	J	22.91 (1.1)	13.94 (2.13	26.66 (1.52)	17.79 (1.89)	65
Mather et al. (2002)	J	16.7 (7.77)	12.6 (6.94)	17.4 (6.86)	13.7 (6.19)	85
Dunn et al. (2005)	J	19.1 (1.8)	9.0 (3.6)	20.5 (2.4)	14.0 (4.9)	30
Singh et al. (2005)	J	DNA	8.5 (5.5)	DNA	14.4 (6.0)	37
Blumenthal et al. (2007)	J	16.4 (3.7)	9.2 (6.1)	17.2 (4.4)	11.1 (7.0)	76
Legrand and Heuze (2007)	J	20.63 (6.50)	10.88 (8.08)	22.29 (4.31)	20.71 (6.26)	15
Foley et al. (2008)	J	28.80 (9.86)	17.10 (11.09)	28.69 (8.59)	21.38 (7.74)	13
de Zeeuw et al. (2010)	J	6.2 (1.5)	3.1 (1.9)	6.8 (1.5)	5.8 (2.2)	27
Chalder et al. (2012)	J	32.1 (9.0)	16.12 (11.3)	32.1 (9.5)	16.87 (12.6)	288

The studies used in the present meta-analysis are in bold printing.
A, abstract; D, doctoral dissertation; DNA, data not available; J, peer-reviewed journal; n, total number of participants (exercise group and control group) used in the meta-analyses.

Table 3. Population characteristics, intervention characteristics, and methodological quality of the included studies

Study	n*	Age	Depression-diagnosis	Depression-outcome	Experiment group type of exercise	Experiment group frequency/dose supervision	Contol group	Duration weeks	Adherence (%)	Al. Conc.	Blind	Intention-to-treat
Hess-Homeier (1981)	13	21–41	Interview/BDI	BDI	Aerobic (running/walking) $n = 5$	4 times/week 30 min/session Dose not provided Supervised	No treatment $n = 6$	8	85	No	No	No
Setaro (1985)	60	18–35	MMPI	MMPI	Aerobic (dance) $n = 25$	2 times/week Length/session & dose not provided Supervised	No treatment $n = 25$	10	83	No	No	No
Epstein (1986)	17	39.4	DSM-III	BDI	Aerobic (walking/running) $n = 7$	3–5 sessions/week 30 min/session Dose: individual pace Supervised	No treatment $n = 10$	8	Not provided	No	No	No
Mutrie (1988)	UC	43.4	Diagnosed by general practitioner	BDI	Aerobic $n = 9$	3 sessions/week 29 min/session Dose not provided Supervised	No treatment $n = 7$	4	UC	No	No	No
Doyne et al. (1987)	UC	28.5	Interview/RDC	BDI	Aerobic (running/walking) $n = 8$	4 sessions/week Length/session not provided 80% max heart rate Supervised	No treatment $n = 11$	8	UC	No	No	No
McNeil et al. (1991)	20	Not provided	BDI	BDI	Aerobic (walking) $n = 10$	3 times/week 20–40 min/session Dose: individual pace Supervised	No treatment $n = 10$	6	100	No	No	Yes
Veale et al. (1992)	83	35.5	Interview/ CIS	BDI	Aerobic (running)/routine care $n = 36$	3 sessions/week Length/session & dose not provided Supervised	Routine care $n = 29$	12	78	Yes	No	No
Mather et al. (2002)	86	64.9	Interview/ ICD-10	HRSD	Nonaerobic (muscle strengthening, endurance & stretching) $n = 42$	2 sessions/week 45 min/session Dose not provided Supervised	Health education classes $n = 43$	10	99	Yes	Yes	No
Singh et al. (2005)	40	69	Interview/ DSM-IV	HRSD	Nonaerobic (progressive resistance training) $n = 18$	3 sessions/week 60 min/session 80% of one repetition max Supervised	Usual care $n = 19$	8	92	Yes	Yes	No

Table 3. (continued)

Study	n^*	Age	Depression-diagnosis	Depression-outcome	Experiment group type of exercise	Experiment group frequency/dose supervision	Contol group	Duration weeks	Adherence (%)	Al. Conc.	Blind	Intention-to-treat
Dunn et al. (2005)	30	33.8	DSM-IV	HRSD	Aerobic (treadmill walking/running, stationary cycling) $n = 17$	3 sessions/week Length/session not provided 17.5 kcal/kg/week Supervised	Placebo (stretching flexibility exercise) $n = 13$	12	47	Yes	Yes	Yes
Blumenthal et al. (2007)	100	52	Interview/ DSM-IV	HDRS	Aerobic (treadmill walking/jogging) $n = 41$	3 sessions/week 30 min/session 70–80% max heart rate Supervised	Placebopills $n = 35$	16	76	Yes	Yes	Yes
Legrand and Heuze (2007)	17	33	BDI	BDI	Aerobic (treadmill, stationary cycling, rowing) $n = 8$	3–5 sessions/week 30 min/session 60–80% max heart rate Supervised	Placebo (low exercise, 1 session/week, 30 min/session) $n = 7$	8	88	No	No	No
Foley et al. (2008)	23	18–55	Interview/ SCI/DSM-IV	MADRS	Aerobic (not specified) $n = 8$	3 sessions/week 30–40 min/session Moderate intensity Supervised	Mild-intensity stretching $n = 5$	12	56	No	Yes	Yes
de Zeeuw et al. (2010)	30	41.1	PHQ-9	PHQ-9	Aerobic (cycling, jogging, walking, climbing on stationary gym equipment) $n = 14$	2 sessions/week 45 min/session 60–80% max heart rate Supervised	No treatment $n = 13$	10	90	No	No	Yes
Chalder et al. (2012)	361	39.8	ICD-10	BDI	Individually chosen exercise $n = 142$	Frequency and dose not provided Supervision UC	Usual care $n = 146$	16	UC	Yes	No	Yes

n^* = the total number of participants randomized in the exercise group and control group. The numbers of participants that completed the treatment/intervention are shown on experiment group and control group respectively. Age refers to mean age or range of age when mean age is not provided.

Al. conc., allocation concealment. BDI, Beck Depression Inventory; CIS, Clinical Interview Scale; DSM-III, Diagnostic and Statistical Manual of Mental Disorders (third edition); DSM-IV, Diagnostic and Statistical Manual of Mental Disorders (fourth edition); HRDS, Hamilton Rating Scale of Depression; ICD-10, International Classification of Diseases (tenth revision); MADRAS, the Montgomery-Åsberg Depression Rating Scale; MMPI, Minnesota Multiphasic Personality Inventory; PHQ-9, Patient Health Questionnaire; RDC, Research Diagnostic Criteria; SCI, Structured Clinical Interview; UC, unclear.

Study name	Hedges's *g*	Standard error	Variance	Lower limit	Upper limit	*Z*-value	*P*-value	Relative weight
Setaro, 1985	–1.556	0.319	0.102	–2.181	–0.931	–4.878	0.000	8.59
Epstein, 1986	–1.116	0.505	0.255	–2.107	–0.126	–2.209	0.027	6.23
Mutrie, 1987	–2.269	0.623	0.388	–3.490	–1.048	–3.643	0.000	5.03
McNeil, 1991	–1.422	0.484	0.234	–2.370	–0.474	–2.939	0.003	6.48
Veale, 1992	–0.049	0.247	0.061	–0.532	0.435	–0.198	0.843	9.56
Mather, 2002	–0.060	0.215	0.046	–0.482	0.361	–0.281	0.779	9.96
Dunn, 2005	–0.833	0.374	0.140	–1.566	–0.099	–2.225	0.026	7.85
Blumenthal, 2007	–0.167	0.228	0.052	–0.614	0.280	–0.731	0.465	9.80
Legrand, 2007	–1.054	0.524	0.274	–2.080	–0.027	–2.012	0.044	6.02
Foley, 2008	–0.408	0.536	0.288	–1.459	0.643	–0.761	0.447	5.89
de Zeuw, 2010	–0.994	0.397	0.158	–1.772	–0.215	–2.501	0.012	7.55
Doyne, 1987	–1.665	0.520	0.270	–2.683	–0.647	–3.205	0.001	6.07
Calder, 2012	–0.062	0.118	0.014	–0.293	0.168	–0.531	0.595	10.98
	–0.772	0.186	0.035	–1.138	–0.407	–4.143	0.000	

Statistics for each study

Hedges's *g* and 95% CI

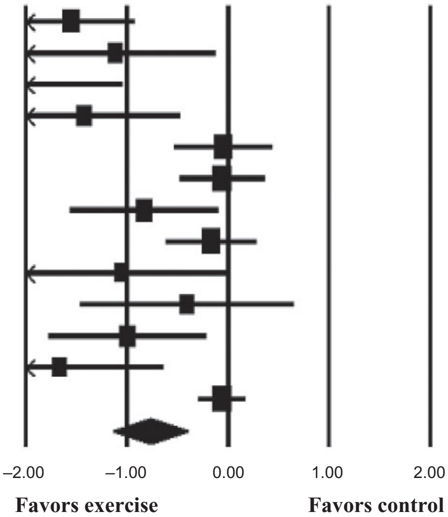

Fig. 1. Meta-analysis of all 13 trials investigating the effect of exercise vs control on depression.

Study name	Hedges's *g*	Standard error	Variance	Lower limit	Upper limit	*Z*-value	*P*-value	Relative weight
Dunn, 2005	–0.833	0.386	0.149	–1.589	–0.076	–2.157	0.031	39.19
Blumenthal, 2007	–0.167	0.231	0.053	–0.619	0.285	–0.723	0.469	60.81
	–0.428	0.325	0.106	–1.065	0.209	–1.316	0.188	

Statistics for each study

Hedges's *g* and 95% CI

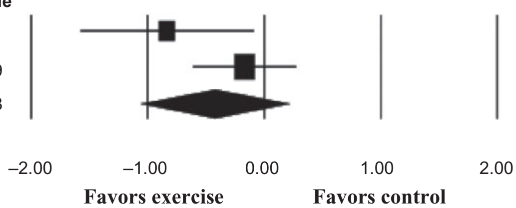

Fig. 2. Meta-analysis on the effect of exercise on depression, including only trials that met all three methodological quality criteria (allocation concealment, blinded outcome assessment, intention-to-treat analysis).

Methodological quality

Methodological quality was assessed using three criteria: adequate allocation concealment, the use of intention-to-treat analysis and blinded outcome assessment. As can be seen in Table 3, six studies were coded as having adequate allocation concealment (Veale et al., 1992; Mather et al., 2002; Dunn et al., 2005; Singh et al., 2005; Blumenthal et al., 2007; Chalder et al., 2012), five studies used blinded outcome assessment (Mather et al., 2002; Dunn et al., 2005; Singh et al., 2005; Blumenthal et al., 2007; Foley et al., 2008), and six studies used intention-to-treat analysis (McNeil et al., 1991; Dunn et al., 2005; Blumenthal et al., 2007; Foley et al., 2008; de Zeeuw et al., 2010; Chalder et al., 2012). Only two studies passed all three criteria (Dunn et al., 2005; Blumenthal et al., 2007). See Table 3.

Comparison between exercise and control condition on the effect of reduction in depression

The result presents the pooled standardized mean difference (calculated by Hedges's *g*, using a random effect model) of all 12 included studies. The overall effect size was –0.77 (95% CI, –1.14, –0.41) indicating a large effect size in favor of exercise compared with the control condition, and it was also significant ($Z = –4.14$, $P < 0.001$). The results indicated heterogeneity among the studies: [$Q = 53.35$, degrees of freedom (d.f.) = 12, $P < 0.001$], $I^2 = 77.5\%$ (see Fig. 1).

Sensitivity analysis: comparison between exercise and control condition on the effect of reduction in depression on trials with high methodological quality

When only those trials that met all three methodological quality criteria was analyzed (in a random effect model), the result showed a substantially reduced, nonsignificant overall effect size in favor of exercise: $g = –0.43$ (95% CI, –1.06, 0.21). The *Q*-value indicated no heterogeneity among the studies ($Q = 2.19$, d.f. = 1, $P = 0.14$), $I^2 = 54.4\%$ (see Fig. 2).

Sensitivity analysis: comparison between exercise and control condition on the effect of reduction in depression on trials that used intention-to-treat analyses

In a model consisting of those trials that used intention-to-treat analyses, a large effect size was found in favor of exercise: $g = –0.70$ (95% CI –1.03, –0.38). The model was significant ($Z = –4.42$, $P < 0.001$), and indicated no heterogeneity among the studies ($Q = 9.27$, d.f. = 5, $P = 0.10$), $I^2 = 46.07\%$ (see Fig. 3).

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Relative weight
McNeil, 1991	–1.422	0.484	0.234	–2.370	–0.474	–2.939	0.003	9.18
Dunn, 2005	–0.833	0.374	0.140	–1.566	–0.099	–2.225	0.026	13.40
Blumenthal, 2007	–0.167	0.228	0.052	–0.614	0.280	–0.731	0.465	23.53
Foley. 2008	–1.202	0.580	0.337	–2.340	–0.065	–2.072	0.038	6.82
de Zeuw, 2010	0.994	0.397	0.158	–1.772	–0.215	–2.501	0.012	12.33
Chalder, 2012	–0.625	0.120	0.014	–0.861	–0.389	–5.187	0.000	34.75
	–0.703	0.165	0.027	–1.027	–0.379	–4.248	0.000	

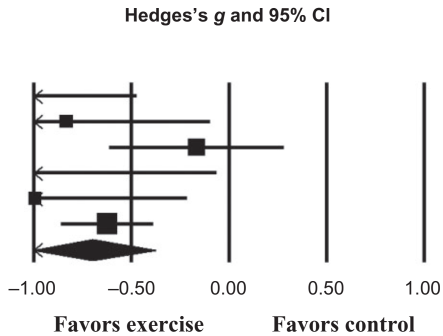

Fig. 3. Meta-analysis on the effect of exercise on depression, including only trials that used intention-to-treat analysis.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value
Setaro, 1985	–1.556	0.319	0.102	–2.181	–0.931	–4.878	0.000
Epstein, 1986	–1.116	0.505	0.255	–2.107	–0.126	–2.209	0.027
Mutrie, 1987	–2.269	0.623	0.388	–3.490	–1.048	–3.643	0.000
McNeil, 1991	–1.422	0.484	0.234	–2.370	–0.474	–2.939	0.003
Mather, 2002	–0.060	0.215	0.046	–0.482	0.361	–0.281	0.779
Dunn, 2005	–0.833	0.374	0.140	–1.566	–0.099	–2.225	0.026
Blumenthal, 2007	–0.167	0.228	0.052	–0.614	0.280	–0.731	0.465
Legrand, 2007	–1.054	0.524	0.274	–2.080	–0.027	–2.012	0.044
Foley, 2008	–0.408	0.536	0.288	–1.459	0.643	–0.761	0.447
de Zeuw, 2010	–0.994	0.397	0.158	–1.772	–0.215	–2.501	0.012
Doyne, 1987	–1.665	0.520	0.270	–2.683	–0.647	–3.205	0.001
	–0.966	0.220	0.048	–1.398	–0.535	–4.393	0.000

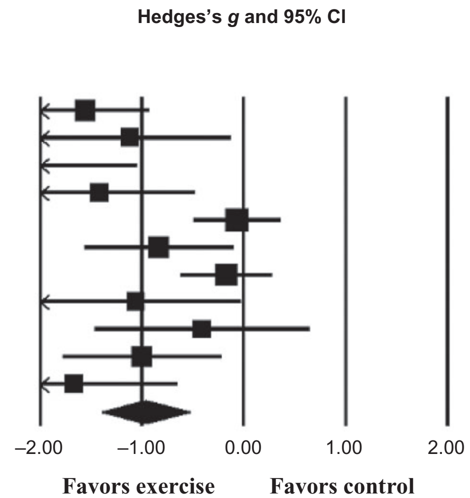

Fig. 4. Meta-analysis on the effect of exercise on depression, including only trials that have compared exercise with no treatment or placebo.

Sensitivity analysis: exercise vs no treatment control and placebo control conditions on the effect of reduction in depression

In this analysis, only trials that compared exercise with no treatment or placebo condition were analyzed in a random effect model. The result showed a larger effect size in favor of exercise than in the main meta analysis where all studies were included in the model (see Fig. 1); $g = –0.97$ (95% CI –1.40, –0.54). The model was also significant ($Z = –4.39$, $P < 0.001$). The Q-value indicated heterogeneity among the studies ($Q = 35.64$, d.f. = 10, $P < 0.001$), $I^2 = 71.94\%$ (see Fig. 4).

Discussion

The results of the meta-analysis offered plausible evidence that physical exercise reduces depression, thereby confirming to greater or lesser extent the results of several earlier meta-analyses (North et al., 1990; Craft & Landers, 1998; Lawlor & Hopker, 2001; Stathopoulou et al., 2006; Landers & Arent, 2007; Mead et al., 2009; Rethorst et al., 2009; Krogh et al., 2011). Despite the fact that the large TREAD study (Chalder et al., 2012) – where exercise had no significant antidepressant effects compared with the control group – was included in the meta-analysis, a large overall effect size was found. A sensitivity analysis showed that the effect size was even larger when only trials with no treatment and placebo control conditions were included in the analysis, suggesting that "usual care" may also have an antidepressant effect. Unfortunately, the content of usual care is seldom specified in clinical research studies, probably due to the fact that clinics tend to administer standard treatment inconsistently and with great variation (Kazdin, 2010). However, it is understandable that researchers still tend to use usual care as a control condition considering the advantages concerning ethical issues; participants in usual care do receive treatment as opposed to no treatment control or placebo. Additionally, usual care may also control for nonspecific therapeutic factors (Kazdin, 2010). Nevertheless, the use of a waiting list control condition may be preferable in studies that aim to determine the effect of exercise on depression.

The results from the present meta-analysis as well as from previous meta-analyses demonstrated that studies with high methodological quality (allocation concealment, intention-to-treat analysis, and blinded outcome assessment) are associated with reduced effect sizes of moderate strength. Previous meta-analyses (e.g., Lawlor & Hopker, 2001, Krogh et al., 2011; Mead et al., 2009) have emphasized that methodological weaknesses of the included trials may seriously have biased the results, leading to a hypothesized exaggerated effect size in favor of exercise. In line with the aforementioned meta-analyses, very few trials in this review met all three criteria for methodological quality (allocation concealment, intention-to-treat analysis, and blinded outcome assessment). In fact, only two of the included studies (Dunn et al., 2005; Blumenthal et al., 2007) passed those criteria. In a subanalysis of these two studies, the effect size was somewhat reduced, although still resulting in a moderate effect size. This is also similar to the results found in previous meta-analyses (e.g., Mead et al., 2009) when the methodologically sound trials are more closely subanalyzed. An exception, however, is Krogh et al. (2011) who found only a small effect size in their subanalysis of three studies considered to be of high methodological quality. Although it should be noted that one of the trials (Krogh et al., 2009) included in that analysis compared exercise with relaxation training, showing that exercise did not differ from relaxation in depression reduction. This is not surprising because relaxation, as previously mentioned, has an antidepressant effect (Reynolds & Coats, 1986). Because the aim in Krogh et al. (2011) was to compare exercise with a placebo condition, the inclusion of this particular study in such a small subanalysis probably biased the result, and subsequently led to an underestimated overall effect size. Furthermore, aside from Krogh et al. (2011), other earlier meta-analyses (Stathopoulou et al., 2006; Mead et al., 2009; Rethorst et al., 2009) have also included trials using placebo conditions that are associated with reductions in depression and anxiety. This may have biased the results in favor of the control group to a certain degree, thus generating an underrated overall effect size. Moreover, Rethorst et al. (2009) showed that adequate allocation concealment and the use of intention-to-treat analysis are associated with higher effect sizes in trials examining the effect of exercise on depression, not the other way around, which has been widely assumed. In this study, however, an additional sensitivity analysis including only trials that had used intention-to-treat analyses showed a large effect size, similar to the overall effect size. Hence, if and to what extent poor methodological quality may have caused exaggerated effect sizes in favor of exercise is still unclear. It is therefore difficult to draw any firm conclusions about the possible influence of methodological weaknesses on the effect size. Meta-analytic results so far indicate that exercise, at least, has a moderate antidepressant effect compared with no treatment and placebo. At best, exercise has a large effect on reductions in depression symptoms and could be categorized as a very useful and powerful intervention. Moreover, earlier meta-analyses also indicate that exercise does not differ from traditional treatment (antidepressant medication, psychotherapy) in reducing depression symptoms (Mead et al., 2009; Rethorst et al., 2009).

The advantages of physical exercise over traditional approaches as an intervention in depressive disorder are manifold; exercise is associated with a wide range of additional physical benefits: stress reduction (Landers & Arent, 2007), decreased blood pressure (Rethorst et al., 2009), reduced risks for coronary artery diseases, weight reduction (Biddle & Mutrie, 2008), increased oxygen uptake (VO2) (Roston et al., 1987), and animal studies also show improvements in cognitive functioning such as learning and memory (van Praag, 2009). In addition, no negative side effects have thus far been reported (as opposed to antidepressant medication) (Biddle & Mutrie, 2008).

Conclusions and directions for future research

To date, it is not possible to determine exactly how effective exercise is in reducing depression symptoms in clinical and nonclinical depressed populations, respectively. However, the results from the present meta-analysis as well as from seven earlier meta-analyses (North et al., 1990; Craft & Landers, 1998; Lawlor & Hopker, 2001; Stathopoulou et al., 2006; Mead et al., 2009; Rethorst et al., 2009; Krogh et al., 2011) indicate that exercise has a moderate to large antidepressant effect. Some meta-analytic results (e.g., Rethorst et al., 2009) suggest that exercise may be even more efficacious for clinically depressed people. Exercise is, of course, not the one and only universal solution to the worldwide growing problem of depression, and this kind of treatment will not be appropriate for all depressed people. On the other hand, this limitation is just as relevant for traditional treatments; that is to say, neither medication nor psychotherapy suits everyone.

In short, our final conclusion is that exercise may well be recommended for people with mild and moderate depression who are willing, motivated, and physically healthy enough to engage in such a program.

There are, nevertheless, several areas in need of further research. First, Callaghan et al. (2011) showed that a specially designed exercise program, including motivational support, was significantly more effective in reducing depression than the regular exercise program. In addition, the tailored exercise group showed significantly greater improvements on a number of other psychological and psychosocial indicators (i.e., quality of life, self-esteem, and mental well-being) compared with the "exercise-as-usual" group. This is a compelling result, suggesting that exercise combined with other features may be more beneficial than exercise alone. This type of promising exercise-based combination treatment tailored for a depressed population needs to be further evaluated in future studies. On the other hand, Chalder et al. (2012) also applied an individually tailored exercise program – the PAF – but this specific type of intervention program did not have a significant effect on depression compared with usual care. Whereas the external validity in Chalder et al. (2012) arguably could be regarded as very good, the internal validity has several weaknesses. For instance, it is not known as to what extent the participants engaged in supervised exercise sessions. Consequently, the study does not tell us whether the exercise sessions were group based or if the participants exercised on their own. More importantly, exercise dose/intensity was not assessed. Because Dunn et al. (2005) showed that only moderate and high exercise dose/intensity significantly reduces depression symptoms (while low exercise dose/intensity does not), this is a crucial aspect to consider when investigating the effect of exercise on depression. Thus, the overall result in Chalder et al. (2012), indicating no antidepressant effect of exercise, needs to be very cautiously interpreted. In fact, the only conclusion that is possible to draw from this study is that this particular exercise intervention – the PAF program – does not differ from usual care in reducing depression. Second, exercise as an adjunct to traditional treatment is also warranted in future studies. Third, because the methodological quality has been poor in many studies, future trials should carefully follow methodological recommendations: having adequate allocation concealment, using intention-to-treat analysis, using blinded outcome assessment, diagnosing depression with a clinical interview, and having a large sample size. Finally, the complex multilevel mechanisms responsible for antidepressant effects associated with exercise need to be explored. Although various attempts have been made to explain the beneficial effects of exercise on depression and anxiety symptoms, the causal link is still unknown, and its "underlying mechanisms are poorly understood" (Cotman et al., 2007; p. 466). Several plausible mechanisms have however been proposed over the years, and the majority of them concern biological, neurological, and physiological changes that are enhanced by exercise, and result in a potential antidepressant effect. Additionally, psychological and psychosocial mechanisms associated with exercise have also been suggested to affect depression (O'Neal et al., 2000; Buckworth & Dishman, 2002; Craft & Perna, 2004; Landers & Arent, 2007; Biddle & Mutrie, 2008; aan het Rot et al., 2009).

Instead of separating biological, neurological, and physiological proposed mechanisms, La Forge (1995) states that several of these mechanisms are overlapping in functions as well as in structure, and because of the complexity of neurobiological systems, he argues that the best possible way to understand why exercise produces mental health effects is to integrate them in one single neurobiological model. Hence, La Forge (1995) emphasizes the importance of understanding the interaction of several neurobiological processes rather than seeing them as separate and isolated mechanisms. Biddle and Mutrie (2008) share La Forge's (1995) aforementioned view and argue that the interactive connection between mind and body is central to employ when seeking to understand the mental health effects of exercise. Consequently, Biddle and Mutrie (2008) suggest an extension of La Forge's (1995) integrative model where not only neurobiological mechanisms are included, but also psychological and psychosocial mechanisms. Hence, in order to get a deeper understanding of how exercise may influence depressive symptoms, an important issue for future research is to develop an integrated model containing biological, physiological, neuorological, and psychological mechanisms.

Perspectives

Physical exercise interventions as a treatment for depression appear to have a moderate to large effect. It is conceivable that that the antidepressant effect could be larger if the exercise-based program would be tailored specifically for a depressed population. Physical exercise may be recommended to the mildly and moderately depressed individual. However, the mechanisms responsible for reductions in depressive symptoms because of exercise are still not very well understood. Thus, examining potential mechanisms is an important issue for future research.

Key words: depression, exercise, meta-analysis, physical activity, review.

References

aan het Rot M, Collins KA, Fitterling HL. Physical exercise and depression. Mount Sinai J Med 2009: 76: 204–214.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th edn. Arlington, VA: APA, 1994.

Astin JA. Why patients use alternative medicine – results of a national study. JAMA 1998: 279: 1548–1553.

Berken GH, Weinstein DO, Stern WC. Weight gain. A side-effect of tricyclic antidepressants. J Affect Disord 1984: 7: 133–138.

Biddle SH, Mutrie N. The psychology of physical activity. London: Routledge, 2008.

Blumenthal JA, Babyak MA, Doraiswamy PM, Watkins L, Hoffman BM, Barbour KA, Herman S, Craighead WE, Brosse AL, Waugh R, Hinderliter A, Sherwood A. Exercise and pharmacotherapy in the treatment of major depressive disorder. Psychosom Med 2007: 69: 587–596.

Borenstein L. Effect sizes for continuous data. In: Cooper H, Hedges LV, Valentine JC, eds. The handbook of research synthesis and meta analysis. New York: Russell Sage Foundation, 2009: 279–293.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester: Wiley, 2009.

Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive

Meta-analysis Version 2. Englewood, NJ: Biostat, 2005.
Buckworth J, Dishman RK. Exercise psychology. Leeds: Human Kinetics, 2002.
Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clin Psychol Rev 2006: 26: 17–31.
Callaghan P. Exercise: a neglected intervention in mental health care? J Psychiatr Ment Health Nurs 2004: 11: 476–483.
Callaghan P, Khalil E, Morres I, Carter T. Pragmatic randomised controlled trial of preferred intensity exercise in women living with depression. BMC Public Health 2011: 11: 465.
Cassano P, Fava M. Depression and public health – an overview. J Psychosom Res 2002: 53: 849–857.
Chalder M, Wiles NJ, Campbell J, Hollinghurst SP, Haase AM, Taylor AH, Fox KR, Costelloe C, Searle A, Baxter H, Winder R, Wright C, Turner KM, Calnan M, Lawlor DA, Peters TJ, Sharp DJ, Montgomery AA, Lewis G. Facilitated physical activity as a treatment for depressed adults: randomised controlled trial. BMJ 2012: 344: e2758.
Chou KL, Lee PW, Yu EC, Macfarlane D, Cheng YH, Chan SS, Chi I. Effect of Tai Chi on depressive symptoms amongst Chinese older patients with depressive disorders: a randomized clinical trial. Int J Geriatr Psychiatry 2004: 19: 1105–1107.
Cohen, J. Statistical power analysis for the behavioral sciences, 2nd edn. New Jersey: Lawrence Erlbaum, 1988.
Cotman CW, Berchtold NC, Christie LA. Exercise builds brain health: key roles of growth factor cascades and inflammation. Trends Neurosci 2007: 30: 464–472.
Craft LL, Landers DM. The effect of exercise on clinical depression and depression resulting from mental illness: a meta-analysis. J Sport Exerc Psychol 1998: 20: 339–357.
Craft LL, Perna FM. The Benefits of Exercise for the Clinically Depressed. Prim Care Companion J Clin Psychiatry 2004: 6: 104–111.
de Zeeuw ELEJ, Tak ECPM, Dusseldorp E, Hendriksen IJM. Workplace exercise intervention to prevent depression: a pilot randomized controlled trial. Ment Health Phys Act 2010: 3: 72–77.
Doyne EJ, Ossipklein DJ, Bowman ED, Osborn KM, McDougallwilson IB, Neimeyer RA. Running versus weight lifting in the treatment of depression. J Consult Clin Psychol 1987: 55: 748–754.
Dunn AL, Trivedi MH, Kampert JB, Clark CG, Chambliss HO. Exercise treatment for depression: efficacy and dose response. Am J Prev Med 2005: 28: 1–8.
Ebmeier KP, Donaghey C, Steele DJ. Recent development and current controversies in depression. Lancet 2006: 367: 153–167.
Epstein D. Aerobic activity versus group cognitive therapy: an evaluative study of contrasting interventions for the alleviation of clinical depression. Reno: University of Nevada, 1986.
Foley LS, Prapavessis H, Osuch EA, De Pace JA, Murphy BA, Podolonsky NJ. An examination of potential mechanisms for exercise as a treatment for depression: a pilot study. Ment Health Phys Act 2008: 1: 69–73.
Forder PM, Gebski VJ, Keech AC. Allocation concealment and blinding: when ignorance is bliss. Med J Aust 2005: 182: 87–89.
Fortney L, Taylor M. Meditation in medical practice: a review of the evidence and practice. Prim Care 2010: 37: 81–90.
Galper DI, Trivedi MH, Barlow CE, Dunn AL, Kampert JB. Inverse association between physical inactivity and mental health in men and women. Med Sci Sports Exerc 2006: 38: 173–178.
Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, Raskob G, Lewis SZ, Schunemann H. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. Chest 2006: 129: 174–181.
Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, Macera CA, Heath GW, Thompson PD, Bauman A. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sport Exer 2007: 39: 1423–1434.
Hassmén P, Koivula N, Uutela A. Physical exercise and psychological well-being: a population study in Finland. Prev Med 2000: 30: 17–25.
Hess-Homeier MJ. A comparison of Beck's cognitive therapy and jogging as treatments for depression. Missoula: University of Montana, 1981.
Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003: 327: 557–560.
Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. Br Med J 1999: 319: 670–674.
Hsu JH, Shen WW. Male sexual side effects associated with antidepressants: a descriptive clinical study of 32 patients. Int J Psychiatry Med 1995: 25: 191–201.
Jorgensen C. Aerobic conditioning in the therapeutic treatment of chronic schizophrenia. Flagstaff, AZ: Northern Arizona University, 1986.
Kazdin AE. Research design in clinical psychology. Boston, MA: Allyn & Bacon, 2010.
Kessler RC, Soukup J, Davis RB, Foster DF, Wilkey SA, Van Rompay MI, Eisenberg DM. The use of complementary and alternative therapies to treat anxiety and depression in the United States. Am J Psychiatry 2001: 158: 289–294.
Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. PLoS Med 2008: 5: 260–268.
Klein MH, Greist JH, Gurman AS, Neimeyer RA, Lesser DP, Bushnell NJ, Smith RE. A comparative outcome study of group psychotherapy vs exercise treatments for depression. Int J Mental Health 1985: 13: 148–176.
Krogh J, Nordentoft M, Sterne JA, Lawlor DA. The effect of exercise in clinically depressed adults: systematic review and meta-analysis of randomized controlled trials. J Clin Psychiatry 2011: 72: 529–538.
Krogh J, Saltin B, Gluud C, Nordentoft M. The DEMO trial: a randomized, parallel-group, observer-blinded clinical trial of strength versus aerobic versus relaxation training for patients with mild to moderate depression. J Clin Psychiatry 2009: 70: 790–800.
La Forge R. Exercise-associated mood alterations: a review of interactive neurobiological mechanisms. Med Exerc Nutr Health 1995: 4: 17–32.
Lan C, Lai J-S, Chen S-Y. Tai Chi Chuan: an ancient wisdom on exercise and health promotion. Sports Med 2002: 32: 217–224.
Landers DM, Arent SM. Physical activity and mental health. In: Tenenbaum G, Eklund RC, eds. Handbook of Sport psychology. Hoboken, NJ: John Wiley & Sons, Inc., 2007: 469–491.
Lawlor DA. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. Author's reply. BMJ 2001: 322: 763.
Lawlor DA, Hopker SW. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. BMJ 2001: 322: 763–767.

Legrand F, Heuze JP. Antidepressant effects associated with different exercise conditions in participants with depression: a pilot study. J Sport Exerc Psychol 2007: 29: 348–364.

Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006: 367: 1747–1757.

Martinsen EW, Medhus A, Sandvik L. Effects of aerobic exercise on depression: a controlled study. Br Med J (Clin Res Ed) 1985: 291: 109.

Mather AS, Rodriguez C, Guthrie MF, McHarg AM, Reid IC, McMurdo MET. Effects of exercise on depressive symptoms in older adults with poorly responsive depressive disorder – randomised controlled trial. Br J Psychiatry 2002: 180: 411–415.

McNeil JK, LeBlanc EM, Joyner M. The effect of exercise on depressive symptoms in the moderately depressed elderly. Psychol Ageing 1991: 6: 487–488.

Mead GE, Morley W, Campbell P, Greig CA, McMurdo M, Lawlor DA. Exercise for depression. Cochrane Database Syst Rev 2009: Art. No: CD004366. doi: 1002/14651858. CD004366.pub4

Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet 2007: 370: 851–858.

Murphy GE, Carney RM, Knesevich MA, Wetzel RD, Whitworth P. Cognitive-behavior therapy, relaxation training, and tricyclic antidepressant medication in the treatment of depression. Psychol Rep 1995: 77: 403–420.

Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: global burden of disease study. Lancet 1997: 349: 1436–1442.

Mutrie N. Exercise as a treatment for moderate depression in the UK National Health Service. Proceedings of sport, health psychology, and exercise symposium (abstract). London: Sports Council and Health Education Authority, 1988: 96–105.

Mutrie N. The relationship between physical activity and clinically defined depression. In: Biddle SJH, Fox KR, Boucher SH, eds. Physical activity and psychological well-being. London: Routledge, 2000: 46–62.

Mutrie N. Healthy body, healthy mind? Psychologist 2002: 15: 412–413.

North TC, McCullagh P, Tran ZV. Effect of exercise on depression. Exerc Sport Sci Rev 1990: 18: 379–415.

O'Neal HA, Dunn AL, Martinsen EW. Depression and exercise. Int J Sport Psychol 2000: 31: 110–135.

Petajan JH, Gappmaier E, White AT, Spencer MK, Mino L, Hicks RW. Impact of aerobic training on fitness and quality of life in multiple sclerosis. Ann Neurol 1996: 39: 432–441.

Pildal J, Chan AW, Hrobjartsson A, Forfang E, Altman DG, Gotzsche PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. BMJ 2005: 330: 1049.

Pinchasov BB, Shurgaja AM, Grischin OV, Putilov AA. Mood and energy regulation in seasonal and non-seasonal depression before and after midday treatment with physical exercise or bright light. Psychiatry Res 2000: 94: 29–42.

Rethorst CD, Wipfli BM, Landers DM. The antidepressive effects of exercise a meta-analysis of randomized trials. Sports Med 2009: 39: 491–511.

Reynolds WM, Coats KI. A comparison of cognitive behavioral-therapy and relaxation training for the treatment of depression in adolescents. J Consult Clin Psychol 1986: 54: 653–660.

Roston WL, Whipp BJ, Davis JA, Cunningham DA, Effros RM, Wasserman K. Oxygen uptake kinetics and lactate concentration during exercise in humans. Am Rev Respir Dis 1987: 135: 1080–1084.

Schultz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. J Am Med Assoc 1995: 273: 408–412.

Segal ZV, Williams M, Teasdale JD. Mindfulness-Based Cognitive Therapy for Depression. New York: Guilford Press, 2002.

Setaro JL. Aerobic exercise and group counseling in the treatment of anxiety and depression. Ann Arbor: University of Maryland, 1985.

Shadish WR, Haddock K. Analyzing effect sizes: fixed-effects models. In: Cooper H, Hedges LV, Valentine JC, eds. The handbook of research synthesis and meta analysis. New York: Russell Sage Foundation, 2009: 257–277.

Singh NA, Clements KM, Fiatarone MA. A randomized controlled trial of progressive resistance training in depressed elders. J Gerontol A Biol Sci Med Sci 1997: 52: M27–M35.

Singh NA, Stavrinos TA, Scarbek Y, Galambos G, Liber C, Singh MAF. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. J Gerontol A Biol Sci Med Sci 2005: 60: 768–776.

Stathopoulou G, Powers MB, Berry AC, Smits JAJ, Otto MW. Exercise interventions for mental health: a quantitative and qualitative review. Clin Psychol Sci Pract 2006: 13: 179–193.

Tloczynski J, Tantriella M. A comparison of the effects of Zen breath meditation or relaxation on college adjustment. Int J Psych Orient 1998: 41: 32–43.

Trindade E, Menon D, Topfer LA, Coloma C. Adverse effects associated with selective serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. CMAJ 1998: 159: 1245–1252.

Tsang HWH, Chan ER, Cheung WM. Effects of mindful and non-mindful exercises on people with depression: a systematic review. Br J Clin Psychol 2008: 47: 303–322.

Tsang HWH, Fung KMT, Chan ASM, Lee G, Chan F. Effect of qigong exercise programme on elderly with depression. Int J Geriat Psychiatry 2006: 21: 890–897.

Unutzer J, Klap R, Sturm R, Young AS, Marmon T, Shatkin J, Wells KB. Mental disorders and the use of alternative medicine: results from a national survey. Am J Psychiatry 2000: 157: 1851–1857.

van der Watt G, Laugharne J, Janca A. Complementary and alternative medicine in the treatment of anxiety and depression. Curr Opin Psychiatry 2008: 21: 37–42.

van Praag H. Exercise and the brain: something to chew on. Trends Neurosci 2009: 32: 283–290.

Veale D, Lefevre K, Pantelis C, Desouza V, Mann A, Sargeant A. Aerobic exercise in the adjunctive treatment of depression – a randomized controlled trial. J Roy Soc Med 1992: 85: 541–544.

Wang CC, Collet JP, Lau J. The effect of Tai Chi on health outcomes in patients with chronic conditions – a systematic review. Arch Intern Med 2004: 164: 493–501.

Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. Can Med Assoc J 2006: 174: 801–809.

Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, Gluud C, Martin RM, Wood AJG, Sterne JAC. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. Br Med J 2008: 336: 601–605.