Review article

Exercise as a treatment for depression: A meta-analysis

Siri Kvam a,*, Catrine Lykkedrang Kleppe b, Inger Hilde Nordhus c, Anders Hovland c,d

a Sogndal Child and Adolescent Psychiatric Outpatient Clinic, Division of Medicine, District General Hospital of Førde, postboks 1000, 6807 Førde, Norway
b Child Habilitation Unit, Department of Pediatrics, Sørlandet Hospital, Norway
c Department of Clinical Psychology, University of Bergen, Norway
d Solli District Psychiatric Centre (DPS), Norway

A R T I C L E   I N F O

Article history:
Received 15 December 2015
Received in revised form 29 February 2016
Accepted 12 March 2016
Available online 20 May 2016

Keywords:
Physical exercise
Depression
Depressive disorder
Treatment outcome
Meta-analysis

A B S T R A C T

Background: This meta-analysis of randomized controlled trials (RCTs) examines the efficacy of physical exercise as treatment for unipolar depression, both as an independent intervention and as an adjunct intervention to antidepressant medication.

Methods: We searched PsycINFO, EMBASE, MEDLINE, CENTRAL, and Sports Discus for articles published until November 2014. Effect sizes were computed with random effects models. The main outcome was reduction in depressive symptoms or remission.

Results: A total of 23 RCTs and 977 participants were included. Physical exercise had a moderate to large reduction in depressive symptoms or remission.

Conclusions: Physical exercise is an effective intervention for depression. It also could be a viable adjunct treatment in combination with antidepressants.
1. Introduction

According to the World Health Organization (WHO), depression is a global disease, with over 350 million people affected (WHO, 2012). The WHO estimates that depression will be the second leading cause of global burden of disease worldwide by 2020 (WHO, 2001). Major depressive disorder (MDD) is the most prevalent mental disorder (Kessler et al., 2005; Wittchen and Jacobi, 2005), with a lifetime prevalence of 6–15% (Bromet et al., 2011). Depression is one of the most common diagnoses in primary health care (WHO, 2001). Depression reduces health more than somatic diseases such as arthritis, angina, and diabetes do (Moussavi et al., 2007), and depression was one of the leading causes of disability in 2012 (WHO, 2012). The mortality rate of depression is about 4%, which is equivalent to that of smoking (NICE, 2013). For patients suffering from somatic diseases such as cancer, cardiovascular diseases, and infections, the mortality risk increases even further with comorbid depression (Mykletun et al., 2009).

Guidelines from the National Institute for Health and Care Excellence (NICE) recommend the psychological treatments cognitive behavioral therapy (CBT) or interpersonal therapy (IPT) as treatment of choice for mild to moderate depression, followed by antidepressant medication (NICE, 2013). Many patients do not achieve sufficient symptom relief despite adequate treatment implementation (Bahr, 2009), and 50% experience at least one new depressive episode (Helsedirektoratet, 2009). Research points to the importance of maximizing the response to treatment as early as possible due to declining prognosis with depression duration and failed treatment responses (NICE, 2013; Trivedi et al., 2006). A subsequent treatment is often required to increase the effect (Major et al., 2011), either by changing treatment or by adding one (Wisniewski et al., 2007). Given the serious consequences of depression, cost-effective and robust interventions to establish recovery and prevent relapse are crucial.

Both the WHO (WHO, n.d.) and the NICE guidelines (NICE, 2013) recommend implementing physical exercise (henceforth referred to as exercise) in the standard treatment of depression. It is thus important to have updated knowledge on the effect of exercise on depression. Several reviews provide support for the antidepressant effect of exercise (Blake et al., 2009; Bridle et al., 2012; Cooney et al., 2013; Joseffson et al., 2014; Krogh et al., 2011; Mura et al., 2014; Rethorst et al., 2009; Robertson et al., 2012; Silveira et al., 2013). However, most of these previous reviews (Blake et al., 2009; Bridle et al., 2012; Cooney et al., 2013; Joseffson et al., 2014; Mura et al., 2014; Rethorst et al., 2009; Robertson et al., 2012) have not distinguished between depressive symptoms and depression as a diagnosis fulfilling certain diagnostic criteria according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD). In evaluation of the efficacy of exercise as a treatment in a clinical setting, Krogh et al. (2011) advocate that inclusion of studies should be limited to those including participants with a diagnosis of depression. To the best of our knowledge, the meta-analyses by Krogh et al. (2011) and Silveira et al. (2013) are the only ones that have investigated the effect of exercise on depression exclusively in participants with a diagnosis of unipolar depression. However, Krogh et al. and Silveira et al. completed the search for articles in 2008 and 2011, respectively, and Silveira et al. included non-randomized controlled trials. Hence, an update in this important domain is warranted, and we seek to fill this gap. In an attempt to increase the generalizability of the knowledge about treatment for the depressed population that most commonly needs health care, we include only studies on participants with a diagnosis of unipolar depression.

The majority of the previous meta-analyses compared exercise with different types of controls (waiting list, placebo, other treatment). This is common practice, but complicates the interpretation of results, as exercise compared to other treatments is likely to yield smaller effect sizes than exercise compared to no intervention. The terms “no intervention”, “waiting list”, and “usual care” have been used interchangeably in previous reviews and studies (for instance, in Rethorst et al. (2009) and Sims et al. (2009)). Furthermore, the term “usual care” can encompass a range of interventions of varying efficacy (Freedland et al., 2011). It is important to ascertain the relative effect of exercise on depression in this regard, as the probability of response to treatment and prognosis decreases with duration of depression (NICE, 2013), and because many patients with depression go untreated or undertreated (González et al., 2010; Health and Social Care Information Centre and Community and Mental Health statistics team, 2014; WHO, 2012). Rethorst et al. (2009) included only no-intervention or waiting list controls; however, they also included nonclinical participants. To investigate whether exercise could provide symptom relief while awaiting access to the treatments of choice for depression, we specifically compare exercise groups to no-treatment or waiting list controls. Since we believe that no intervention or waiting list is different from usual care, we include a
separate analysis with usual care controls. To assess the effect of exercise compared to treatments recommended by the NICE (2013), we also compared exercise to psychological treatments and antidepressant medication. Rethorst et al. (2009) and Cooney et al. (2013) have also compared exercise to these controls, but as previously described, they also included nonclinical participants in their analyses. Thus, the present study is, to the best of our knowledge, the first meta-analysis to compare the effects of exercise with these control conditions for patients with unipolar depression.

Due to declining prognosis with depression duration and failed treatment responses (NICE, 2013), the need for maximizing the effect of interventions for depression is paramount. The WHO recommends exercise in combination with antidepressants or psychotherapy (WHO, n.d.), and a recent review indicates that exercise in combination with antidepressants can be effective (Mura et al., 2014). To the best of our knowledge, no meta-analysis has previously investigated this in a population suffering from unipolar depression. We therefore assess the relative effect of the combined treatment of antidepressant medication and exercise for persons with unipolar depression.

The Cochrane Library has published several comprehensive and rigorous reviews regarding the effect of exercise on depression (Cooney et al., 2013; Mead et al., 2009; Rimer et al., 2012). We have extended and refined the thorough electronic searches executed in these Cochrane reviews to provide an exhaustive compilation of studies according to our current research aims. Unlike these previous Cochrane reviews, we will only include studies in which participants had a clinical diagnosis of unipolar depression.

Considering both the frequent occurrence and the severe impact of depression, updated knowledge on a readily available intervention, such as exercise, is highly warranted, and the objective of the present meta-analysis is to investigate the effect of exercise as a treatment for unipolar depression. We will examine the effect of exercise compared to active control conditions (psychological treatments and antidepressant medication), as well as to usual care and to no intervention. We will also assess the effect of exercise as an adjunct to treatment with antidepressant medication. These five comparisons have not previously been specifically investigated for patients suffering from the common and severe diagnosis of unipolar depression.

2. Methods

2.1. Eligibility criteria

We evaluated randomized controlled trials (RCTs) published in Norwegian, Danish, Swedish, English, and Spanish for inclusion. The allocation of participants to conditions had to be described as randomized, including terms such as “randomly”, “randomization”, and “random”, and the exercise group had to be compared to a control group.

2.2. Participants

The participants were adults 18 years or older, of both genders, in any setting, and with a diagnosis of unipolar depression according to DSM- or ICD criteria. Some of the studies also included a small portion of participants diagnosed with minor depression or dysthymia (Doyne et al., 1987; Pinchasov et al., 2000; Singh et al., 2005, 1997), and in line with the previous meta-analyses by Krogh et al. (2011) and Silveira et al. (2013), we included these studies in the current meta-analysis. We excluded studies in which participants had seasonal depression or bipolar diagnosis. Comorbid mental disorders were accepted as long as unipolar depression was the primary mental disorder.

2.3. Types of conditions

Any kind of aerobic exercise (e.g., walking, running, cycling) and nonaerobic exercise (e.g., resistance exercise, strength exercise, weight lifting) was included, even if in combination with other antidepressant treatment. We excluded studies with tai chi, qigong, yoga, etc., and studies that only gave guidance about exercise but did not provide an exercise program. The control group could receive placebo treatment, usual care, another kind of treatment (e.g., antidepressant medication, psychotherapy, or alternative interventions), no treatment, or be on a waiting list. We excluded studies on the acute effect of exercise, with measures taken immediately after one exercise bout, studies with an exercise program lasting less than a week, and studies without a nonexercising control group.

2.4. Outcome measures

The studies had to include a primary outcome measure on symptoms of depression, assessed with a validated scale on depression severity or indicated by remission. Numeric data had to be reported. We used the outcome data from the main outcome measure of depression (indicated by the authors as such, or the first result mentioned in the abstract and results section) in estimation of effect size. If the main outcome was dichotomous, a continuous validated outcome measure was preferred if available. When the study had multiple arms, we used the arms that gave the largest clinical effect.

2.5. Literature search

2.5.1. Electronic searches

The present meta-analysis builds on and extends the electronic searches in the Cochrane reviews “Exercise for depression” (Cooney et al., 2013; Mead et al., 2009; Rimer et al., 2012). We searched the databases Sports Discus, PsycINFO, MEDLINE, EMBASE, and Cochrane Central Register of Controlled Trials (CENTRAL), with keywords such as “exercise”, “physical activity”, “running”, “walking”, “depression”, “depressive disorder”, “randomized controlled trial”, “controlled clinical trial”, and “randomly”. For Sports Discus, our searches covered articles published from January 2007 to November 24, 2014, and in the remaining databases, from January 2010 to November 20, 2014. Appendix A-E (Supplemental Materials) provides detailed search strategies for all databases.

2.5.2. Search in other resources

We went through the bibliography of 9 reviews (Cooney et al., 2013; Greer and Trivedi, 2009; Josefsen et al., 2014; Kelley et al., 2015; Krogh et al., 2011; Mura et al., 2014; Perraton et al., 2010; Rethorst et al., 2009; Robertson et al., 2012) and the lists of “studies awaiting assessment” and “ongoing studies” in two of these (Cooney et al., 2013; Rimer et al., 2012). Further, we checked the bibliography of the latest published studies included in our review (Gary et al., 2010; Hemat-Far et al., 2012; Krogh et al., 2012; Mota-Pereira et al., 2011; Salehi et al., 2014; Schuch et al., 2011) and contacted the main authors through email to inquire about unpublished material or ongoing studies.

2.6. Study selection and quality assessment

After scrutinizing the records, two of the authors (SK and CLK) independently evaluated full-text articles based on the eligibility criteria. The same authors also assessed three aspects of risk of
bias in the included studies: use of allocation concealment, blinding of assessor of main outcome assessment (henceforth referred to as blinded outcome), and use of intention-to-treat analysis. When uncertainty or disagreement occurred, we contacted the main author of the study. If further information was not provided, we evaluated the study as not fulfilling the particular aspect. We also assessed how many of the studies included in the current meta-analysis were registered in a trial registry, in line with current recommendations.

2.7. Data analysis

The program Comprehensive Meta-Analysis version 2 was used to conduct all data analyses (Borenstein et al., 2005), and effect sizes were calculated with a random effects model. Hedges’ g was used as a measure of effect size. By convention, we interpreted effect sizes of 0.2 as a small effect, 0.5 as a moderate effect, and 0.8 as a large effect (Schünenmann et al., 2008). The level of significance was set to p < 0.05. We used the Q-statistic together with I² to determine the level of heterogeneity. To investigate and adjust for publication bias, we examined the fail-safe N and the funnel plot, and we performed a trim-and-fill algorithm for analysis of the main effect.

3. Results

3.1. Results of the literature search

The electronic searches yielded 3346 records. After deletion of duplicates, we screened records to exclude obvious nontarget records manually, before reading abstracts and continuing to remove irrelevant records. Search in additional sources yielded one study of Blumenthal et al. (1999) are reported in Babyak et al. (2000); from Blumenthal et al. (2007), in Hoffman et al. (2011); following reasons: participants did not have a diagnosis of unipolar depression (Akandere and Demir, 2011; Annesi and Gorpala, 2010; Blumenthal et al., 2012a; Bonnet, 2005; Bradas et al., 2010; Brenes et al., 2007; Cakt et al., 2010; Carraro and Gobbi, 2014; Cassilhas et al., 2010; Chen et al., 2012; Christensen et al., 2012; Chrysouhoou et al., 2014; Chu, 2008; Collins et al., 2011; Conradsson et al., 2010; Dalgas et al., 2010; De Zeeuw et al., 2010; Effing et al., 2011; Elrdige et al., 2011; Ellard et al., 2014; Faulconbridge et al., 2012; Fremont and Craighead, 1987; Gallagher et al., 2014; Ha and Choi, 2014; Harris et al., 2010; Hein et al., 2011; Hess-Homeier, 1981; Hoffman et al., 2010; Holmgren et al., 2010; Howarter et al., 2014; Hughes et al., 2012; Imayama et al., 2011; Jhonnasson et al., 2011; Khalaf and Fathy, 2011; Kline et al., 2012; Korstjens et al., 2011; Koudi et al., 2010; Lee et al., 2010; Legrand and Heuze, 2007; Legrand and Thatcher, 2011, 2014; Levinger et al., 2011; Lim and Hong, 2010; Lincoln et al., 2011; López-Rodríguez et al., 2012; Maci et al., 2012; Mailey et al., 2010; Marzolini et al., 2009; Mata et al., 2013; Matthews et al., 2011; McAuley et al., 2010; McGale et al., 2011; McKenna et al., 2013; McNeil et al., 1991; Mehnerst et al., 2011; Mendes et al., 2010; Meshcheryakova et al., 2010; Middleton et al., 2012; Midgaard et al., 2011; Moros et al., 2010; Mortazavi et al., 2012; Mosquera-Valderrama et al., 2012; Murtezani et al., 2011; Nabkasorn et al., 2008; Norman et al., 2010; Orth, 1979; Pala Özdemir et al., 2010; Penttinen et al., 2011; Perna et al., 2010; Pfaff et al., 2014; Pinniger et al., 2012; Piscbek et al., 2010; Reid et al., 2010; Rethorst et al., 2011, 2010; Robledo-Colonia et al., 2012; Rosenberg et al., 2010; Sabapathy et al., 2011; Sañudo et al., 2012, 2011, 2010; Scheewe et al., 2013; Setaro, 1985; Shahidi et al., 2011; Sprod et al., 2010; Sturm et al., 2012; Saarto et al., 2012; Van Citters et al., 2010; Vural et al., 2014; Williams and Tappen, 2008; Wise et al., 2012; Yang et al., 2011; Österås et al., 2012), the article did not specify what kind of mood disorders the participants suffered from and more precise information could not be obtained (Mather et al., 2002), the studies included participants with other primary mental disorders (Knubben et al., 2007; Oeland et al., 2010), they were not RCTs (Banting et al., 2014; Cerda et al., 2011; Dereli and Yalım, 2010; Gutierrez et al., 2012; Justine and Hamid, 2010; Milani et al., 2011; Silveira et al., 2010), they lacked a control group (Beckie et al., 2011; Callaghan et al., 2011; Chan et al., 2012; Craft et al., 2007; Greenwood et al., 2012; Kerr et al., 2008; Mata et al., 2012; White et al., 2009) or a nonexercising control group (Baek et al., 2014; Toups et al., 2011; Trivedi et al., 2011), they used an exercise intervention not in line with the eligibility criteria (Arcos-Carmona et al., 2011; Casasas et al., 2012; Chalder et al., 2012; García-Toro et al., 2012; Ma et al., 2010; Martiny et al., 2012; McClure et al., 2011; Piette et al., 2011; Ström et al., 2013; Watkins et al., 2012), the participants were younger than 18 years old (Roshan et al., 2011), they did not provide numerical data (Greist et al., 1979; McCann and Holmes, 1984), or they used a single-subject design (Bonnet, 2005). This evaluation led us with 23 studies for the present meta-analysis (Blumenthal et al., 2012b, 2007, 1999; Doyne et al., 1987; Dunn et al., 2005; Epstein, 1986; Foley et al., 2008; Gary et al., 2010; Hemat-Far et al., 2012; Klein et al., 1985; Krogh et al., 2012, 2009; Martinsen et al., 1985; Mata-Pereira et al., 2011; Mutrie, 1986, 1988; Pilu et al., 2007, 2012; Salehi et al., 2014; Schuch et al., 2011; Sims et al., 2009; Singh et al., 2005, 1997; Veale et al., 1992). Results from follow-up assessments are reported for seven studies (Blumenthal et al., 2007, 1999, 2011; Gary et al., 2010; Klein et al., 1985; Krogh et al., 2007, 2009; Sims et al., 2009; Singh et al., 1997). Follow-up measures from the study of Blumenthal et al. (1999) are reported in Babya et al. (2011); from Blumenthal et al. (2007), in Hoffman et al. (2011); and from Singh et al. (1997), in Singh et al. (2001). Results from the literature search and evaluation of studies are summarized in Fig. 1.

3.2. Study characteristics

Table 1 shows an overview of study characteristics. In addition to MDD, four studies included participants with minor depression and dysthymia (Doyne et al., 1987; Pinchasov et al., 2000; Singh et al., 2005, 1997). In two studies, the participants had treatment-resistant MDD (Pilu et al., 2007; Salehi et al., 2014). In one study, 47 of 100 participants met the criteria for unipolar depression (Blumenthal et al., 2012b), and in another study (Pinchasov et al., 2000), 18 of 63 participants had nonseasonal depression. Outcome data for only these subgroups were included in the analyses. Three studies explicitly stated that the participants had a somatic comorbid diagnosis of coronary heart disease (Blumenthal et al., 2012b), heart failure (Gary et al., 2010), and stroke (Sims et al., 2009). Sixteen studies provided aerobic exercise (Blumenthal et al., 2012b, 2007, 1999; Dunn et al., 2005; Epstein, 1986; Foley et al., 2008; Gary et al., 2010; Hemat-Far et al., 2012; Klein et al., 1985; Krogh et al., 2012; Martinsen et al., 1985; Mata-Pereira et al., 2011; Pinchasov et al., 2000; Salehi et al., 2014; Schuch et al., 2011; Veale et al., 1992), four studies provided resistance training (Pilu et al., 2007; Sims et al., 2009; Singh et al., 2005, 1997), and three studies provided both types of exercise (Doyne et al., 1987; Krogh et al., 2009; Mutrie, 1986). The outcome assessor was blinded in 10 studies (Blumenthal et al., 2012b, 2007, 1999; Dunn et al., 2005; Gary et al., 2010; Krogh et al., 2012, 2009; Mata-Pereira et al., 1997).
allocation to condition was adequately concealed in 11 studies (Blumenthal et al., 2012b, 2007, 1999; Dunn et al., 2005; Krogh et al., 2012, 2009; Martinsen et al., 1985; Salehi et al., 2014; Sims et al., 2009; Singh et al., 1997, 2005), and intention-to-treat analysis was used in 12 studies (Blumenthal et al., 2012b, 2007, 1999; Dunn et al., 2005; Foley et al., 2008; Gary et al., 2010; Krogh et al., 2012, 2009; Martinsen et al., 1985; Mutrie, 1986; Sims et al., 2009; Singh et al., 1997). Table 1 provides more details on study characteristics. We found that 5 of the 23 included studies in the present meta-analysis were pre-registered in a trial registry (Blumenthal et al., 2012b, 2007; Krogh et al., 2012, 2009; Salehi et al., 2014).-F.

3.3. Effect of interventions

3.3.1. Analysis 1. Exercise versus control

Based on 23 studies (977 participants), we found that reduction in depressive symptoms after treatment showed a moderate to large and significant effect in favor of exercise (Fig. 2), \( g = -0.68 \) (95% CI = \(-0.92\) to \(-0.44\), \( p < 0.001\)). The heterogeneity between studies was significant and moderate to high (\( Q (22) = 68.737, p < 0.001, I^2 = 67.99\)). See Appendix F Table 2 (Supplemental Materials) for data on means, standard deviations, and sample size post treatment for the included studies.

3.3.2. Assessment of publication bias

The fail-safe N analysis for the main analysis (i.e., analysis 1) showed that 463 studies with no effect of exercise are required before the \( p \)-value is no longer statistically significant. The funnel plot (Fig. 3) was asymmetrical. The trim-and-fill analysis indicated that zero studies are missing from the left side of the analysis and eight studies from the right (Fig. 3). If the missing studies were added, the adjusted effect size \( g \) for this analysis would have been \(-0.38\) (95% CI = \(-0.62\) to \(-0.13\), \( p = 0.002\)), which is a small to moderate and significant effect size.

3.3.3. Subgroup analyses

We performed two subgroup analyses to compare studies with or without blinded outcome or intention-to-treat. In the first subgroup analysis (Fig. 4), the studies without blinded outcome (13 studies, 377 participants) showed a large and significant effect in favor of exercise, \( g = -0.91 \) (95% CI = \(-1.22\) to \(-0.61\), \( p < 0.001\)), while the effect for the studies with blinded outcome (10 studies, 600 participants) was moderate and significant, \( g = -0.40 \) (95% CI = \(-0.63\) to \(-0.16\), \( p = 0.001\)).
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>N</th>
<th>Age</th>
<th>Sex</th>
<th>Conditions</th>
<th>Exercise</th>
<th>Duration</th>
<th>Main outcome</th>
<th>Quality</th>
</tr>
</thead>
</table>
| Blumenthal (1999)     | Community volunteers recruited through flyers, media advertisements, letters to physicians and mental health facilities | 156  | 57 (6.5)  | Mix  | 1) Supervised exercise  
2) Combination: Exercise + Sertraline  
3) Medicine: Sertraline (SSRI) | Walking/jogging  
3 sessions/week  
45 min  
70–85% of HRmax | 16       | HAMD-17   | B/A/I   |
| Blumenthal (2007)     | Outpatients recruited through media                                           | 202  | 52 (8)    | Mix  | 1) Supervised aerobic exercise  
2) Home-based aerobic exercise: Same prescription, minimal contact from study staff  
3) Medicine: Sertraline (SSRI)  
4) Placebo-pill | Walking/jogging  
3 sessions/week  
45 min  
70–85% of HRmax | 16       | Remission MDD | B/A/I   |
| Blumenthal (2012b)    | Outpatients with coronary heart disease                                       | 101  | 63,9      | Mix  | 1) Exercise  
2) Medicine: Sertraline (SSRI)  
3) Placebo-pill | Walking/jogging  
3 sessions/week  
30 min  
70–85% of HRmax | 16       | HAMD-17   | B/A/I   |
2) Nonaerobic exercise: Weight lifting, 50–60% HRmax  
3) Waiting list | Individually walking/running  
4 session/week  
7 min intervals (not stated for how long)  
80% HRmax | 8        | BDI       |         |
| Dunn (2005)           | Community volunteers                                                          | 80   | 35.9 (6.4) | Mix  | 1) Low dose LD/3: 3 sessions/week, 7 kcal/kg/week  
2) LD/5: 5 sessions/week  
3) Public Health dose PHD/3: 3 sessions/week  
4) PHD/5: 5 sessions/week  
5) Placebo: Stretching flexibility exercise 3 sessions/week, 15–20 min | Supervised individual aerobic exercise (treadmill or stationary bicycle)  
PHD: 17.5 kcal/kg/week | 12       | HAMD-17   | B/A/I   |
| Epstein (1986)        | Recruited through media advertisement.                                       | 26   | 39,4      | Mix  | 1) Group aerobic exercise  
2) Group CT: 1 session/week, 90 min  
3) Waiting list | Walking/running  
3 sessions/week  
30 min  
70% HRmax | 8        | BDI       |         |
| Foley (2008)          | Recruited through advertisement in media and referral from a psychiatrist     | 23   | 18–55 years | No   | 1) Supervised aerobic exercise  
2) Stretching: Supervised, 3 sessions/week, 30–40 min, mild-intensity | 3 sessions/week  
30–40 min, moderate intensity | 12       | BDI-II    | 1        |
| Gary (2010)           | Outpatients with a diagnosis of NYHA Class II to III heart failure and a BDI-II score of 10 or higher were asked to participate in the study | 74   | 65,8 (13,5) | Mix  | 1) Home-based exercise  
2) Individual CBT  
3) Usual care: No information or counselling from their health care provider other than that normally provided  
4) Combination: Exercise + CBT | Not supervised, monitored by home visits once a week Advised to walk up to 1 h at moderate intensity  
3 sessions/week | 12       | HAMD-17   | B/I     |
<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Population</th>
<th>Sample Size</th>
<th>Age (Mean ± SD)</th>
<th>Setting</th>
<th>Intervention Details</th>
<th>Outcome Measures</th>
<th>Duration and Frequency</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemat-Far (2012)</td>
<td>University students selected by psychiatrist</td>
<td>20</td>
<td>18–25 years F</td>
<td>1) Exercise 2) Control: Asked to pursue their normal life without doing exercise</td>
<td>Supervised running 3 sessions/week 40–60 min 60–65% HRmax</td>
<td>BDI</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Klein (1985)</td>
<td>Outpatients recruited through newspaper advertisement</td>
<td>74</td>
<td>30,1 (6,7) Mix</td>
<td>1) Running 2) Group therapy: Semistructured, interpersonal therapy and CT, 2 h/week 3) Group meditation: Supervised, 2 h/week</td>
<td>Supervised, individually 2 Sessions/week 45 min</td>
<td>SCL-90</td>
<td>12</td>
<td>FU: 9 m</td>
</tr>
<tr>
<td>Krogh (2009)</td>
<td>Outpatients referred by a medical doctor or a psychologist</td>
<td>165</td>
<td>38,9 (9,46) Mix</td>
<td>1) Supervised aerobic exercise: Different aerobic exercises (cycling, rowing, running), 70–89% of HRmax 2) Supervised strength exercise 3) Relaxation: 2 sessions/week. Supervised, relaxation and balance exercises, back massage</td>
<td>Circuit training, 2 Sessions/week 90 min 50–75% of 1RP</td>
<td>HAMD-17 B/A/I</td>
<td>16</td>
<td>FU: 12 m</td>
</tr>
<tr>
<td>Krogh (2012)</td>
<td>Outpatients referred from a clinical setting by a physician or a psychologist</td>
<td>115</td>
<td>41,6 Mix</td>
<td>1) Supervised aerobic exercise 2) Group stretching exercise: Supervised, 3 sessions/week</td>
<td>Stationary ergometer cycle 3 Sessions/week 45 min</td>
<td>HAMD-17 B/A/I</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Martinsen (1985)</td>
<td>Inpatients</td>
<td>49</td>
<td>40 Mix</td>
<td>1) Supervised aerobic exercise 2) Control: Occupational therapy</td>
<td>3 Sessions/week 60 min 50–70% of HRmax</td>
<td>BDI A/I</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Mata-Pereira (2011)</td>
<td>Outpatients at a psychiatry clinic</td>
<td>33</td>
<td>47 Mix</td>
<td>1) Supervised aerobic exercise 2) Control: No treatment first 4 weeks, combination of aerobic and nonaerobic exercise after 4 weeks</td>
<td>Home-based walking 5 Sessions/week (1 walk per week supervised) 30–45 min</td>
<td>HAMD-17 B</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Mutrie (1986)</td>
<td>Referred by GPs</td>
<td>36</td>
<td>42,1 (13.4) Mix</td>
<td>1) Aerobic exercise 2) Nonaerobic exercise: Vice versa of exercise prescription for aerobic group 3) Control: No treatment first 4 weeks, combination of aerobic and nonaerobic exercise after 4 weeks</td>
<td>Only aerobic exercise (minimum 20 min) first 4 weeks, strength- and stretching-exercises added after 4 weeks</td>
<td>BDI 4&quot;</td>
<td>8</td>
<td>4&quot;</td>
</tr>
<tr>
<td>Pilu (2007)</td>
<td>Recruited from University Psychiatric if diagnosed with MDD and no response to pharmacological treatment</td>
<td>30</td>
<td>Not stated</td>
<td>1) Medicine: SSRI, SNRI, NARI, TCA 2) Combination: Exercise + Medicine</td>
<td>Supervised Group strengthening exercise 2 Sessions/week 60 min</td>
<td>HAMD</td>
<td>8 m</td>
<td></td>
</tr>
<tr>
<td>Pinchasov (2000)</td>
<td>Nondepressed group recruited from medical staff. Depressed participants recruited from patients at the hospital were the investigation was carried out</td>
<td>63 18'</td>
<td>35,2 (9,3) F</td>
<td>1) BLT: 2500 lux cool-white incandescent light, administered daily between 14.00 and 16.00 h 2) Exercise</td>
<td>Stationary bicycle Daily for 1 week 60 min</td>
<td>HAMD-21</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Salehi (2014)</td>
<td>Inpatients with TR-MDD</td>
<td>60</td>
<td>31,45 (6,75) Mix</td>
<td>1) Supervised aerobic exercise 2) ECT: 3 times/week 3) Combination: Exercise + ECT</td>
<td>Individual treadmill exercise 3 session/week 30 min 60–75% of VO2max</td>
<td>HAMD-21 B/A</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>N</td>
<td>Age</td>
<td>Sex</td>
<td>Conditions</td>
<td>Exercise</td>
<td>Duration</td>
<td>Main outcome</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Schuch (2011)</td>
<td>Inpatients at a University hospital</td>
<td>26</td>
<td>42,7</td>
<td>Not stated</td>
<td>1) Conventional treatments of pharmacological and/or ECT 2) Combination: Supervised exercise + Conventional treatments</td>
<td>Individual, aerobic exercise of own choice (stationary bicycle, a treadmill or an elliptic) 3 Sessions/week 16.5 kcal/kg/week</td>
<td><strong>HAMD-17</strong></td>
<td></td>
</tr>
<tr>
<td>Sims (2009)</td>
<td>Recruited from hospital databases of discharged stroke patients, GPs and newspaper</td>
<td>45</td>
<td>67,13 (15,2)</td>
<td>Mix</td>
<td>1) Supervised progressive resistance training (PRT) 2) Usual care: Asked to not do any progressive resistance training</td>
<td>Group strengthening exercises using machine weights 2 Sessions/week 80% 1RP</td>
<td>10 FU: 6 m</td>
<td>CESD</td>
</tr>
<tr>
<td>Singh (1997)</td>
<td>Recruited from community through two volunteer databases</td>
<td>32</td>
<td>71.3 (1,2)</td>
<td>Mix</td>
<td>1) Supervised PRT 2) Health education program: 1–9 subjects in group, asked not to begin any exercise during intervention, 2 sessions/week</td>
<td>1–8 subjects exercised simultaneously 3 Sessions/week 45 min 80% 1RP</td>
<td>10 FU: 26 m</td>
<td>BDI</td>
</tr>
<tr>
<td>Singh (2005)</td>
<td>Recruited through their GPs</td>
<td>60</td>
<td>69</td>
<td>Mix</td>
<td>1) <em>High-intensity PRT</em> 2) Low-intensity PRT: Same regimen, except at 20% 1RM and did not progress 3) <em>Standard care from GP</em></td>
<td>3 Sessions/week 45 min 80% of 1RM</td>
<td>8 HAMD</td>
<td></td>
</tr>
<tr>
<td>Veale (1992)</td>
<td>Depressed outpatients currently receiving standard care</td>
<td>83</td>
<td>35,5</td>
<td>Mix</td>
<td>1) Supervised aerobic exercise 2) Control: No extra treatment besides standard care</td>
<td>Running program 3 Sessions/week</td>
<td>12 BDI</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**

- **Participants:** MDD = Major depressive disorder. TR-MDD = Treatment-resistant major depressive disorders. N = Total number of participants included in the study. Age: In mean, standard deviation in parenthesis. Age range when mean not stated. Sex: Mix = Both male and female. W = Women.
- **Conditions:** When more than than conditions, the ones used are in italic. BLT = Bright light therapy. CBT = Cognitive behavior therapy. CT = Cognitive therapy. ECT = Electroconvulsive therapy. NARI = Selective noradrenaline reuptake inhibitors. SNRI = Selective serotonin norepinephrine reuptake inhibitors. TCA = Tricyclic antidepressants.
- **Exercise:** HRmax = Maximum heart rate is the highest heart rate an individual can achieve. 1RP = One repetition maximum is the maximum amount of force that can be generated in one maximal contraction. VO2max = Maximal oxygen consumption.
- **Duration:** In weeks when not stated otherwise. M = Months. FU = Follow-up.
- **Main outcome:** BDI = Beck Depression Inventory. CESD = Centre for Epidemiologic Studies Depression Scale. HAMD = Hamilton Rating Scale for Depression. SCL-90 = Symptom Checklist 90.
- **Quality:** Stated in the article that the study had used this methodological characteristic: B = Blinding of assessor of main outcome assessment. A = Allocation concealment. I = Intention-to-treat analysis.

When a study includes more than two conditions, the conditions in italics in the table indicate the conditions used in analysis 1.

* Blumenthal (2012): 47 of 101 participants had MDD. Pinchasov (2000): 18 of 63 participants had nonseasonal MDD.
* Mutrie (1986): Outcome data at 4 weeks are used.
* Schuch (2011): Specific duration of intervention not stated, last outcome data is from ‘discharge’ from hospital.
CI = -0.70 to -0.11, p = .01). The difference in effect sizes between the subgroups was significant (p = .018).

In the second subgroup analysis (Fig. 5), the effect size for the studies that did not use intention-to-treat (11 studies, 334 participants) was large and significant, \( g = -0.80 \) (95% CI = -1.15 to -0.46, \( p < 0.001 \)), and the effect size for those studies that used intention-to-treat (12 studies, 643 participants) was moderate and significant, \( g = -0.56 \) (95% CI = -0.87 to -0.25, \( p < 0.001 \)). The difference in effect sizes between the subgroups was not significant (\( p = .311 \)).

3.3.4. Analysis 2. Exercise versus control: follow-up

Seven studies (348 participants) were included in the analysis of follow-up studies (Fig. 6), and the controlled effect of exercise on depressive symptoms after treatment in follow-up studies was small and nonsignificant, \( g = -0.22 \) (95% CI = -0.53 to 0.09, \( p = .16 \)). Heterogeneity in effect between studies was moderate and not significant (\( Q (6) = 11.311, p = .079, I^2 = 46.96 \)).

3.3.5. Analysis 3. Exercise versus no intervention

Four studies (77 participants) were included in this analysis (Fig. 7). Reduction in depressive symptoms when comparing exercise to no intervention yielded a large and significant effect in favor of exercise, \( g = -1.24 \) (95% CI = -1.83 to -0.65, \( p < 0.001 \)). The heterogeneity was small to moderate and not significant (\( Q (3) = 4480, p = .214, I^2 = 33.04 \)).

3.3.6. Analysis 4. Exercise versus usual care

Four studies (180 participants) were included in this analysis (Fig. 8). Reduction in depressive symptoms when comparing exercise to usual care yielded a moderate and significant effect in favor of exercise, \( g = -0.48 \) (95% CI = -0.80 to -0.16, \( p < 0.001 \)). The heterogeneity was neither substantial nor significant (\( Q (3) = 3582, p = .310, I^2 = 16.25 \)).
3.3.7. Analysis 5. Exercise versus psychological treatment

Three studies (79 participants) were included in this analysis (Fig. 9). Reduction in depressive symptoms when comparing exercise to psychological treatments yielded a small and non-significant effect in favor of exercise, $g = -0.02$ (95% CI = -0.65 to 0.21, $p = 0.31$). There was no substantial heterogeneity ($Q(2) = 1991$, $p = 0.370$, $I^2 = 0.00$).

3.3.8. Analysis 6. Exercise versus antidepressant medication

Based on three studies (236 participants), we found no significant effect on depressive symptoms in favor of exercise when compared to antidepressant medication (Fig. 10), $g = -0.08$ (95% CI = -0.33 to 0.18, $p = 0.55$). There was no substantial heterogeneity ($Q(2) = 1057$, $p = 0.589$, $I^2 = 0.00$).

3.3.9. Analysis 7. Antidepressant medication combined with exercise versus medication only

Four studies (188 participants) were included in this comparison (Fig. 11), and the combination of exercise and antidepressants as a treatment for depression compared to antidepressant medication only yielded a moderate but nonsignificant effect in favor of the combined treatment, $g = -0.50$ (95% CI = -1.10 to 0.11, $p = 0.11$).
### Analysis 2. Exercise versus control: Follow-up

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size</th>
<th>Statistics for each study</th>
<th>Hedge's g and 95% CI</th>
<th>Relative weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exercise</td>
<td>Control</td>
<td>Hedge's g</td>
</tr>
<tr>
<td>Blumenthal 1996</td>
<td>29</td>
<td>-0.09</td>
<td>0.56</td>
<td>0.43</td>
</tr>
<tr>
<td>Blumenthal 2007</td>
<td>43</td>
<td>0.05</td>
<td>-0.45</td>
<td>0.55</td>
</tr>
<tr>
<td>Gary 2010</td>
<td>17</td>
<td>0.02</td>
<td>-0.67</td>
<td>0.71</td>
</tr>
<tr>
<td>Klein 1985</td>
<td>8</td>
<td>-0.57</td>
<td>1.18</td>
<td>0.53</td>
</tr>
<tr>
<td>Krogh 2009</td>
<td>46</td>
<td>0.15</td>
<td>-0.28</td>
<td>0.59</td>
</tr>
<tr>
<td>Sims 2009</td>
<td>23</td>
<td>-0.90</td>
<td>1.51</td>
<td>-0.30</td>
</tr>
<tr>
<td>Singh 1997</td>
<td>15</td>
<td>0.62</td>
<td>-1.33</td>
<td>0.09</td>
</tr>
</tbody>
</table>

**Fig. 6.** Forest plot for analysis 2. Exercise versus control: follow-up.

### Analysis 3. Exercise versus no intervention

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size</th>
<th>Statistics for each study</th>
<th>Hedge's g and 95% CI</th>
<th>Relative weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exercise</td>
<td>Control</td>
<td>Hedge's g</td>
</tr>
<tr>
<td>Doyne 1987</td>
<td>13</td>
<td>-1.19</td>
<td>-2.04</td>
<td>-0.35</td>
</tr>
<tr>
<td>Epstein 1996</td>
<td>7</td>
<td>0.77</td>
<td>-1.72</td>
<td>0.18</td>
</tr>
<tr>
<td>Hermt-Far 2012</td>
<td>10</td>
<td>-0.99</td>
<td>-1.89</td>
<td>-0.10</td>
</tr>
<tr>
<td>Mith 1986</td>
<td>9</td>
<td>7</td>
<td>-2.89</td>
<td>-3.64</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>38</td>
<td>-1.24</td>
<td>-1.83</td>
</tr>
</tbody>
</table>

**Fig. 7.** Forest plot for analysis 3. Exercise versus no intervention.

### Analysis 4. Exercise versus usual care

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size</th>
<th>Statistics for each study</th>
<th>Hedge's g and 95% CI</th>
<th>Relative weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exercise</td>
<td>Control</td>
<td>Hedge's g</td>
</tr>
<tr>
<td>Gary 2010</td>
<td>18</td>
<td>0.17</td>
<td>-0.84</td>
<td>0.50</td>
</tr>
<tr>
<td>Sims 2009</td>
<td>23</td>
<td>-0.53</td>
<td>-1.11</td>
<td>0.06</td>
</tr>
<tr>
<td>Singh 2005</td>
<td>18</td>
<td>-1.00</td>
<td>-1.87</td>
<td>-0.33</td>
</tr>
<tr>
<td>Veale 1992</td>
<td>36</td>
<td>-0.33</td>
<td>-0.81</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>-0.48</td>
<td>-0.80</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

**Fig. 8.** Forest plot for analysis 4. Exercise versus usual care.

### Analysis 5. Exercise versus psychological treatment

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size</th>
<th>Statistics for each study</th>
<th>Hedge's g and 95% CI</th>
<th>Relative weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exercise</td>
<td>Control</td>
<td>Hedge's g</td>
</tr>
<tr>
<td>Epstein 1986</td>
<td>7</td>
<td>-0.81</td>
<td>-1.75</td>
<td>0.17</td>
</tr>
<tr>
<td>Gary 2010</td>
<td>18</td>
<td>0.63</td>
<td>-0.61</td>
<td>0.66</td>
</tr>
<tr>
<td>Klein 1985</td>
<td>14</td>
<td>-0.22</td>
<td>-0.94</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>-0.22</td>
<td>-0.65</td>
<td>0.21</td>
</tr>
</tbody>
</table>

**Fig. 9.** Forest plot for analysis 5. Exercise versus psychological treatment.
The heterogeneity in effect between studies was large and significant \((Q(3) = 10.300, p = .016, I^2 = 70.87)\).

### 3.3.10. Analysis 8. Exercise versus control: blinded outcome

Ten studies (600 participants) were included in the analysis of studies that compared exercise versus control with blinded outcome (Fig. 12). The analysis indicated that exercise had a moderate and significant effect on depressive symptoms compared to the control conditions, \(g = -0.40\) (95% CI = -0.69 to -0.11, \(p = .01\)).

The heterogeneity in effect between studies was moderate to high and significant \((Q(9) = 26.351, p = .002, I^2 = 65.85)\).

### 3.3.11. Analysis 9. Exercise versus control: Allocation concealment, intention-to-treat, blinded outcome

Six studies (461 participants) were included in the analysis with studies using blinded group allocation, blinded outcome, and intention-to-treat analysis (Fig. 13). Reduction in depressive symptoms after treatment indicated a small and nonsignificant
controlled effect from exercise, $g = -0.26$ (95% CI $= -0.61$ to $0.08$, $p = .14$). The heterogeneity in effect between studies was high and significant ($Q(S) = 15.888$, $p = .007$, $I^2 = 68.53$).

4. Discussion

The current meta-analysis is the first to determine the efficacy of exercise for patients suffering from unipolar depression compared to the most common and real alternatives for these patients: psychological treatment, antidepressant medication, usual care, and no intervention. This is important, as a large proportion of these patients do not seek or receive recommended treatment for their depression, or they prefer an alternative treatment to established treatments. It is also the first meta-analysis to assess the effect of exercise in combination treatment with antidepressant medication. Our study also provides a necessary update on the main overall effect of exercise for this group of patients against all control conditions.

The main analysis indicates that exercise as a treatment for unipolar depression has a moderate to large effect ($g = -0.68$) compared to control conditions. Control conditions here range from waiting list to treatment of choice for depression, such as CBT and medication. This makes the interpretation of this finding challenging, as it would be expected that a viable treatment for depression would be superior to waiting list. Further, it seems unreasonable to expect exercise to be superior to treatments that are the recommended treatment options for this condition, as discussed elsewhere (Blumenthal and Doraiswamy, 2014). These results are however in line with previous findings by Krogh et al. (2011) and Silveira et al. (2013). The present meta-analysis included more studies, and thus extends and validates these previous findings. In the subgroup analyses for studies with or without blinded outcome or intention-to-treat, the difference between effect sizes in the first subgroup analysis was significant, while the difference between effect sizes in the second was not significant. This indicates that the effect size for the studies using blinded outcome is indeed different from the effect size for the studies that do not.

Exploration of publication bias in the main analysis supports its results. A fail-safe N of five times the number of included studies in the meta-analysis plus 10 ($5k + 10$) is suggested as a standard for considering the results robust (Rosenthal, 1979). Application of this equation tells us that the results are robust and that the main effect in the present meta-analysis is, with great possibility, a true effect of the experimental intervention and not only a result of bias in publication. The funnel plot was asymmetrical. This does not necessarily imply publication bias, but can indicate small-study effects or true heterogeneity (Sterne et al., 2008). The trim-and-fill method corrects for funnel plot asymmetry: Adding the eight missing studies would decrease the estimated effect size, and the intervention would yield a moderate, albeit smaller, and still significant effect. The method assumes publication bias to be the only reason for asymmetry, and the recomputed effect size should be interpreted with caution (Sterne et al., 2008). Overall, the examination of publication bias establishes the results from the main analysis as valid.

In the comparison between exercise and no intervention, we found a large effect in favor of exercise. This indicates that exercise is an efficacious treatment compared to no treatment. Considering that people with depression often have to wait for psychotherapy, which is the recommended treatment of choice by NICE (2013), after referral to services (Health and Social Care Information Centre and Community and Mental Health statistics team, 2014), this result indicates that exercise can be a worthwhile intervention for those awaiting treatment. The importance of rapid access to treatment is emphasized by the fact that prognosis declines with duration of depression (NICE, 2013; Trivedi et al., 2006). Further, over half of the persons who are suffering from depression globally do not receive any treatment at all (González et al., 2010; Health and Social Care Information Centre and Community and Mental Health statistics team, 2014; WHO, 2012). An alternative treatment like exercise could potentially be offered as a low-threshold option.

We found a moderate and significant effect in favor of exercise when exercise was compared to usual care. As expected, the effect sizes in the analyses with no intervention control or usual care were different, where the effect size for exercise compared to no intervention was larger than the effect size for exercise compared to no intervention ($g = -1.24$ and $-0.48$, respectively). Our finding is comparable to the results of Rethorst et al. (2009) with the same control conditions (however, they pooled no treatment and waiting list together) when they used clinical population as a moderator variable.

When exercise was compared to established psychological treatments (CBT, IPT, and cognitive therapy) or antidepressant medication, the effect sizes were small or close to zero, respectively. The number of studies in this analysis was small, but the current interpretation of these results is that exercise does not function differently from these established treatments for depression.

Analysis of exercise as an adjunct to antidepressant medication indicated a moderate effect in favor of the combination treatment, but this effect was only trending toward significance. This analysis included only four studies and the large confidence interval for this analysis indicates poor precision, so the true effect size can be
higher but also possibly approaching zero. More knowledge about combination treatment for depression is important because we know that many patients need several treatments to achieve satisfactory symptom relief (Casacalenda et al., 2002; Trivedi et al., 2006; Wisniewski et al., 2007). In particular, knowledge about the underlying mechanisms would be of great interest. This could inform how mechanisms for one treatment enhance or moderate mechanisms for the other treatment in the combination. Results from animal studies suggest that different types of antidepressant medication could contribute differently to the joint effect of exercise and medication as treatment for depression (Russo-Neustadt et al., 2001, 2004). One proposed biomarker for the successful treatment of depression is the brain-derived neurotrophic factor (BDNF) (Polyakova et al., 2015). In animal studies, both antidepressant medication and exercise influenced BDNF expression in several hippocampal areas (Russo-Neustadt et al., 2001, 2004). Interestingly, the effect was even more evident for the animals that received combination treatment, as well as in cases where antidepressants alone had failed. Combining different pharmacological drugs increased the risk of adverse effects or drug-drug interaction (Ferguson, 2001; Gillman, 2007), which necessitates further exploration of adjunction with nonpharmacological treatments. The relative efficacy of combining exercise with psychotherapy should also be assessed, but for the time being, we are only aware of three studies (Bonnet, 2005; Gary et al., 2010; Reuter et al., 1984) that have used this combination, of which one (Gary et al., 2010) is included in this meta-analysis.

The effect size was moderate and significant for the analysis including only studies with blinded outcome assessment. When extracting the studies that in addition to blinded outcome had used allocation concealment and intention-to-treat analysis, the effect size decreased to small and was no longer significant. This is equivalent to results in other meta-analyses (Cooney et al., 2013; Krogh et al., 2011). However, due to the small number of studies in this analysis, these findings should be interpreted with caution, as the contribution of other characteristics of these individual studies could be considerable. Nevertheless, this smaller number of studies lends itself to a closer consideration of these studies to understand how the results should be interpreted compared to the overall findings. As discussed above, the efficacy of the control condition is important for the effect of the experimental condition (Borkovec and Sibra, 2005). For instance, when considering the well-conducted study by Blumenthal et al. (1999), it should be kept in mind that exercise is compared to antidepressant medication, which is the treatment of choice for moderate to severe depression (NICE, 2013). At first glance, the results indicate that exercise has no effect. However, since the control group received antidepressant medication, the results imply that exercise had an effect equivalent to that of medication, but that adding exercise to medication did not confer an additional benefit. The two rigorous studies by Krogh and colleagues (Krogh et al., 2012, 2009) also conform to the three quality criteria used in the current and other meta-analyses (Cooney et al., 2013; Krogh et al., 2011). However, the contribution of other important study characteristics is not readily visible in the results of this analysis, such as compliance to the delivered exercise, and this is indeed noted by the authors themselves (Krogh et al., 2012, 2009). Further, overall treatment dose in these studies is well below recommendations (Haskell et al., 2007). Interestingly, in the most recent study by Krogh et al. (2012), there is not a lack of effect by exercise; both exercise and stretching interventions yielded large within-group effects. Findings from several studies suggest that even exercise of low intensity might affect cognitive variables and neuroplasticity (Casilhas et al., 2012, 2007; Ruscheweyh et al., 2011); thus, the apparent lack of a controlled effect of exercise could be explained by the control condition yielding an antidepressant effect as well, which seems to be the case in the study by Krogh et al. (2012). Finally, it should be noted that when the effect of psychotherapy was assessed in a relatively recent meta-analysis (Cuiper et al., 2010), the overall effect size was reduced to small but significant when only high-quality studies were considered. Thus, psychotherapy, an established treatment recommended as treatment of choice, also yields smaller effects when rigorous methods are used, and the authors conclude that this does not necessarily imply that the treatment is not valuable.

The small and nonsignificant effect for the follow-up data is similar to the findings by Cooney et al. (2013) and Krogh et al. (2011). For the time being, it seems like the effects of exercise diminish after the intervention has ceased, but the reason for this is not evident. As stated above, one explanation can be that the treatment the control group receives is as effective as exercise at follow-up. Another explanation is the phenomena of spontaneous remission, where remission occurs without treatment. In support of this, a recent review found that over 50% experience remission within a year without treatment for depression (Whiteford et al., 2012). Interestingly, the authors observed a possible symptom severity association. In accordance with this, all the subjects who had minor depression in the study by Gary et al. (2010) improved over time independent of group assignment. A third explanation can be that the mechanisms which exercise functions through, are short-lived. This is supported by the fact that the increased BDNF level after exercise seems impermanent (Knaepen et al., 2010). More research on mechanisms can possibly identify important features for a longer-lasting effect, or how often subsequent “doses” of exercise are needed to maintain the effect. Two follow-up studies found that self-reported exercise during the follow-up period, but not initial treatment group assignment, predicted depressive symptoms at follow-up (Babyak et al., 2000; Hoffman et al., 2011). If this is the case, then knowledge about which types of exercise interventions cause behavioral change would be particularly valuable. Lasting behavioral change is favorable for both physical and mental well-being (Bahr, 2009), and behavioral change would be beneficial from a public health perspective and cost-benefit perspective.

On the basis of the findings in the current meta-analysis, exercise as a treatment for depression can be recommended as a stand-alone treatment and as an adjunct to antidepressant medication. However, several issues may limit the interpretation and generalizability of the findings. These issues are not specific for the current meta-analysis, but apply to all reviews on the field. First, as mentioned above, the effect appears to subside after exercise cessation. Second, it is impossible to blind the patient to an exercise intervention, and this may introduce bias. To ensure that the trial is at least single-blind, use of a blinded clinical-rated outcome measure is recommended (Krogh et al., 2011). Third, the “file drawer problem”, where nonsignificant or negative findings are less likely to be published than positive findings (Rosenthal, 1979), may have led to an overestimation of the effect of psychological treatments and antidepressant medication for depression (Driesen et al., 2015) and this problem could also apply to exercise as a treatment for depression. One way of minimizing the impact of publication bias in a review is to ensure that all relevant studies are identified. Despite our extensive systematic literature search and requests for unpublished material or ongoing studies from the main authors of the included studies, we may not have succeeded in identifying all studies relevant for this meta-analysis. To augment transparency and access to clinical trials, and to prevent selective reporting and publishing of research outcomes, registration of trials in public registries has been increasingly recommended and common in recent years. Despite this, we found that only 5 of the 23 included studies in the present meta-analysis were pre-registered in a trial registry (Blumenthal et al., 2012b,
2007; Krogh et al., 2012, 2009; Salehi et al., 2014). In the current study, we tried to control for publication bias by use of the statistical analyses of fail-safe N and trim-and-fill analysis. We found that the main effect was somewhat reduced following the trim-and-fill procedure, but it was still moderate and significant. It should however be noted that these approaches have been criticized for their shortcomings (Borenstein, 2005).

Ultimately, the current meta-analysis includes randomized controlled trials (efficacy studies), while practical clinical studies (effectiveness studies) are required to establish the effect and feasibility of exercise for depression in a “real-life” clinical setting (Marley, 2000). Three recent reviews discuss the feasibility of exercise as a treatment for depression (Blumenthal et al., 2012c; Nyström et al., 2015; Rethorst and Trivedi, 2013), and one obvious threat to the effectiveness of exercise as a viable treatment is patient adherence. Exercise may be perceived as more demanding than other available treatments such as antidepressant medication for depressed patients, who lack energy and motivation (Blumenthal et al., 2012c). The credibility of evidence from RCTs is weakened if there is substantial attrition. Statistical techniques, such as intention-to-treat analysis, target the attrition bias from dropouts. However, they ignore other factors that can introduce bias, such as noncompliance, protocol deviations, and withdrawal (Gupta, 2011). A recent systematic overview (Nyström et al., 2015) recommends that exercise should be tailored to the preferences of the patient to raise the likelihood of adherence in a clinical setting. The authors further highlight consideration of subjective experience and additional mechanisms besides increased physical fitness. This is supported by a recent study where context, social support, and social engagement explained the negative association between leisure-time physical activity and depressive symptoms better than biological changes did (Harvey et al., 2010). We recommend that future studies investigate different types of possible mechanisms to provide insight on how exercise best could be tailored to increase its effectiveness as a treatment for unipolar depression.

4.1. Strengths and limitations

The rationale for the inclusion criterion of a diagnosis of unipolar depression was to provide better knowledge about treatment specifically for this common diagnosis. The findings in the main analysis are based on comprehensive and exhaustive available data, and with the separate analyses comparing exercise to no intervention, usual care, and established treatments, and exercise in combination treatment, the present meta-analysis contributes new important knowledge.

We recognize several limitations. Only 7 of the 23 included studies used blinded outcome assessment, allocation concealment, and intention-to-treat analysis. The lack of studies of the highest methodological quality complicates the interpretation of the findings. As expected, the heterogeneity was moderate to high and significant in the majority of the analyses. As noted above, we have used the arms with the largest clinical effect instead of largest dose, and this may overestimate the effect of exercise. Finally, an important issue that we want to highlight is the use of antidepressant medication before enrollment in the studies. Besides the studies that specifically used exercise as an adjunct to medication (which we included in analysis 4), it is not uncommon that studies that do not investigate combination treatment include participants who use antidepressant medication at baseline. In 11 of the included studies (Blumenthal et al., 2012b, 2007, 1999; Doyne et al., 1987; Dunn et al., 2003; Klein et al., 1983; Krogh et al., 2012; Mutrie, 1986; Pinhasov et al., 2000; Singh et al., 1997, 2005), use of antidepressant medication was an exclusion criterion, while in 7 studies (Foley et al., 2008; Gary et al., 2010; Krogh et al., 2009; Martinsen et al., 1985; Salehi et al., 2014; Sims et al., 2009; Veale et al., 1992), some or all of the participants used antidepressant medication. We do not know the impact of the initial use of antidepressants on the overall results in the present meta-analysis, and this can especially confound the results of analysis 3 (comparison of exercise with no intervention), where two of the studies (Epstein, 1986; Hemat-Far et al., 2012) did not mention how they dealt with this issue.

5. Conclusions

Findings from the current meta-analysis indicate that exercise is an effective intervention for depression compared with various types of controls. The effect of exercise as an independent treatment is evident, and the effect is particularly high when compared to no intervention. Thus, exercise may serve as an alternative for patients who do not respond to the given treatment, patients who are awaiting treatment, or those who for different reasons do not receive or want traditional treatment. However, the issue of publication bias and lack of studies of the highest methodological quality complicate the interpretation of the current findings. The present meta-analysis is the first to investigate exercise adjunct to antidepressant medication, and the moderate effect size trending toward significance is interesting and should be further investigated in future studies and meta-analyses.

Potential conflict of interest

All authors declare that they have no conflict of interest.

Acknowledgements

All authors declare that they have no conflict of interest. This meta-analysis was funded by a research grant from the District General Hospital of Førde. The funding source had no influence on the design, execution, or interpretation of the meta-analysis. We are grateful to the following authors for being so helpful and forthcoming to our enquiries: James Blumenthal, Rebecca A. Gary, Jesper Krogh, Jorge Mota-Pereira, Felipe Schuch, Egil W. Martinsen. We thank librarian Kjersti Aksnes-Hopland at the Faculty of Psychology, University of Bergen, for assistance in the literature retrieval.

Appendix A-F. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2016.03.063.

References


