Aerobic exercise for adult patients with major depressive disorder in mental health services. A systematic review and meta-analysis

Morres, Ioannis; Hatzigeorgiadis, Antonis ; Stathi, Afroditi; Comoutos, Nikos ; Arpin-Cribbie, Chantal ; Charalampos Krommidas, Charalampos ; Theodorakis, Yannis

DOI:
10.1002/da.22842

Citation for published version (Harvard):
Morres, I, Hatzigeorgiadis, A, Stathi, A, Comoutos, N, Arpin-Cribbie, C, Charalampos Krommidas, C & Theodorakis, Y 2018, 'Aerobic exercise for adult patients with major depressive disorder in mental health services. A systematic review and meta-analysis', *Depression and Anxiety*, vol. 36, no. 1, pp. 39-53. https://doi.org/10.1002/da.22842

Aerobic exercise for adult patients with major depressive disorder in mental health services. A systematic review and meta-analysis

Short title: Exercise and Major Depression

Ioannis D. Morres[1*], Antonis Hatzigeorgiadis[1], Afroditi Stathi[2], Nikos Comoutos[1], Chantal Arpin-Cribbie[3], Charalampos Krommidas[1], Yannis Theodorakis[1]

[1]School of Physical Education and Sport Science, University of Thessaly, Trikala, Greece

[2]School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK

[3]Department of Psychology, Laurentian University, Ontario, Canada

*Corresponding author:

Ioannis D. Morres, Ph.D.

Postdoctoral Fellow

University of Thessaly, School of Physical Education & Sport Science

Trikala, Karyes, 42100,

Greece

Tel: +30 24310 47009

Email: iomorres@pe.uth.gr

Keywords:

Depression, treatment, exercise, empirically supported treatments, mood disorders

Abstract

Although exercise is associated with depression relief, the effects of aerobic exercise (AE) interventions on clinically depressed adult patients have not been clearly supported. The purpose of this meta-analysis was to examine the antidepressant effects of AE vs. non-exercise comparators exclusively for depressed adults (18-65 years) recruited through mental health services with a referral or clinical diagnosis of major depression. Eleven e-databases and bibliographies of nineteen systematic reviews were searched for relevant randomized controlled trials. A random effects meta-analysis (Hedges' g criterion) was employed for pooling post-intervention scores of depression. Heterogeneity and publication bias were examined. Studies were coded considering characteristics of participants and interventions, outcomes and comparisons made, and study design; accordingly, sensitivity and subgroup analyses were calculated. Across 11 eligible trials (13 comparisons) involving 455 patients, AE was delivered on average for 45minutes, at moderate intensity, three times/week, for 9.2 weeks and showed a significantly large overall antidepressant effect ($g = -0.79$, 95% CI = -1.00, -0.57, $p < .00$) with low and non-statistically significant heterogeneity ($I^2 = 21\%$). No publication bias was found. Sensitivity analyses revealed large or moderate to large antidepressant effects for AE ($I^2 < 33\%$) among trials with lower risk of bias, trials with short-term interventions (up to 4 weeks), and trials involving individual preferences for exercise. Subgroup analyses revealed comparable effects for AE across various settings and delivery formats, and in both outpatients and inpatients regardless symptom severity. Notwithstanding the small number of trials reviewed, AE emerged as an effective antidepressant intervention.

Introduction

Depression is a life threatening and disabling mental illness affecting increasingly large proportions of the society at an alarming rate worldwide (Global Burden of Disease Study 2013 Collaborators, 2015; Üstün, Ayuso-Mateos, Chatterji, Mathers, & Murray, 2004). Major depressive disorder (also referred to as clinical depression) is the most common type of depression seriously challenging health systems especially since it is often recurrent and treatment resistant.

Physical exercise is widely recommended in depression treatment (NICE, 2009; CANMAT, 2016; Stanton & Reaburn, 2014). It has been associated with depression relief in various meta-analytic reviews (Craft & Landers, 1998; Rethorst, Wipfli, & Landers, 2009; Robertson, Robertson, Jepson, & Maxwell, 2012; Schuch et al., 2016; Silveira et al., 2013; Stanton & Reaburn, 2014), even after risk of bias was considered (e.g., Rethorst et al., 2009; Schuch et al., 2016). In contrast, other meta-analyses set this association into question after coding for lower risk of bias (Cooney et al., 2013; Krogh, Hjorthøj, Speyer, Gluud, & Nordentoft, 2017; Krogh, Nordentoft, Sterne, & Lawlor, 2011).

Considering carefully the attributes of these meta-analytic studies, however, we identified methodological aspects that could potentially justify these equivocal conclusions. Particularly, in a number of trials reported in these meta-analyses the samples included participants who were recruited through mental health services but also through media advertisements. Also, in a number of trials the diagnosis of depression was not based on valid diagnostic criteria. Three meta-analytic studies, Krogh et al. (2011), Krogh et al. (2017) and Kvam, Kleppe, Nordhus, and Hovland (2016) have exclusively focused on exercise trials for patients with a diagnosis of major depression based on valid diagnostic criteria including the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) or the International Classification of Disease (World Health Organization, 1992).

However, some of the trials in these three studies included media respondents. Media respondents may have strong outcome expectations or motivation to lifestyle changes; Blumenthal and Ong (2009) reported that community volunteers of exercise trials for depression are typically motivated to exercise. Instead, depressed patients recruited through health services do not seem to be comparably motivated, as they show high dropout rates from exercise on referral schemes (Crone, Johnston, Gidlow, Henley, & James, 2008; James et al., 2008; Tobi, Kemp, & Schmidt, 2017). In addition, a number of trials reviewed by these three meta-analyses included samples with other mood disorders (e.g., dysthymia), or samples with older adult (+65 years) or both adult (18-65 years) and older adult depressed patients. Older depressed adults show distinct differences in depression (Fiske, Wetherell, & Gatz, 2009), demonstrate higher depression relief through exercise (Silveira et al., 2013), and are more likely to complete exercise on prescription programs than adult peers (James et al., 2008). Finally, these three meta-analyses included trials comparing exercise to other exercise activities (e.g., stretching). Such exercise activities, however, have been also shown to improve depression, thus confounding the true antidepressant effect of exercise intervention and misleading relevant conclusions (Schuch, Morres, Ekkekakis, Rosenbaum, & Stubbs, 2017).

In addition to the issues identified above, the quality of the relevant meta-analysis could improve through a more appropriate assessment for the risk of bias. While exercise reviews for depression are typically coding trials for risk of bias, attributable to the association of systematic errors with overestimation of treatment efficacy (Moher et al., 1998; Schulz, 2001; Schulz, Chalmers, Hayes, & Altman, 1995), this coding has not been based on tools designed for physical therapy interventions, such as physical exercise. This seems essential as intervention-specific design aspects including dropouts may remain undetected and continue to influence conclusions, especially because exercise is perceived as requiring higher time- and effort-demands than other health behaviors (Turk, Rudy, & Salovey, 1984). Relevant coding has been recently addressed by meta-analytic studies in the exercise-anxiety area (Ensari, Greenlee, Motl, & Petruzzello, 2015; Stonerock, Hoffman, Smith, & Blumenthal, 2015). Finally, we identified in the literature that the antidepressant effects of aerobic exercise (AE) in particular remain unexplored by meta-analyses, although AE (e.g., walking) is the most pursued type of physical activity in mental health services (Sørensen, 2006).

Taking into account all the issues identified above, the purpose of our study was to explore the antidepressant effects of AE intervention, when compared to non-exercise comparators in adult patients, 18-65 years of age, recruited through mental health services with a referral or a clinical diagnosis of major depressive disorder. Towards this goal, coding for risk of bias was addressed based on a tool structured for physical therapy interventions.

Method

This review was conducted in accordance to the Preferred Items for Systematic Reviews and Meta-Analyses (PRISMA) statement that ensures quality through a standard list of 27 items (Moher, Liberati, Tetzlaff, & Altman, 2009).

Literature search

Eligibility criteria of study characteristics included; (1) participants 18-65 years old, recruited via mental health services with a referral or with a previous diagnosis of major depression (without psychotic features) as a primary disorder and not as a result of a mental or medical disorder/condition; (2) AE interventions as defined by the American College of Sports Medicine (ACSM) (Pollock et al., 1998); (3) comparison of AE interventions to any treatment form (e.g., psychotherapy, medication) or condition (e.g., waiting list) excluding exercise activities; (4) outcome measures of depression as a primary outcome; (5) studies with a design of a randomized controlled trial (RCT). Given, however, the limited number of exercise trials for clinically depressed adults seen in literature, it was an a priori decision of this meta-analysis to (i) include comparative efficacy trials and (ii) compare AE to a second intervention (excluding exercise activities) should three-arm RCTs were allocated. Previous meta-analyses for exercise on depression are seen to typically include comparative efficacy trials (e.g., Krogh et al., 2017; Kvam et al., 2016). Finally, the search covered the period from 1980 to March 2017 for trials written in English.

Two authors independently conducted the literature search. A total of ten electronic databases were searched: SCOPUS, PubMed, PEDro, and the PsyINFO, SPORTDiscus, Academic Search Complete, Education Resource Information Centre via EBSCO. Also, the Proquest Dissertations and Theses database was searched for unpublished studies. In addition, we searched the Trials Register of Promoting Health Interventions (TRoPHI; EPPI Centre), ClinicalTrials.Gov, and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP). We used medical subject headings (e.g., MESH) when possible or text word terms including: depressive disorder, depression, diagnosis, patients, adult (limiters 18-65years), and exercise, aerobic, running, swimming, jogging, walking or bicycling. The full search strategy for the PubMed is given in supplementary material 1.

Trials were initially screened through titles and abstracts. Full text versions were obtained subject to positive initial screening. Hand searching was also conducted; we screened the bibliographies of all major systematic reviews including Cochrane reviews in the exercise-depression area in the last 20 years. This led to the revision of 19 systematic reviews (supplementary material 2).

Data extraction

Data were extracted and checked for accuracy and for duplicates onto prepared forms by the first and the fourth author. Subsequently, the two authors reached a consensus on the eligible trials. Researchers of eligible trials were contacted by e-mail to provide additional information on clinical/methodological aspects to their studies. Two reminders were sent within a period of two months. All but one author provided clinical clarifications involving delivery formats of exercise interventions (supervision, location, individually or group) and status of participants (outpatients or inpatients). Also, three authors were contacted to provide methodological clarifications. Relevant information was obtained and included number of dropouts, concealed allocation, blind assessment or baseline balance.

Coding

Coding of eligible trials was based on the PICOS criteria that refer to Participants, Interventions, Comparisons, Outcome measures and Study design (PICOS) (Moher et al., 2009). Hence, coding included the following study characteristics: (1) participants: hospitalized or non-hospitalized (inpatients or outpatients); severity of depression (mild, moderate or severe); (2) intervention: duration ($\leq$ 4weeks), frequency (< 3 days/week, 3 days/week or > 3 days/week), intensity (lower, moderate or vigorous), setting (outdoors or indoors), social format (in groups or individually), delivery format (supervised or non-supervised); (3) comparisons: AE vs. non-exercise comparators; (4) outcome measures: rating (self- or clinician-rated); (5) study design: RCTs or comparative efficacy trials with higher or lower risk of bias scores of methodological qualities (score of < 6 or $\geq$ 6) on the Physiotherapy Evidence Database scale (PEDro) (de Morton, 2009).

Risk of bias

The PEDro scale is a well-established comprehensive measure of methodological quality in the literature of physical therapy (Bhogal, Teasell, Foley, & Speechley, 2005), which has shown good psychometric properties (de Morton, 2009; Macedo et al., 2010; Maher, Sherrington, Herbert, Moseley, & Elkins, 2003). Also, the PEDro scale is increasingly used in various research areas (e.g., Knols, de Bruin, Shirato, Uebelhart, & Aaronson, 2010; Pan, Wang, Xie, Du, & Guo, 2014; Pinto et al., 2012), including the area of exercise and anxiety (e.g., Ensari et al., 2015). In the area of exercise for depression, two systematic reviews have employed the PEDro scale (Perraton, Kumar, & Machotka, 2010; Stanton & Reaburn, 2014), but neither was a meta-analysis. Internal validity on the PEDro scale is assessed on a point system favoring between-groups comparisons and point estimates/variability measures; blinding patients, therapists and assessors; baseline balance, intention-to-treat, and drop-outs. The maximum score for the internal validity criteria is 10; in our study the maximum score was 8, as it is difficult, if not impossible, to blind patients/therapists in exercise trials for depression. A cut-off score of 6 was employed for classification of higher quality RCTs (lower risk of bias). This score represents the point of reference for high quality trials when using the PEDro scale (Maher et al., 2003). The first and the third authors independently assessed the methodological quality of each trial, and sought consensus on the relevant evaluations. Cohen's Kappa statistic was computed, and interpreted based on the Landis and Koch reference to estimating the inter-rater agreement (Landis & Koch, 1977).

Statistical analysis

The software Comprehensive Meta-Analysis (Version 2.0, Biostat, Englewood, New Jersey) was used to calculate intervention effects. A random effects model using the Hedges' g criterion was employed to estimate standardized mean differences in depression scores from pre- to post-intervention between exercise and non-exercise comparators (Borenstein, Hedges, Higgins, & Rothstein, 2010). The selection of a random effects model lies upon the assumption that there is a sampling error (within-study error) and between-study variance. Also, the Hedges' g criterion was selected because it prevents overestimation of an effect-size when the retrieved studies are less than 20. Hedge's g algorithms were interpreted with the Cohen's d standards (Cohen, 1992) where values of .20, .50 and .80 point a small, moderate and large intervention effect, respectively.

Statistical heterogeneity was assessed with the Cochran's Q and I^2 statistics for each trial (Higgins, Thompson, Deeks, & Altman, 2003) taking into account that I^2 values up to 40% are unlikely to be important (Higgins & Green, 2011). Publication bias was assessed by means of visual inspection of the funnel plots and the Begg-Mazumbar Kendall's tau (Begg & Mazumdar, 1994) and Egger bias test (Egger, Smith, Schneider, & Minder, 1997) for the main composite analysis. In case of significant publication bias, the trim and fill statistical procedure was considered on the right and left side of the plot (Duval & Tweedie, 2000). This procedure adds or removes studies to balance an asymmetrical funnel plot and adjusts the effect-size accordingly. In this manner, an unbiased estimate of the effect is provided. Further, the fail and safe criterion (Rosenthal, 1979) was employed to calculate the number of studies needed to nullify significant effects (e.g., >.05) for the main composite analysis. A fail-safe number of five times the number of reviewed comparisons plus 10 (5K+10) is seen as the cutoff score for considering the results robust (Rosenthal, 1979).

Also, three sensitivity analyses were performed to explore the robustness of our main finding (overall effect-size). First, we coded lower risk of bias trials (PEDro score ≥ 6) because meta-analytic reviews have reported equivocal conclusions on the association of exercise with antidepressant effects in high quality trials (e.g., Cooney et al., 2013; Krogh et al., 2017; Krogh et al., 2011; Rethorst et al., 2009; Schuch et al., 2016). In the second and third sensitivity analyses we coded trials with short-term (up to four weeks) or preference-oriented exercise interventions, respectively, because both interventional aspects comprised a promising strategy in the only currently available pragmatic RCT for exercise in adult depressed women who were living in the community and recruited though health services (Callaghan, Khalil, Morres, & Carter, 2011); both analyses would potentially contribute to the clarification of the translational value of our main finding for pragmatic settings (routine practice).

In addition, eight subgroup analyses were computed based on the PICOS criteria in order to explore in detail the effect of AE on depression across various delivery formats, comparisons and settings, in both out- and in-patients with various depressive symptom severities, and outcome measures used. Differences in effects-sizes for subgroup comparisons were considered significant if non-overlapping 95% confidence intervals for the different subgroup was found (Higgins & Green, 2011). Subgroup analyses involved consideration that less than five comparisons as estimates of effect may reveal imprecise results (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Publication bias tests including the Begg-Mazumbar Kendall's tau and Egger bias test were also performed, taking into account, however, that ten or more comparisons are needed to provide more accurate results (Sterne et al., 2011). Also, risk difference (RD) investigated potential differences in the number of dropouts between the AE and non-exercise comparators. Finally, a p-value of 0.05 was considered for significance in all computations.

Results

The PRISMA flow diagram for study selection is presented in Figure 1. Eleven trials involving 455 patients recruited via mental health services were eligible for inclusion. These studies are presented in Table 1 and 2 and are marked (*) in the reference list. One study (Veale et al., 1992) reported two trials, but only the first compared AE to a non-exercise condition and was allocated as Veale et al. (1992-a). Also, Salehi et al. (2016) was a three-arm comparative efficacy trial that employed AE, electroconvulsive therapy (ECT), and AE + ECT. In line to our a priori decision to review comparative efficacy trials and exclude trials with exercise controls, we reviewed only the comparisons of AE vs. ECT and AE + ECT vs. ECT and allocated as Salehi et al. (2016-a) and Salehi et al. (2016-b), respectively. Also, in the three-arm trial of Sadeghi et al. (2016), AE was compared to the non-exercise comparator of "group discussions". This trial, however, compared also cognitive therapy as a second intervention to the same non-exercise comparator. Due to small number of exercise trials for depression, we compared AE to “group discussion” and allocated it as Sadeghi et al. (2016-a), whereas the comparison of AE to cognitive therapy was allocated as Sadeghi et al. (2016-b). The eleven eligible trials yield 13 comparisons and used continuous outcomes at pre/post-intervention with higher scores indicating more severe depression. Based on available evidence from six of the reviewed trials (see Table 1), an average of 65% of eligible patients entered into studies. Finally, trials were implemented in North and South America, Europe and Asia (Table 1).

Study characteristics

Only one eligible trial did not employ fully supervised-based AE (Mota-Pereira et al., 2011); it included one supervised session (in a hospital gymnasium) and four non-supervised home-based sessions per week. Also, AE was delivered on average three times/week, at moderate intensity, with a session length of 45minutes, and total program duration of 9.2 weeks. A comparable number of trials recruited outpatients or inpatients, used equipment-based or -free modalities, administrated clinician- or self-rated outcomes, employed group or individual formats, and were delivered inside or outside a hospital. Also, a comparable number of trials compared AE to treatment as usual (TAU), antidepressant medication or psychological therapies. In two trials, AE was compared to non-exercise comparators of waiting list (Oertel-Knöchel et al., 2014) and self-administrative group discussion (Sadeghi et al., 2016). Seven trials were conducted indoors, and seven recruited patients with moderate-severe or severe depression. The average dropout rate for the intervention and non-exercise comparators was 14.8% and 14.5%, respectively.

Risk of bias

Seven trials (64%) received a score of ≥ 6 on the PEDro scale that indicates lower risk of bias. In contrast, four trials showed higher risk of bias, as they received a score of < 6 on the PEDro scale (see Table 2). Cohen's Kappa statistic was .77, portraying a substantial inter-rater reliability on the PEDro scoring (Landis & Koch, 1977).

Meta-analysis

Pooled results showed that AE revealed a significantly large overall antidepressant effect compared to non-exercise comparators ($g = -0.79$, 95% CI -1.00, -0.57, $p < .00$) with low and non-significant heterogeneity ($Q = 15.20$, $p = 0.23$, $I^2 = 21\%$). The Begg-Mazundar Kendall's tau (tau -0.21, $p = 0.29$) and Egger bias test (intercept -2.48, $p = 0.12$) indicated no publication bias (see Table 3 and Figure 2). Also, the funnel plot did not show evidence of asymmetry (Figure 3). In addition, the fail-safe algorithm indicated that 218 studies with no antidepressant effect for AE would be required to nullify the significance of the main result. This indicates no publication bias, given that the relevant fail-safe standard (5k+10) for this review is 75 studies (5 X 13 comparisons +10). Therefore, computation of trim and fill analysis did not appear to be essential.

Sensitivity analyses

The sensitivity analysis for lower risk of bias trials displayed a moderate to large effect for AE on depression ($g = -0.70$, 95% CI -0.94, -0.45, $p < .00$) with negligible heterogeneity ($I^2 = 2\%$). The sensitivity analysis for trials involving individual exercise preferences displayed a large effect for AE ($g = -0.84$, 95% CI -1.17, -0.51, $p < .00$) with very low and non-significant heterogeneity ($I^2 = 7\%$). The sensitivity analysis for trials with short-term interventions displayed a moderate to large effect for AE ($g = -0.71$, 95% CI -1.09, -0.34, $p < .00$) with low and non-significant heterogeneity ($I^2 = 30\%$). Details are presented in Table 3.

Subgroup Analyses

Eight subgroup analyses were conducted on the basis of the PICOS criteria (Moher et al., 2009) excluding the study design criterion (S) that was coded by our first sensitivity analysis. All the analyses revealed statistically significant effects for AE. Furthermore, none of the analyses provided evidence for significant differences for effect-sizes between the different groups; no publication bias was recorded; and levels of heterogeneity were non-significant. The analyses are described below and the full statistics are presented in Table 4.

Participants

Aerobic exercise showed large and moderate to large effects for outpatients and inpatients, and for patients with any severity classification in depression (g range = -.71 to -.97). In these samples heterogeneity was low and non-significant (I^2 from 2% to 41%).

Intervention

Large effects were found in trials where AE included equipment-free modalities, and was conducted in individual or group formats, and in indoor, outdoor or non-hospital settings (g range = -0.77 to -1.07). In addition, moderate to large or moderate effects were seen in equipment-based AE (g = -0.67, 95% CI -.98, -0.35, $p < .00$) or in hospital settings (g = -0.61, 95% CI -0.96, -0.27, $p < .00$). In all computations, levels of heterogeneity were non-significant and ranged from 0% to 53%. Finally, no dropout differences were found between the aerobic and the non-exercise comparators (RD = 0.01, 95% CI -0.03, 0.05, p = 0.59, I^2 = 0%).

Comparisons

Compared to antidepressants/TAU, AE showed large antidepressant effect (g = -0.75, 95% CI -1.01, -0.48, $p < .00$). Similar effects were seen in the comparison of AE to psychological treatments when these treatments were performed as mono-therapy/part of multi-therapeutic program (g = -0.85, 95% CI -1.21, -0.48, $p < .00$). Heterogeneity was low (I^2 0% and 42%, respectively) and non-significant.

Outcome measures

Large and moderate to large, respectively, antidepressant effects for AE were found in trials with self-rated ($g = -0.97$, 95% CI -1.35, -0.59, $p < .00$) or clinician-rated ($g = -0.69$, 95% CI -0.94, -0.44, $p < .00$) outcome measures. Heterogeneity was low (I^2 26% and 9%, respectively) and non-significant. Also, a reduction of depression score by 4.5 points was seen in trials employing the clinician-rated outcome of the Hamilton Rating Scale for Depression (HAMD) (-4.48, 95% CI = -7.69, -1.25, standard error = 1.64, $p < .00$), and by 7 points in trials employing the self-rated outcome of the Beck Depression Inventory-I or II (-6.96, 95% CI = -12.49, -1.42, standard error = 2.83, $p = 0.01$).

Discussion

In this review, AE showed a significant large overall antidepressant effect ($g = -0.79$) on adult patients recruited via mental health services with a referral or a clinical diagnosis of major depression. Heterogeneity was low and non-significant ($I^2 = 21\%$), and no signs of publication bias were found.

Our study included only trials with depressed patients recruited via mental health services; this could be considered a step forward in the literature. Previous meta-analyses included a number of trials with depressed persons recruited via media advertisements. However, media respondents may have a non-clinical depression, despite a high score in depression checklists, or a possible diagnosis given before/at study entry. Moreover, they may have strong outcome expectations and determination for lifestyle change; Blumenthal and Ong (2009) report that community volunteers for exercise trials for depression are typically motivated to exercise. Depressed patients instead, may have a more challenging clinical profile given that they have suffered tenacious symptoms, including psychosocial impairment that led to a mental health service, and they may often experience failure or disappointment because the service use uncovers the disease severity/complexity and the need for systematic care (Bursztajn & Barsky, 1985; Maguire, Cullen, O'Sullivan, & O'Grady-Walshe, 1995;

Morgan, 1989). To this extent, depressed patients referred to exercise on referral schemes are, unsurprisingly, showing high dropout rates, often the highest among all health referrals (Crone et al., 2008; James et al., 2008; Tobi et al., 2017). Therefore, our findings that stem from trials with patients with a referral or clinical diagnosis of depression are representative to routine practice and, thus, of additional value.

Another positive aspect of our study dealt with the inclusion of trials with only adult depressed patients of 18-65 years old. This separation appeared to be essential because older adult (+65years) depressed patients manifest distinct clinical differences in depression (Fiske et al., 2009) and higher depression relief through exercise (Silveira et al., 2013), and are more likely to complete exercise on prescription programs (James et al., 2008).

Importantly, 64% of our trials mirrored lower risk of bias (PEDro score ≥ 6). To the best of our knowledge, only two systematic (but not meta-analytic) reviews for exercise and depression have previously employed the PEDro scale (Perraton et al., 2010; Stanton & Reaburn, 2014). These studies reported a score of ≥ 6 on the PEDro scale for 43% (Perraton et al., 2010) and 100% (Stanton & Reaburn, 2014) of their reviewed trials. Exercise and anxiety meta-analyses (Ensari et al., 2015; Stonerock et al., 2015) that employed the PEDro scale have reported substantially fewer trials with lower risk of bias (33%) compared to our review.

Also, three sensitivity analyses were performed. The first recorded a significant moderate to large antidepressant effect for AE among lower risk of bias trials. This is in line to Rethorst et al. (2009) or Schuch et al. (2016) who found similar effects for exercise on depression after coding for lower risk of bias, but in contrast to reviews that reported no antidepressant effect for exercise after relevant coding (Krogh et al., 2017; Krogh et al., 2011). The present finding is of key-importance given that high risk of bias is linked with overestimation of treatment efficacy (Moher et al., 1998; Schulz, 2001; Schulz et al., 1995). Further, the use of the PEDro scale, which was developed to evaluate risk of bias for physical therapy interventions such as physical exercise, suggests that when intervention-specific design aspects (e.g., dropouts) are included in the relevant evaluation do not confound results.

In the second sensitivity analysis, short-term AE (up to 4 weeks) showed a moderate to large effect on depression highlighting its vital role at the early stage of care, as the most frequent treatment of pharmacotherapy is typically requiring a period of 4 weeks before providing any benefit (Gartlehner et al., 2011; John M. Eisenberg Center for Clinical Decisions Communications Science, 2011; Seehusen & Sheridan, 2013). In further support of its vital role at the early stage of mental health care, AE has been also shown to improve well-being of major depressed patients after a single session (Bartholomew, Morrison, & Ciccolo, 2005), and to decrease depression after ten consecutive daily sessions in a sample comprising both bipolar and major depressed patients (Knubben et al., 2007). In the third sensitivity analysis AE involving individual preferences revealed large antidepressant effects. Preference-based exercise appears to be a promising strategy in adolescent and adult general population segments (e.g., Hamlyn-Williams, Freeman, & Parfitt, 2014; Rose & Parfitt, 2007), in patients with various health disorders (Morton, Biddle, & Beauchamp, 2008) and in depressed patients participating in pragmatic RCTs (Callaghan et al., 2011; Carter et al., 2015; Morres, Stathi, Martinsen, & Sørensen, 2014). Therefore, this strategy warrants further attention by researchers and practitioners.

The findings of the second and third sensitivity analyses (up to four-week exercise and involvement of individual preferences) provide potentially important translational evidence for routine practice as they concur with the only currently available AE trial with a pragmatic design for adult depressed women (≤ 65years) recruited through health services (Callaghan et al., 2011). This RCT reported antidepressant effects for a four-week and preference-based AE program implemented in public gyms for depressed women living in the community (Callaghan et al., 2011).

Based on the subgroup analyses, AE brought about a large or moderate to large improvement in depression in a wide range of delivery formats; through equipment-based or equipment-free modalities, inside or outside a hospital, outdoors or indoors, in groups or individually, and in cohorts with outpatients or inpatients, and with different depressive symptom severity. In the two remaining subgroup analyses, AE favored over psychological treatments or antidepressants/TAU, demonstrating large effects on depression. Overall, AE was found to be comparably effective across all eight subgroup analyses. Noteworthy, three of our reviewed trials consisted of treatment-resistant depressed samples (Mota-Pereira et al., 2011; Oertel-Knöchel et al., 2014; Pilu et al., 2007). Another notable finding was that supervised AE had clinically meaningful antidepressant effects; in trials using the clinician-rated HAMD or the self-rated BDI outcomes, we found a reduction in depression score by 4.5 and 7 points, respectively. Both reductions are considered clinically meaningful, as the relevant cutoff score is 3 points (NICE, 2009). Finally, AE showed similar dropout rates to non-exercise comparators. Given that exercise is perceived as requiring higher time- and effort-demands than other health behaviors (Turk et al., 1984), the lack of differences in dropout rates between AE and non-exercise treatment modalities for depression is of major importance.

Several limitations of the study require consideration. First, a small number of trials were allocated. However, this is widely seen as a limitation of the field. Also, we tried to run a robust analysis to offset threats caused by publication bias. In addition, most of our sensitivity and subgroup analyses (89%) computed at least five arms, which is considered an essential number to avoid imprecise results (Borenstein et al., 2009). Second, we did not code for side-effects of exercise because only Legrand and Neff (2016) explored side-effects, with 37% of the sample reporting transient joint/muscular soreness, diarrhea or fatigue. Although these effects did not preclude patients from completing the interventions (Legrand & Neff, 2016), exploration of side-effects of exercise via standardized tests is an important priority. Based on limited data, almost 12% of mental health patients experience a side-effect by exercise (side-effects and patients' diagnoses were not clarified) (Sørensen, 2006). Also, only 6% of depressed patients (including older adults) assigned to physical activity programs describe the potential of a side-effect (Searle et al., 2011).

Third, our reviewed trials did not fully describe design criteria or clinical characteristics. Although all but one author provided comprehensive feedback to our inquiries, future trials need to detail intervention and non-exercise comparators. Of particular importance, the contact time of exercisers with supervisors and/or peers may have a confounding impact on the favourable comparison of exercise to non-exercise comparators. Such confounding bias, however, did not seem to occur in our study; in our subgroup analyses AE was more effective than non-exercise comparators of either increased or decreased contact time with service providers/users (e.g., psychological treatments or antidepressant medication). Nevertheless, future trials should equalize groups in contact time and present relevant information in order to allow researchers draw even firmer conclusions on whether AE per se favours non-exercise comparators in depressed patients.

A final issue that needs to be discussed concerns the representativeness of our findings to routine practice. Specifically, 35% of the patients identified as eligible for participation by our reviewed trials appeared to be reluctant to entering into studies. However, considering the pessimistic profile of depression and the high time and effort demands of exercise, this percentage appears rather low. Nevertheless, barriers to entering into studies with AE programs need to be identified by future trials. Only two of our reviewed trials provided such information, reporting lack of interest as a predominant barrier (Legrand & Neff, 2016; Schuch et al., 2015). Since lack of interest is a key symptom of depression, recruiters may find it difficult to trigger interest to exercise these patients. However, a study conducted in routine practice for mental health patients from both hospital and day-care centers has provided promising evidence (Sørensen, 2006). Particularly, exercise participation was predicted solely by intrinsic regulation, reflecting inherent pleasure and participation for its own shake (Sørensen, 2006). To this extent, autonomy supportive environments, providing options for the type and intensity of exercise, which promote intrinsically regulated exercise behavior (Ryan & Deci, 2000), might trigger interest in eligible but reluctant to exercise depressed patients.

Conclusion

This is the first meta-analytic study to compare the antidepressant effects of AE to treatments for depression excluding exercise activities, in adult patients (18-65) with a referral or clinical diagnosis of major depression, who were recruited through mental health services and not through media advertisements. Supervised AE compared favorably to treatments for depression across various delivery formats, comparisons or settings, and regardless symptom severity and type of outcome measure. Importantly, the antidepressant effect of AE was not affected among trials with lower risk of bias, trials with short-term (up to four weeks) exercise interventions, or trials with interventions involving individual preferences for exercise. Notwithstanding the limited number of trials reviewed, AE was found to be an effective antidepressant intervention.

Acknowledgments:

We would like to thank the following researches for providing additional information on their trials: Brand, S., Carta, M.G., Kahl, K.G., Legrand, F., Martinsen, E.W., Mota-Pereira, J., Oertel-Knöchel, V., Rueter, M., Schuch, F.B., Veale, D.

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). *Washington, DC: Author.*

Bartholomew, J. B., Morrison, D., & Ciccolo, J. T. (2005). Effects of Acute Exercise on Mood and Well-Being in Patients with Major Depressive Disorder. *Medicine and Science in Sports and Exercise, 37*(12), 2032-2037. doi:10.1249/01.mss.0000178101.78322.dd

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 1088-1101. doi:10.2307/2533446

Bhogal, S. K., Teasell, R. W., Foley, N. C., & Speechley, M. R. (2005). The PEDro scale provides a more comprehensive measure of methodological quality than the Jadad scale in stroke rehabilitation literature. *Journal of Clinical Epidemiology, 58*(7), 668-673. doi:10.1016/j.jclinepi.2005.01.002

Blumenthal, J. A., & Ong, L. (2009). A commentary on 'Exercise and Depression' (Mead et al., 2009): And the verdict is…. *Mental Health and Physical Activity, 2*(2), 97-99. doi:10.1016/j.mhpa.2009.08.001

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Random-Effects Model *Introduction to Meta-Analysis* (pp. 69-75): John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97-111. doi:10.1002/jrsm.12

Bursztajn, H., & Barsky, A. J. (1985). Facilitating patient acceptance of a psychiatric referral. *Archives of Internal Medicine, 145*(1), 73-75. doi:10.1001/archinte.1985.00360010097015

Callaghan, P., Khalil, E., Morres, I., & Carter, T. (2011). Pragmatic randomised controlled trial of preferred intensity exercise in women living with depression. *BMC Public Health, 11*(1), 465. doi:10.1186/1471-2458-11-465

Carter, T., Guo, B., Turner, D., Morres, I., Khalil, E., Brighton, E., . . . Callaghan, P. (2015). Preferred intensity exercise for adolescents receiving treatment for depression: a pragmatic randomised controlled trial. *BMC Psychiatry, 15*(1), 1-12. doi:10.1186/s12888-015-0638-z

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science, 1*(3), 98-101. doi:10.1111/1467-8721.ep10768783

Cooney, G., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., . . . Mead, G. E. (2013). Exercise for depression. *Cochrane Database Syst Rev, 9*, CD004366. doi:10.1002/14651858.CD004366.pub6

Craft, L. L., & Landers, D. M. (1998). The effect of exercise on clinical depression and depression resulting from mental illness: a meta-analysis. *Journal of Sport & Exercise Psychology, 20*, 339-357. doi:10.1123/jsep.20.4.339

Crone, D., Johnston, L. H., Gidlow, C., Henley, C., & James, D. V. B. (2008). Uptake and Participation in Physical Activity Referral Schemes in the UK: An Investigation of Patients Referred with Mental Health Problems. *Issues in Mental Health Nursing, 29*(10), 1088-1097. doi:10.1080/01612840802319837

de Morton, N. A. (2009). The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Australian Journal of Physiotherapy, 55*(2), 129-133. doi:10.1016/S0004-9514(09)70043-1

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics, 56*(2), 455-463. doi:10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634. doi:10.1136/bmj.315.7109.629

Ensari, I., Greenlee, T. A., Motl, R. W., & Petruzzello, S. J. (2015). Meta-analysis of acute exercise effects on state anxiety: An update of randomized controlled trials over the past 25 years. *Depression and Anxiety, 32*(8), 624-634. doi:10.1002/da.22370

Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in Older Adults. *Annual Review of Clinical Psychology, 5*, 363-389. doi:10.1146/annurev.clinpsy.032408.153621

Gartlehner, G., Hansen, R. A., Morgan, L. C., Thaler, K., Lux, L., Van Noord, M., . . . Wilkins, T. (2011). Comparative Benefits and Harms of Second-Generation Antidepressants for Treating Major Depressive DisorderAn Updated Meta-analysis. *Annals of Internal Medicine, 155*(11), 772-785. doi:10.7326/0003-4819-155-11-201112060-00009

Global Burden of Disease Study 2013 Collaborators. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The lancet, 386*(9995), 743-800. doi:10.1016/S0140-6736(15)60692-4

Hamlyn-Williams, C. C., Freeman, P., & Parfitt, G. (2014). Acute affective responses to prescribed and self-selected exercise sessions in adolescent girls: an observational study. *BMC Sports Science, Medicine and Rehabilitation, 6*(1), 1-9. doi:10.1186/2052-1847-6-35

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions* (Vol. 4): John Wiley & Sons.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560. doi:10.1136/bmj.327.7414.557

James, D. V., Johnston, L. H., Crone, D., Sidford, A. H., Gidlow, C., Morris, C., & Foster, C. (2008). Factors associated with physical activity referral uptake and participation. *Journal of Sports Sciences, 26*(2), 217-224. doi:10.1080/02640410701468863

John M. Eisenberg Center for Clinical Decisions Communications Science. (2011). Second-Generation Antidepressants for Treating Adult Depression: An Update *Comparative Effectiveness Review Summary Guides for Clinicians. Rockville (MD): Agency for Healthcare Research and Quality (US) 2007*.

Kerling, A., Tegtbur, U., Gützlaff, E., Kück, M., Borchert, L., Ates, Z., . . . Kahl, K. G. (2015). Effects of adjunctive exercise on physiological and psychological parameters in depression: A randomized pilot trial. *Journal of Affective Disorders, 177*, 1-6. doi:10.1016/j.jad.2015.01.006

Knols, R. H., de Bruin, E. D., Shirato, K., Uebelhart, D., & Aaronson, N. K. (2010). Physical activity interventions to improve daily walking activity in cancer survivors. *BMC Cancer, 10*, 406. doi:10.1186/1471-2407-10-406

Knubben, K., Reischies, F. M., Adli, M., Schlattmann, P., Bauer, M., & Dimeo, F. (2007). A randomised, controlled study on the effects of a short-term endurance training programme in patients with major depression. *British Journal of Sports Medicine, 41*(1), 29-33. doi:10.1136/bjsm.2006.030130

Krogh, J., Hjorthøj, C., Speyer, H., Gluud, C., & Nordentoft, M. (2017). Exercise for patients with major depression: a systematic review with meta-analysis and trial sequential analysis. *BMJ open, 7*(9). doi:10.1136/bmjopen-2016-014820

Krogh, J., Nordentoft, M., Sterne, J. A., & Lawlor, D. A. (2011). The effect of exercise in clinically depressed adults: systematic review and meta-analysis of randomized controlled trials. *The Journal of Clinical Psychiatry, 72*(4), 529-538. doi:10.4088/JCP.08r04913blu

Kvam, S., Kleppe, C. L., Nordhus, I. H., & Hovland, A. (2016). Exercise as a treatment for depression: A meta-analysis. *Journal of Affective Disorders, 202*, 67-86. doi:10.1016/j.jad.2016.03.063

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310

*Legrand, F. D., & Neff, E. M. (2016). Efficacy of exercise as an adjunct treatment for clinically depressed inpatients during the initial stages of antidepressant pharmacotherapy: An open randomized controlled trial. *Journal of Affective Disorders, 191*, 139-144. doi:10.1016/j.jad.2015.11.047

Macedo, L. G., Elkins, M. R., Maher, C. G., Moseley, A. M., Herbert, R. D., & Sherrington, C. (2010). There was evidence of convergent and construct validity of physiotherapy evidence database quality scale for physiotherapy trials. *Journal of Clinical Epidemiology, 63*(8), 920-925. doi:10.1016/j.jclinepi.2009.10.005

Maguire, N., Cullen, C., O'Sullivan, M., & O'Grady-Walshe, A. (1995). What do Dublin GPs expect from a psychiatric referral? *Irish Medical Journal, 88*(6), 215-216.

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M., & Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy, 83*(8), 713-721. doi:10.1093/ptj/83.8.713

*Martinsen, E. W., Medhus, A., & Sandvik, L. (1985). Effects of aerobic exercise on depression: a controlled study. *BMJ, 291*(6488), 109-109. doi:10.1136/bmj.291.6488.109

Moher, D. Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Journal of Clinical Epidemiology, 62*(10), 1006-1012. doi:10.1016/j.jclinepi.2009.06.005

Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., . . . Klassen, T. P. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The lancet, 352*(9128), 609-613. doi:10.1016/s0140-6736(98)01085-x

Morgan, D. (1989). Psychiatric cases: an ethnography of the referral process. *Psychological Medicine, 19*(03), 743-753. doi:10.1017/s003329170002434x

Morres, I., Stathi, A., Martinsen, E. W., & Sørensen, M. (2014). Physical exercise and major depressive disorder in adult patients. In A. G. Papaioannou & D. Hackfort (Eds.), *Routledge Companion to Sport and Exercise Psychology: Global perspectives and fundamental concepts* (pp. pp. 823-834). London: Taylor & Francis.

Morton, K. L., Biddle, S. J., & Beauchamp, M. R. (2008). Changes in self-determination during an exercise referral scheme. *Public Health, 122*(11), 1257-1260. doi:10.1016/j.puhe.2007.11.006

*Mota-Pereira, Silverio, J., Carvalho, S., Ribeiro, J. C., Fonte, D., & Ramos, J. (2011). Moderate exercise improves depression parameters in treatment-resistant patients with major depressive disorder. *Journal of Psychiatric Research, 45*(8), 1005-1011. doi:10.1016/j.jpsychires.2011.02.005

NICE. (2009). Depression in adults: recognition and management. *NICE guidelines [CG90], London, UK.*

*Oertel-Knöchel, V., Mehler, P., Thiel, C., Steinbrecher, K., Malchow, B., Tesky, V., . . . Hänsel, F. (2014). Effects of aerobic exercise on cognitive performance and individual psychopathology in depressive and schizophrenia patients. *European Archives of Psychiatry and Clinical Neuroscience, 264*(7), 589-604. doi:10.1007/s00406-014-0485-9

Pan, L., Wang, M., Xie, X., Du, C., & Guo, Y. (2014). Effects of anabolic steroids on chronic obstructive pulmonary disease: a meta-analysis of randomised controlled trials. *PloS One, 9*(1), e84855. doi:10.1371/journal.pone.0084855

*Perraton, L. G., Kumar, S., & Machotka, Z. (2010). Exercise parameters in the treatment of clinical depression: a systematic review of randomized controlled trials. *Journal of Evaluation in Clinical Practice, 16*(3), 597-604. doi:10.1111/j.1365-2753.2009.01188.x

*Pilu, A., Sorba, M., Hardoy, M., Floris, A., Mannu, F., Seruis, M., . . . Carta, M. (2007). Efficacy of physical activity in the adjunctive treatment of major depressive disorders: preliminary results. *Clinical Practice and Epidemiology in Mental Health, 3*(1), 8. doi:10.1186/1745-0179-3-8

Pinto, R. Z., Maher, C. G., Ferreira, M. L., Ferreira, P. H., Hancock, M., Oliveira, V. C., . . . Koes, B. (2012). Drugs for relief of pain in patients with sciatica: systematic review and meta-analysis. *BMJ, 344*. doi:10.1136/bmj.e497

Pollock, M. L., Gaesser, G. A., Butcher, J. D., Després, J.-P., Dishman, R. K., Franklin, B. A., & Garber, C. E. (1998). ACSM Position Stand: The Recommended Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory and Muscular Fitness, and Flexibility in Healthy Adults. *Medicine and Science in Sports and Exercise, 30*(6), 975-991. doi:10.1097/00005768-199806000-00032

Ravindran, A. V., Balneaves, L. G., Faulkner, G., Ortiz, A., McIntosh, D., Morehouse, R. L., . . . the, C. D. W. G. (2016). Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 5. Complementary and Alternative Medicine Treatments. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie, 61*(9), 576-587. doi:10.1177/0706743716660290

Rethorst, C. D., Wipfli, B. M., & Landers, D. M. (2009). The antidepressive effects of exercise. *Sports Medicine, 39*(6), 491-511. doi:10.2165/00007256-200939060-00004

Robertson, R., Robertson, A., Jepson, R., & Maxwell, M. (2012). Walking for depression or depressive symptoms: a systematic review and meta-analysis. *Mental Health and Physical Activity, 5*(1), 66-75. doi:10.1016/j.mhpa.2012.03.002

Rose, E. E., & Parfitt, G. (2007). A quantitative analysis and qualitative explanation of the individual differences in affective responses to prescribed and self-selected exercise intensities. *Journal of Sport and Exercise Psychology, 29*(3), 281-309. doi:10.1123/jsep.29.3.281

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641. doi:10.1037/0033-2909.86.3.638

Rueter, M. A. (1980). *The effect of running on individuals who are clinically depressed.* (MSc Thesis Master Thesis), Pennsylvania State University, USA.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. doi:10.1037/0003-066X.55.1.68

*Sadeghi, K., Ahmadi, S. M., Ahmadi, S. M., Rezaei, M., Miri, J., Abdi, A., . . . Jamshidi, K. (2016). A comparative study of the efficacy of cognitive group therapy and aerobic exercise in the treatment of depression among the students. *Global journal of health science, 8*(10), 1. doi:10.5539/gjhs.v8n10p1

Sadeghi, K., Ahmadi, S. M., Ahmadi, S. M., Rezaei, M., Miri, J., Abdi, A., . . . Jamshidi, K. (2016-a). A comparative study of the efficacy of cognitive group therapy and aerobic exercise in the treatment of depression among the students. *Global journal of health science, 8*(10), 1. doi:10.5539/gjhs.v8n10p1

Sadeghi, K., Ahmadi, S. M., Ahmadi, S. M., Rezaei, M., Miri, J., Abdi, A., . . . Jamshidi, K. (2016-b). A comparative study of the efficacy of cognitive group therapy and aerobic exercise in the treatment of depression among the students. *Global journal of health science, 8*(10), 1. doi:10.5539/gjhs.v8n10p1

*Salehi, I., Hosseini, S. M., Haghighi, M., Jahangard, L., Bajoghli, H., Gerber, M., . . . Brand, S. (2016). Electroconvulsive therapy (ECT) and aerobic exercise training (AET) increased plasma BDNF and ameliorated depressive symptoms in patients suffering from major depressive disorder. *Journal of Psychiatric Research, 76*, 1-8. doi:10.1016/j.jpsychires.2016.01.012

Salehi, I., Hosseini, S. M., Haghighi, M., Jahangard, L., Bajoghli, H., Gerber, M., . . . Brand, S. (2016-a). Electroconvulsive therapy (ECT) and aerobic exercise training (AET) increased plasma BDNF and ameliorated depressive symptoms in patients suffering from major depressive disorder. *Journal of Psychiatric Research, 76*, 1-8. doi:10.1016/j.jpsychires.2016.01.012

Salehi, I., Hosseini, S. M., Haghighi, M., Jahangard, L., Bajoghli, H., Gerber, M., . . . Brand, S. (2016-b). Electroconvulsive therapy (ECT) and aerobic exercise training (AET) increased plasma BDNF and ameliorated depressive symptoms in patients suffering from major depressive disorder. *Journal of Psychiatric Research, 76*, 1-8. doi:10.1016/j.jpsychires.2016.01.012

Schuch, F. B., Morres, I. D., Ekkekakis, P., Rosenbaum, S., & Stubbs, B. (2017). A critical review of exercise as a treatment for clinically depressed adults: time to get pragmatic. *Acta Neuropsychiatrica, 29*(2), 65-71. doi:10.1017/neu.2016.21

Schuch, F. B., Vancampfort, D., Richards, J., Rosenbaum, S., Ward, P. B., & Stubbs, B. (2016). Exercise as a treatment for depression: A meta-analysis adjusting for publication bias. *Journal of Psychiatric Research, 77*, 42-51. doi:10.1016/j.jpsychires.2016.02.023

*Schuch, Vasconcelos-Moreno, M. P., Borowsky, C., Zimmermann, A. B., Rocha, N. S., & Fleck, M. P. (2015). Exercise and severe major depression: effect on symptom severity and quality of life at discharge in an inpatient cohort. *Journal of Psychiatric Research, 61*(0), 25-32. doi:10.1016/j.jpsychires.2014.11.005

Schulz, K. F. (2001). Assessing allocation concealment and blinding in randomised controlled trials: why bother? *Evidence Based Nursing, 4*(1), 4-6. doi:10.1136/ebn.4.1.4

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA, 273*(5), 408-412. doi:10.1001/jama.1995.03520290060030

Searle, A., Calnan, M., Lewis, G., Campbell, J., Taylor, A., & Turner, K. (2011). Patients' views of physical activity as treatment for depression: a qualitative study. *The British Journal of General Practice, 61*(585), e149-e156. doi:10.3399/bjgp11X567054

Seehusen, D. A., & Sheridan, R. (2013). Second-generation antidepressants for depression in adults. *American Family Physician, 88*(10), 687-689.

Silveira, H., Moraes, H., Oliveira, N., Coutinho, E. S., Laks, J., & Deslandes, A. (2013). Physical exercise and clinically depressed patients: a systematic review and meta-analysis. *Neuropsychobiology, 67*(2), 61-68. doi:10.1159/000345160

Sørensen, M. (2006). Motivation for physical activity of psychiatric patients when physical activity was offered as part of treatment. *Scandinavian Journal of Medicine and Science in Sports, 16*(6), 391-398. doi:10.1111/j.1600-0838.2005.00514.x

Stanton, R., & Reaburn, P. (2014). Exercise and the treatment of depression: A review of the exercise program variables. *Journal of Science and Medicine in Sport, 17*(2), 177-182. doi:10.1016/j.jsams.2013.03.010

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*. doi:10.1136/bmj.d4002

Stonerock, G. L., Hoffman, B. M., Smith, P. J., & Blumenthal, J. A. (2015). Exercise as Treatment for Anxiety: Systematic Review and Analysis. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine, 49*(4), 542-556. doi:10.1007/s12160-014-9685-9

Tobi, P., Kemp, P., & Schmidt, E. (2017). Cohort differences in exercise adherence among primary care patients referred for mental health versus physical health conditions. *Primary Health Care Research & Development, 18*(5), 463-471. doi:10.1017/S1463423617000214

Turk, D. C., Rudy, T. E., & Salovey, P. (1984). Health protection: Attitudes and behaviors of LPNs, teachers, and college students. *Health Psychology, 3*(3), 189-210. doi:10.1037/0278-6133.3.3.189

Üstün, T. B., Ayuso-Mateos, J. L., Chatterji, S., Mathers, C., & Murray, C. J. L. (2004). Global burden of depressive disorders in the year 2000. *The British Journal of Psychiatry, 184*(5), 386-392. doi:10.1192/bjp.184.5.386

*Veale, D., Le Fevre, K., Pantelis, C., de Souza, V., Mann, A., & Sargeant, A. (1992). Aerobic exercise in the adjunctive treatment of depression: a randomized controlled trial. *Journal of the Royal Society of Medicine, 85*(9), 541-544. doi:10.1177/014107689208500910

Veale, D., Le Fevre, K., Pantelis, C., de Souza, V., Mann, A., & Sargeant, A. (1992-a). Aerobic Exercise in the Adjunctive Treatment of Depression: A Randomized Controlled Trial. *Journal of the Royal Society of Medicine, 85*(9), 541-544. doi:10.1177/014107689208500910

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*: Geneva: World Health Organization.

Table 1. Description of reviewed trials

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Mota-Pereira et al., (2011) Portugal outpatients	Eligible patients entered into study: 73% A. Aerobic Exercise (more depressed) Walking for 30-45min/5times/week/12weeks. One weekly session was individual/supervised by a sport scientist at hospital on treadmill, 5.0km/h/0° grade equal to 3.7-4 METs (moderate intensity). Home-based walking 4times/week, intensity prescription via accelerometers (>1952 counts) or perceived exertion (e.g., no shortness of breath). Compliance support: reminders, staff emphasized the exercise benefits, consulted family members on social support for exercise.	22/ 57.9	48.6/ 2.3	91 14	--- ---	100 [A]	HAMD-17 severe/ improved to mild [§, †]
	B. Usual pharmacotherapy (control group) Non-sedating antidepressants.	11/ 80	45.3/ 3.1	--- 9	--- ---	---	mild/ not improved
	Major Inclusion Criteria : treatment resistance after 9-15 months of combined pharmacotherapy, no regular exercising, no psychotherapeutic treatment, no change in medication in the last 6weeks						
Martinsen et al., (1985) Norway inpatients	A. Aerobic Exercise Group-based/supervised outdoor jogging & walking for 60min, 3times/week/9weeks at 50-70% VO_2max, supplemented by bicycling and/or swimming sessions at the patient's preference.	28/ ---	--- ---	--- 14.2	VO_{2max} [§, †]	32 [A] 100 [P,OT]	BDI moderate/ improved to mild [§, †]
	B. Occupational therapy (control group) Occupational therapy 60min, 3 times/week/9weeks.	21/ ---	--- ---	--- 10.5	not improved	66 [A] 100 [P]	moderate/ improved [§]
	Major Inclusion Criteria: both males and females						

Table 1 continued

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Schuch et al., (2015) Brazil inpatients	Eligible patients entered into study: 47% A. Aerobic Exercise Exercise for 45min, 3times/week/3weeks at self-selected type (stationary bicycle, treadmill, or stepper) and intensity exercise, provided that 16kcal/kg/week are completed. Individual/supervised at hospital.	25/ 72	38.8/ 11.5	--- 8	--- ---	80 [A] 8 [ECT]	HAMD-17 severe/ improved to normal [§, †]
	B. Treatment as usual (control group) All patients were prescribed antidepressants. ECT was prescribed to 8% of the patients.	25/ 76	41.7/ 10.4	--- 4	--- ---	--- ---	severe/ improved to mild [§]
	Major Inclusion Criteria: a score of ≥25 on HAMD-17						
Rueter, (1980) USA outpatients	A. Aerobic Exercise Individual/supervised (co-running) running for at least 20min, 3times/week/10weeks (public indoors sport track). Pace allowed talking, non-competitive, no distance/speed criteria.	11/ ---	--- ---	--- 18	--- ---	100 [A]	BDI moderate/ improved to minimal [§, †]
	B. Counseling Therapy (control group) Group & individual counseling for 10weeks, at least 30min/week. Also, on waiting list for participating in the aerobic exercise program.	11/ ---	--- ---	--- 18	--- ---	---	moderate/ not improved
	Major Inclusion Criteria: both males and females, a score of >15 on BDI						

Table 1 continued

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Sadeghi et al., (2016) Iran outpatients	A. Aerobic exercise Supervised, 50-60min/8weeks. Warm up: 10min stretching, breathing, exercises of upper/lower limbs, running in place at low intensity. The same exercises 60-80%MHR, 30-35min. Cooling down: same exercises at low intensity, 10-15min.	16/ 81	20.9/ 1.0	--- ---	--- ---	---	BDI moderate/ improved to mild §, †
	B. Cognitive Therapy Cognitive therapy for 8 weeks, 2 times/week/4weeks, and 1 time/week in the remaining 4 weeks.	16/ 75	21.1/ 1.2	--- ---	--- ---	---	moderate/ improved to mild §, †
	C. Group meetings (control group) Self-administrative group meetings of 45-60min to discuss issues raised by the group members.	14/ 79	20.9/ 1.2	--- ---	--- ---	---	moderate/not improved
	Major Inclusion Criteria: BDI score 13-28, no medication						
Salehi et al., (2016) Iran inpatients	Eligible patients entered into study: 86%						
	A. Aerobic exercise Individual/supervised cycling/treadmill for 40-45min, 3times/ week/4weeks at 60-75% VO2max.	20/ 25	29/ 5.1	100 0	--- ---	100 [A]	HAMD-17 severe/improved to mild §
	B. Electroconvulsive Therapy ECT 3times/week/4weeks.	20/ 35	29.7/ 6.28	100 0	--- ---	100 [A]	severe/ improved to mild §
	C. Aerobic Exercise + Electroconvulsive Therapy Aerobic exercise + ECT as described above.	20/ 35	30.2/ 6.2	100 0	--- ---	100 [A]	severe/ improved to mild §
	Major Inclusion Criteria: age 25 to 40, HAMD score of >2						

Table 1 continued

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Pilu et al., (2007) Italy outpatients	Eligible patients entered into study: 71%						
	A. Aerobic Exercise Group/supervised for 60min/2times/week/32weeks, 5min warm up, 50min on a self-selected cardio-fitness machine, change every 4min (20 options), 5min cooling down. Community-based gym.	10/ 100	--- ---	--- 0	--- ---	100 [A]	HAMD-17 severe/ improved to normal-mild [§, †]
	B. Pharmacotherapy (control group) 70% on SSRIs, 10% on SNRIs, and 5% of the patients on SSRI + SNRI + Tricyclic antidepressants.	20/ 100	--- ---	--- 0	--- ---	---	severe/not improved
	Major Inclusion Criteria: female, 2month persistence of score of >13 on HAMD-17 despite medication						
Legrand et al., (2016) France inpatients	Eligible patients entered into study: 73%						
	A. Aerobic exercise Supervised/individual daily brisk walking or jogging for 30min/10days at 65-75% of MHR, park outside hospital, 92% individual sessions.	14/ 64	45.3/ 10.0	--- 7.1	--- ---	100 [A]	BDI severe/improved to mild-moderate [§, †]
	B. Stretching exercise Supervised daily stretching 30min/10days on thighs, calves, gluteal, back, shoulders 60secs (60secs break), hospital.	11/ 72	41.0/ 13.2	--- 18	--- ---	100 [A]	severe/ improved [§]
	C. Pharmacotherapy (control group) All patients were treated with antidepressants; 70% on SSRI's, 20% on SSNRI's, 10% on dopamine agonist.	10/ 70	49.1/ 16.5	--- 10.5	--- ---		severe/not improved
	Major Inclusion Criteria: BDI score of >28, antidepressant drug therapy initiated < 2 weeks before participation in the trial						

Table 1 continued

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Veale et al., (1992) UK outpatients	Eligible patients entered into study: 71% A. Aerobic Exercise Group/supervised running 3imes/week/12 weeks in a public park just outside hospital. Warm-up/stretching before running.	48/ ---	--- ---	--- 25	VO_{2max} §, †	45 [A] 100 [P]	CIS moderate-severe/ improved §, †
	B. Treatment as usual (control group) Supportive psychotherapy.	35/ --	--- ---	--- 17.2	not improved	34 [A]	moderate-severe/ improved §
	Major Inclusion Criteria: >17 total weighted score and >2 depression severity score on the CIS						
Oertel-Knöchel et al., (2014) Germany inpatients	A. Aerobic exercise Group-based, supervised, for 45min, 3 times/week/4weeks at 60-70%MHR. Warm-up for 10min, 25min cardio training (aerobic exercise, aerobic with boxing exercises and circuit training) and 10min cool-down.	8/ 50	36.6/ 12.9	drop outs: 32% across all groups	--- ---	100 [CT]	BDI moderate/ improved to mild moderate §
	B. Relaxation Group-based, supervised, for 45min, 3times/week/4weeks. Breathing exercises, “enjoy exercises”, “imaginary journey”, relaxation or acceptance and awareness training. No yoga, no muscle progressive relaxation.	6/ 66	41.3/ 15.6		--- ---	100 [CT]	moderate/ improved §
	C. Waiting (control) No intervention.	8/ 38	42.2 8.3		--- ---	---	moderate/ not improved §
	Major Inclusion Criteria: stable medication 1 month before and during the trial, disease duration at least 5years						

Table 1 continued

Trial, Country, Sample	Interventions	N/ ♀ %	age m/sd	attend. dropouts %	physiolog. gains	concurr. psychiatr. therap. %	depression pre/post
Kerling et al., (2015) Germany inpatients	A. Aerobic exercise Group-based, supervised by study nurses, for 51min, 3times/week/6 weeks at moderate intensity and at 13max at the Borg scale. Warm up for 6min, 25min on a bicycle ergometer, and for 20min on arm ergometer, cross trainer, stepper, treadmill, and recumbent or rowing ergometer at personal preference.	22/ 45	44.2/ 8.5	>90 0	RHR§ VO_{2peak}§ LMax§ VAT§	77 [A] 100 [CBT]	MADRS moderate-severe/ improved to mild-moderate §
	B. Treatment as usual (control group) Antidepressant drugs and CBT were prescribed to 75% and 100% of the patients, respectively. As part of their routine treatment, all patients attended a daily supervised-based physical activity program of ball games, walking, and stretching at moderate intensity for 20min.	20/ 30	40.9/ 11.9	--- 0	not improved	100 [PA]	moderate-severe/ improved to mild-moderate §
	Major Inclusion Criteria: no acute/chronic infectious disease, no acute/lifetime immunological disease						

A: Antidepressants, BDI: Beck Depression Inventory, C: Counseling, CBT: Cognitive behavioral therapy, CIS: Clinical Interview Schedule, CT: Cognitive training, ECT: Electroconvulsive therapy, HAMD-17: Hamilton Rating Scale Depression, LMax: Maximum lactate concentration, MADRS: Montgomery-Åsberg Depression Rating Scale, Max: Maximum, M: Mean, METs: Metabolic Equivalents, MHR: Maximum heart rate, MIC: Major inclusion criteria, OT: Occupational therapy, P: Psychotherapy, PA: Physical activity, RHR: Resting heart rate, SD: Standard deviation, SNRI: Serotonin-norepinephrine reuptake inhibitor, SSRI: Selective serotonin reuptake inhibitor, VO_{2peak}: Peak oxygen uptake, VAT: Ventilatory anaerobic threshold, §: Statistically significant difference within group, †: Statistically significant difference in comparison to control group.

Table 2. Consensus scores of design quality for reviewed trials

PEDro Criteria	Random Allocation	Concealed Allocation	Baseline Balance	Blinding Patient/Therapist/Assessor	Dropout (<15%)	ITT	Statistical Comparison Between Groups	Point and Variability Measures	PEDro Total Score
Trials									
Schuch et al. (2015)	1	1	1	0/0/1	1	1	1	1	8
Kerling et al. (2015)	1	1	1	0/0/1	1	1	1	1	8
Legrand and Neff (2016)	1	1	1	0/0/0	1	1	1	1	7
Salehi et al. (2016)	1	1	1	0/0/1	1	1	1	1	7
Pilu et al. (2007)	1	0	1	0/0/0	1	1	1	1	6
Mota-Pereira et al. (2011)	1	1	0	0/0/1	1	0	1	1	6
Oertel-Knöchel et al. (2014)	1	0	1	0/0/1	0	1	1	1	6
Martinsen et al. (1985)	1	0	1	0/0/0	1	1	1	0	5
Sadeghi et al. (2016)	1	0	1	0/0/1	0	0	1	1	5
Veale et al. (1992-a)	1	0	1	0/0/0	0	0	1	1	4
Rueter (1980)	1	0	1	0/0/0	0	0	1	1	4

PEDro: Physiotherapy Evidence Database scale; ITT: Intention to Treat

Table 3. Meta-analytic findings of the effect of aerobic exercise on depression; overall effect and sensitivity analyses

	Trials/Arms	Treatment Effectiveness			Publication Bias		Heterogeneity	
		g	CI 95%	p value	Egger intercept	Begg-Mazumbar Kendall's tau	Cochrane Q	I^2
Overall effect	11/13	-0.79	-1.00, -0.57	0.00	-2.48, p=0.12	-0.22, p=0.29	15.20, p=0.23	21%
PEDro score of ≥6 [1,2,4-6,9,10]	7/8	-0.70	-0.94, -0.45	0.00	-1.87, p=0.42	-0.03, p=0.90	7.16, p=0.41	2%
Up to 4 weeks [2,5,9,10]	4/5	-0.71	-1.09, -0.34	0.00	-1.18, p=0.74	-0.00, p=1.00	5.79, p=0.22	30%
Exercise preferences [1,3,6,10]	4/4	-0.84	-1.17, -0.51	0.00	-3.73, p=0.52	-0.17, p=0.73	3.23, p=0.36	7%

CI 95%: Confidence Intervals; g: Hedge's g; PEDro: Physiotherapy of Evidence Database Scale.
[1]Kerling et al. 2015; [2]Legrand and Neff, 2016; [3]Martinsen et al. 1985; [4]Mota-Pereira et al. 2011; [5]Oertel-Knöchel et al. 2014; [6]Pilu et al. 2007; [9]Salehi et al. 2016; [10]Schuch et al. 2015.

Table 4. Meta-analytic findings of the effects of aerobic exercise on depression; subgroup analyses

	Trials/Arms	Treatment Effectiveness			Publication Bias		Heterogeneity	
		g	CI 95%	p value	Egger intercept	Begg-Mazumbar Kendall's tau	Cochrane Q	I^2
Participants								
Outpatients [4,6-8,11]	5/6	-0.84	-1.16, -0.51	0.00	-3.14, p=0.13	-0.53, p=0.13	6.12, p=0.29	18%
Inpatients [1-3,5,9,10]	6/7	-0.75	-1.06, -0.44	0.00	-1.36, p=0.67	-0.00, p=1.00	8.89, p=0.18	32%
Mild-moderate/moderate [3,5,7,8]	4/5	-0.97	-1.43, -0.51	0.00	-2.58, p=0.48	-0.10, p=0.81	6.79, p=0.15	41%
Moderate-severe/severe [1,2,4,6,9-11]	7/8	-0.71	-0.94, -0.48	0.00	-1.92, p=0.42	-0.18, p=0.54	7.16, p=0.41	2%
Intervention								
Equipment-based [1,5,6,9,10]	5/6	-0.67	-0.98, -0.35	0.00	-1.59, p=0.63	-0.00, p=1.00	6.59, p=0.25	24%
Equipment-free [2,3,7,8,11]	5/6	-0.94	-1.28, -0.60	0.00	-3.00, p=0.17	-0.40, p=0.26	6.96, p=0.22	28%
Group exercise [1,3,5,6,11]	5/5	-0.80	-1.09, -0.51	0.00	-0.62, p=0.80	-0.00, p=1.00	3.48, p=0.48	0%
Individual exercise [2,4,7,9,10]	5/6	-0.87	-1.30, -0.43	0.00	-4.39, p=0.14	-0.40, p=0.26	10.78, p=0.06	53%
Indoors [1,4-7,9,10]	7/8	-0.77	-1.10, -0.44	0.00	-3.54, p=0.13	-0.25, p=0.39	12.05, p=0.09	41%
Outdoors [2,3,11]	3/3	-0.94	-1.30, -0.58	0.00	-1.95, p=0.59	-0.00, p=1.00	1.37, p=0.50	0%
Hospital [1,5,9,10]	4/5	-0.61	-0.96, -0.27	0.00	-0.64, p=0.88	-0.00, p=1.00	5.62, p=0.23	28%
Non-hospital [2,3,6,7,11]	5/5	-1.07	-1.41, -0.72	0.00	-3.08, p=0.09	-0.30, p=0.46	4.65, p=0.33	14%
Comparisons								
Antidepressants or TAU [1,2,4,6,10,11]	6/6	-0.75	-1.01, -0.48	0.00	-1.38, p=0.35	-0.27, p=0.45	1.87, p=0.87	0%
Psychological treatments [1,3,7,8,11]	5/6	-0.85	-1.21, -0.48	0.00	-3.69, p=0.16	-0.40, p=0.26	8.64, p=0.12	42%
Outcomes								
Self-rated [2,3,5,7,8]	5/6	-0.97	-1.35, -0.59	0.00	-2.50, p=0.41	-0.13, p=0.71	6.81, p=0.24	26%
Clinician-rated [1,4,6,9-11]	6/7	-0.69	-0.94, -0.44	0.00	-1.56, p=0.62	-0.09, p=0.76	6.65, p=0.35	9%

CI 95%: Confidence Intervals; g: Hedge's g; TAU: Treatment as usual

[1]Kerling et al. 2015; [2]Legrand and Neff, 2016; [3]Martinsen et al. 1985; [4]Mota-Pereira et al. 2011; [5]Oertel-Knöchel et al. 2014; [6]Pilu et al. 2007; [7]Rueter, 1980; [8]Sadeghi et al. 2016; [9]Salehi et al. 2016; [10]Schuch et al. 2015; [11]Veale et al. 1992.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

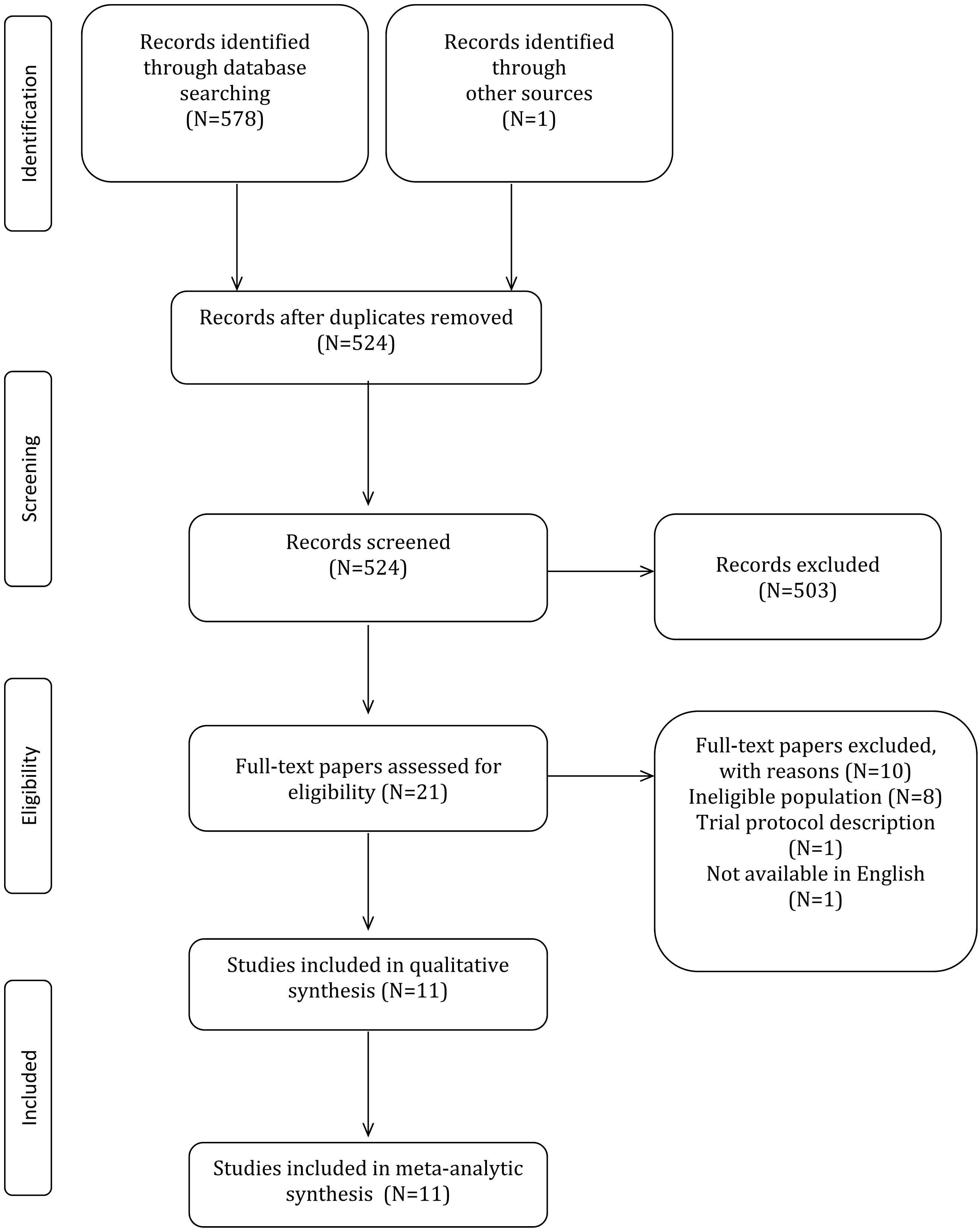

Figure 2. Meta-analysis of depressive symptom score.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Salehi 2016-a	-0,12	0,31	0,10	-0,73	0,49	-0,39	0,70
Kerling 2015	-0,45	0,31	0,09	-1,06	0,15	-1,48	0,14
Sadeghi 2016-b	-0,46	0,35	0,12	-1,14	0,23	-1,31	0,19
Ortel-Knoechel 2014	-0,63	0,49	0,24	-1,58	0,32	-1,30	0,19
Mota-Perreira 2011	-0,68	0,39	0,15	-1,44	0,09	-1,74	0,08
Veale 1992	-0,74	0,26	0,07	-1,25	-0,24	-2,89	0,00
Schuch 2015	-0,80	0,29	0,08	-1,37	-0,23	-2,76	0,01
Sadeghi 2016-a	-0,85	0,37	0,14	-1,58	-0,12	-2,29	0,02
Legrand 2016	-1,01	0,44	0,19	-1,87	-0,15	-2,31	0,02
Pilu 2007	-1,04	0,40	0,16	-1,82	-0,25	-2,60	0,01
Salehi 2016-b	-1,13	0,33	0,11	-1,79	-0,48	-3,38	0,00
Martinsen 1985	-1,23	0,33	0,11	-1,87	-0,58	-3,73	0,00
Reuter 1980	-2,01	0,56	0,31	-3,11	-0,91	-3,59	0,00
	-0,79	0,11	0,01	-1,01	-0,57	-7,17	0,00

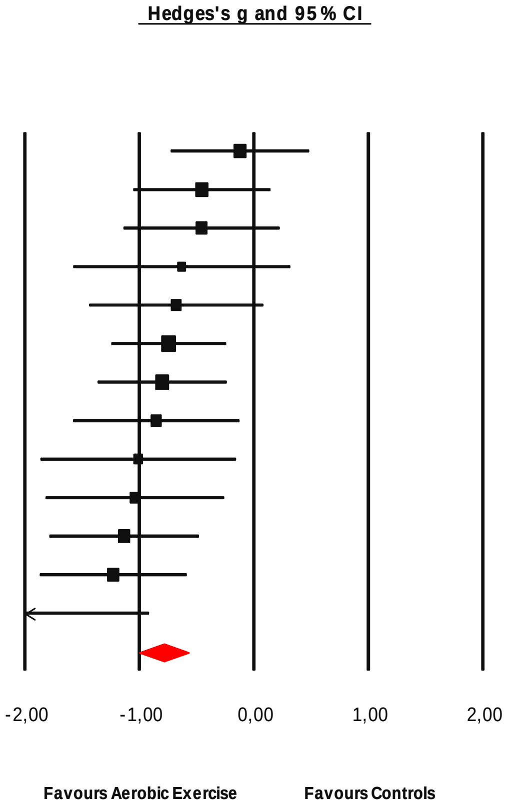

Figure 3. Funnel Plot.