Title: **Effectiveness of exercise-based interventions in reducing depressive symptoms in non-depressed people: a systematic review and meta-analysis of randomized controlled trials**

Authors: **Juan Ángel Bellón, MD, PhD[1,2,3,4], Sonia Conejo-Cerón, PhD[2,3], Almudena Sánchez-Calderón, MSc[4], Beatriz Rodríguez-Martín, PhD[2,5], Darío Bellón, MSc[6], Emiliano Rodríguez-Sánchez, MD, PhD[2,7], Juan Manuel Mendive, MD, PhD[2,8], Ignacio Ara, PhD[6,9,10], and Patricia Moreno-Peral, PhD[2,3]**

[1] 'El Palo' Health Centre, Health District of Primary Care Málaga-Guadalhorce, **SAS**, Málaga, Spain.

[2] Prevention and Health Promotion Research Network (**redIAPP**), **ISCIII**, Spain.

[3] Biomedical Research Institute of Malaga (**IBIMA**), Málaga, Spain.

[4] Faculty of Medicine, Department of Public Health and Psychiatry, University of Málaga (**UMA**), Spain.

[5] Faculty of Health Sciences, Talavera de la Reina, Toledo, Universidad de Castilla-La Mancha (**UCLM**), Spain.

[6] Faculty of Sport Sciences, Universidad de Castilla-La Mancha (**UCLM**), Toledo, Spain.

[7] 'Miguel Armijo' Health Centre, Castilla and León Health Service (**SACyL**). Department of Medicine, University of Salamanca (**USAL**), Salamanca, Spain.

[8] 'La Mina' Health Centre, Institut Català de la Salut (**ICS**), Barcelona, Spain.

[9] GENUD Toledo Research Group, **UCLM**, Toledo, Spain.

[10] CIBER of Frailty and Healthy Aging (**CIBERFES**), Madrid, Spain

Correspondence to:

Dr. Juan Ángel Bellón

Departamento de Salud Pública y Psiquiatría
Facultad de Medicina
Universidad de Málaga
Campus de Teatinos
29071 Málaga, Spain
jabellon@uma.es

Abstract word count: 249

Text word count: 4465

ABSTRACT

Background: In most trials and systematic-reviews that evaluate exercise-based interventions in reducing depressive symptoms, it is difficult to separate treatment from prevention.

Aim: To evaluate the effectiveness of exercise-based interventions in reducing depressive symptoms in non-depressed people.

Method: We searched PubMed, PsycINFO, Embase, WOS, SPORT-Discus, CENTRAL, Open-Grey and other sources up to May 25, 2020. We selected randomized controlled trials (RCTs) that compared exclusively exercise-based interventions with control groups, enrolling non-depressed participants, as measured using validated instruments, and whose outcome was reduction of depressive symptoms and/or incidence of new cases of depression. Pooled standardized mean differences (SMD) were calculated using random-effect models. (PROSPERO:CRD42017055726).

Results: A total of 14 RCTs (18 comparisons) evaluated 1737 non-depressed adults-elderly individuals from 8 countries and 4 continents. The pooled SMD was -0.34 (95%CI:-0.51 to -0.17; P<0.001) and sensitivity analyses confirmed the robustness of this result. We found no statistical evidence of publication bias and heterogeneity was moderate (I^2=54%; 95%CI:22% to 73%). Only 2 RCTs had an overall low risk of bias and 3 had long-term follow-up. Multivariate meta-regression found that a larger sample size, country (Asia), and selective prevention were associated with lower effectiveness, although only sample size remained significant when adjustment for multiple tests was considered. According to GRADE, the quality of evidence was low.

Conclusions: Exercise-based interventions have a small effect on the reduction of depressive symptoms in non-depressed people. It could be an alternative or complement to psychological programs, although further higher-quality trials with larger samples and long-term follow up are needed.

Key words: Depression; Prevention; Exercise; Randomized Controlled Trial; Systematic Review; Meta-Analysis.

INTRODUCTION

According to the WHO, 322 million people suffer from depression worldwide.[1] Between 2007 and 2017, the depression burden measured as years lived with disability increased by 14.1% and 14.8% for women and men, respectively,[2] ranking third (women) and fifth (men) in the world among 354 diseases. By 2030, it is estimated that depression will be the main cause of disease burden in high-income countries.[3]

Although effective therapies are available for depression, they only reduce disease burden by 30%.[4] In the theoretical situation that all people with depression received the appropriate treatment, the reduction of the disease burden would be limited due to the continuous occurrence of new cases of depression.[5]

The prevention of depression –which avoids the development of the disease– emerges as a plausible approach to reducing its disease burden.[6] The term primary prevention is reserved for only those interventions that occur before the onset of a disorder. Approaches to prevent the onset of depressive episodes have targeted people with prodromal symptoms not yet meeting the diagnostic criteria of a depressive disorder (*indicated prevention*), people at elevated risk because they have been exposed to risk factors (*selective prevention*), and the full population (*universal prevention*).[7] The overall aim of these three types of preventive intervention is the reduction of the occurrence of new cases. Usually, this is done through a risk reduction model, and even if outcomes are in the distant future and the goal of fewer cases have not yet been established, the decrease in risk and/or increase in protective factors can be documented,[7] even including estimations of the individual probability of suffering depression in the future.[8] Depressive symptoms are a good predictor of future incidence of depression,[9] and their reduction can be seen as an indicator of decreased risk. Additionally, the aims of indicated preventive interventions might be to reduce the length of time the early symptoms continue and to halt a progression of severity so that the individuals do not meet, nor do they come close to meeting, DSM (Diagnostic and Statistical Manual of Mental Disorders) diagnostic levels.[7]

In the last two decades, dozens of systematic reviews and meta-analyses (SR/MA) of the primary prevention of depression through psychological and psycho-educational interventions have been conducted, and from their global analysis it was concluded that these interventions have a small preventive effect, with the quality of evidence being high.[10-11]

The 2018 Physical Activity Guidelines Advisory Committee Scientific Report suggests strong evidence demonstrates that physical activity reduces the risk of experiencing depression and reduces depressive symptoms in individuals with and without major depression across the lifespan.[12] However, in most of the trials included in the SR/MA considered in this report it is difficult to separate treatment from prevention.

A recent meta-analysis of prospective studies found that higher levels of physical activity are consistently associated with lower odds of developing future depression.[13] The authors of a meta-analysis reported that exercise training is effective in reducing the symptoms of depression in sedentary patients with a chronic disease.[14] However, the reportedly higher effectiveness of physical activity in patients with mild to moderate depression at baseline might be due to the treatment rather than to the primary prevention of depression. In addition, improvement in symptoms of depression was the primary endpoint in only 3 of the 90 trials

included. Another meta-analysis revealed that symptoms of depression also improved through physical activity programs in patients without a diagnosis of clinical depression.[15] Yet, a large number of the trials included had either a before-after design, were non-randomized, or evaluated multi-component interventions (eg, exercise + diet) in which the preventive effect of exercise cannot be measured separately. The goal of our study was to evaluate the effectiveness of exercise-based interventions for the reduction of depressive symptoms in non-depressed individuals.

METHODS

We applied PRISMA guidelines for reporting SR/MA.[16] The study protocol was registered with the International Prospective Register of Systematic Reviews, (PROSPERO: CRD42017055726). As this meta-analysis is based on published data, no ethical approval was required.

Data Sources

We searched 6 electronic databases: PubMed, PsycINFO, Embase, Web of Science (WOS), SPORTDiscus, OpenGrey (System for Information on Grey Literature in Europe), and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to May 25, 2020. We performed a complementary search of the references provided in 56 relevant SR/MA (see **Appendix A**) and those listed in the studies selected. Experts were also contacted to identify further potentially relevant studies.

Search terms included: “physical activity or exercise,” “randomized controlled trial,” “depressive disorder,” “intervention,” and “prevention.” To increase search sensitivity, search terms were used in their broadest sense. We designed the search strategy based on a preliminary search of PubMed and adapted this for further searches of other databases (see **Appendix B**). Following the removal of duplicates, titles and abstracts were reviewed using our inclusion and exclusion criteria. Then, we reviewed the full text of the publications selected. In the cases in which the information required to determine the eligibility of a study was not provided, the authors were contacted to obtain the information required. Two pairs of reviewers (ASC with BRM and DB with SCC), each evaluating half of the records and full-text reports, independently performed the selection. Disagreements within each of the two pairs of reviewers were resolved by consensus or by the intervention of another reviewer (PMP), when appropriate.

Study Selection

Design

We focused the search on randomized controlled trials (RCT), as they provide evidence of causality and are considered the gold-standard for clinical trials.[17]

Participants

To ensure that the results obtained in our search were related to the reduction of depressive symptoms in non-depressed people, we only included RCTs in which participants with depression at baseline, detected

through structured standardized interviews (e.g., SCID) or validated self-reports with standard cut-off points (e.g., BDI-II), were excluded. The RCTs enrolling participants with and without depression at baseline were included when results were provided separately for both types of participants, but only including in the meta-analysis the subsample of non-depressed individuals at baseline. No restrictions were followed regarding diseases, pathologies or pregnancy. There were also no restrictions based on sociodemographic characteristics (age, sex, education level, etc.), setting (community, primary care, hospital, etc.) or publication language.

Intervention: exercise-physical activity

Although exercise is a subtype of physical activity, for the purpose of this review, we used the terms interchangeably. Exercise and physical activity were defined as any bodily movement generated by skeletal muscles that resulted in energy expenditure above resting levels.[18] The studies selected involved exercise-based interventions and provided data about the frequency, intensity, duration and type of exercise. Studies (or study arms) assessing the effectiveness of interventions combining exercise with another type of intervention known to be effective in the prevention of depression (e.g., exercise + psychological intervention) were excluded. Studies had to provide objective (e.g., assessed by accelerometer) or subjective (e.g., assessed by a questionnaire) evidence that physical activity was performed.

Comparators

Control groups could be usual care, no treatment (only evaluations), waiting list, attention control, or any type of placebo. The studies in which the comparator was an intervention that was proven to be effective in the prevention of depression (e.g., cognitive-behavioral therapy) were excluded.

Outcomes

The outcomes –either primary or secondary– of eligible studies included the reduction of symptoms of depression (as measured through validated scales of symptoms of depression) and/or the incidence of new cases of depression during follow-up (as measured through standardized interviews or validated scales of depressive symptoms as standard cut-off).

Data extraction

Data extracted from each study were recorded in an evidence table and extracted by two independent reviewers (PMP and SCC). Disagreements between the two reviewers were resolved by consensus or by the intervention of another reviewer (JAB), when necessary. Missing data was resolved by contacting authors when appropriate.

Evaluation of the risk of bias

The methodological quality of the RCTs was assessed using the Cochrane risk-of-bias tool (version 1).[19] From a qualitative approach, RCTs that were assessed to be at low risk of bias for all domains (sequence

generation, allocation concealment, blinding of outcome assessors, incomplete data analysis and selective reporting addressed) were considered to have an overall low risk of bias. The item 'blinding of participants and personnel' was excluded from this criterion because the nature of exercise-based interventions makes them difficult to blind. To treat the risk of bias as a quantitative variable for meta-regression analysis, each of the six criteria of the Cochrane tool were assigned 0 points when the risk was low, 1 point when the risk was uncertain, and 2 points when the risk was high. Therefore, the risk of bias of an RCT ranged from 0 (the lowest) to 12 (the highest). Two trained, independent reviewers (SCC and PMP) assessed the risk of bias. Disagreements between the two reviewers were resolved by consensus or by the intervention of another reviewer (JAB), when necessary.

Statistical analysis and synthesis

All statistical analyses were performed using 'STATA' version 14.2 and 'Comprehensive Meta-Analysis' (CMA) version 2.2.064.

Measure of effect

We used the Standardized Mean Difference (SMD) between the intervention and the control group as a measure of effect. For each RCT, we calculated the SMD by combining the SMD at different post-test follow-up times into a single estimate as the average, as well as its 95% confidence interval (CI). Negative SMDs indicated a better outcome (reduction of depressive symptoms) in the intervention group. Cohen proposed the following interpretation for this effect size: -0.2 is small; -0.5 medium, and -0.8 large.[20] For any RCT that included two different intervention groups (i.e., aerobic versus resistance) and one control group, standard errors in nested comparisons in the same RCT were inflated, following the recommendation of Cates.[21] A priori, we selected a random-effects model under the assumption that the RCTs to be included in our meta-analysis were performed in heterogeneous 'populations' that may differ from each other.[22]

Heterogeneity

We assessed the heterogeneity using I-squared,[23] which is expressed as a percentage, where heterogeneity is indicated as follows: 0% to 40% irrelevant, 30% to 60% moderate, 50% to 90% substantial, and 75% to 100% considerable.[19] We also calculated Cochran's *Q* test and its *P* value.

Publication bias

We evaluated publication bias by inspecting the funnel plot[24] and the Duval & Tweedie trim-and-fill procedure.[25] We also performed the Begg & Mazumdar rank correlation test.[26]

Analysis of sensitivity

We conducted sensitivity analyses to assess the robustness of the results by repeating the calculation of the pooled SMD in the first and last follow-up evaluation after the intervention, using Hedges's g and the profile likelihood method (an alternative to the DerSimonian-Laird method that is more conservative and

convenient when the number of studies is small), excluding the RCT that caused the greatest increase in heterogeneity and including only those RCTs with an overall low risk of bias.

Subgroup analysis and meta-regression

We used a mixed-effects model for subgroup analyses with the following a priori subgroups:

a) Participant characteristics: country of origin, clinical status (with or without chronic disease), sex, age (adult or elderly) and type of prevention (indicated, selective, or universal).

b) Exercise-based intervention characteristics: format (individual, group) verification (objective, subjective), supervision, walking, yoga, intervention duration, session duration, frequency, volume of exercise and intensity.

c) Methodological characteristics: measure of depression to exclude patients with clinical depression at baseline, measure of outcome, subsample (studies that included depressed and non-depressed participants at baseline but give separate outcomes for non-depressed participants), type of outcome (primary or secondary), type of comparator, qualitative and quantitative level of risk of bias, duration of follow-up, and sample size.

We performed multivariate random-effect meta-regressions to assess the impact of study characteristics (considered in advance) on the effect size, adjusting for other covariates and to explain heterogeneity across studies. A priori, we forced the variables "risk of bias'" and 'sample size' (as a proxy to quantify publication bias) into the multivariate meta-regression models to adjust for two of the five domains of the GRADE quality of evidence criteria, for which quantitative variables are available.[27,28,29] The post hoc analysis strategy to explain the maximum heterogeneity consisted of obtaining the most parsimonious meta-regression model (including the least number of variables) with the best goodness of fit. The normality of the quantitative covariates was checked using the skewness and kurtosis normality test;[30] and transformations were conducted, when appropriate, to approximate normality. We calculated standard errors and confidence intervals using the Knapp & Hartung method.[31] We calculated correlations between covariates to assess potential collinearity. To consider adjustment for multiple tests, we calculated *P* values with the permutation test following recommendations by Higgins & Thompson.[32] Finally, we used a normal probability plot of standardized shrunken residuals to estimate the goodness of fit of the meta-regression models.

The quality of evidence

The quality of evidence in the domains of risk of bias, consistency, directness, precision and publication bias were taken into account according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group methodology.[33]

RESULTS

Search Results

A total of 7640 articles were identified after eliminating duplicates. Of these, 418 articles were included for full-text review and 14 different RCTs[34-47] met the inclusion criteria, which included 18 valid comparisons for the meta-analysis (see **Figure 1**).

[Figure 1 here]

Characteristics of Included Studies

The characteristics of the 14 RCTs included are described in **eTable 1**. The RCTs were conducted in North America (N=6),[33,38,40,42,46,47] Europe (N=3),[36,41,45] Asia (N=3),[35,39,43] and South America (N=2).[37,44] All were published between 1999 and 2019. Overall, the 14 RCTs evaluated a total of 1737 participants (intervention group=1008; control group=729). Sample sizes ranged from 19 to 501 (median=77; IQR 27 to 124).

Regarding the target population, 9 RCTs were aimed at the adult population, and 5 predominantly concerned the elderly.[34,35,41,42,44] Three RCTs included only women (2 pre-postnatal[38,45] and 1 multiple sclerosis[43]), and another 5 included participants with physical chronic diseases (2 knee osteoarthritis,[42,44] 1 low back pain,[46] 1 multiple sclerosis[43] and 1 lung cancer[35]). The type of prevention was selective in 9 RCTs.

Regarding the exercise-based interventions, 9 RCTs provided aerobic exercise. Two RTCs included exclusively walking interventions[35,40] and 2 Iyengar yoga.[46,47] Most of the interventions were supervised (10 RCTs) and objectively verified (10 RCTs), with sessions of moderate intensity (10 RCTs), under 60 minutes (9 RCTs) and a frequency of 2 to 4 sessions per week (10 RCTs). Most interventions lasted 12 weeks or less (9 RCTs).

Methodological aspects included the following: in 13 RCTs, the primary outcome was the reduction of depressive symptoms; 7 RCTs had follow-up <6 months and only 3 RTCs between 12 and 24 months;[37,41,42] 9 RCTs had usual care as a comparator. All RCTs assessed reduction of depressive symptoms as an outcome and only one RCT also assessed the incidence of depression by standardized interview.[38]

Risk of Bias

The risk of bias of each of the RCTs is detailed in **eTable2**. Taking into account our standard for considering an RCT as having a qualitative overall low risk of bias, only 2 RCTs achieved this,[34,38] although the pooled SMD of these was slightly higher than the rest of the RCTs (see **Table 1**).

Effectiveness of the Interventions to Prevent Depression

Figure 2 shows each of the SMDs of the 18 comparisons for the 14 RCTs. The pooled SMD was -0.34 (95%CI: -0.51 to -0.17; P<0.001) for the random-effects model, which indicates that exercise-based interventions had a small, though significant effect on the reduction of depressive symptoms in non-

depressed people. There was moderate heterogeneity across the studies (I^2=54%; 95%CI: 22% to 73%) that was significant (Q=37.14; *P*=0.003). The primary analysis changed very little in the sensitivity analyses (**Table 1**).

[Figure 2 here]

[Table 1 here]

Publication Bias

The Begg & Mazumdar test to detect publication bias was not significant (z= -1.21; *P*=0.240). The Duval and Tweedie procedure did not impute any missing RCTs, and the funnel plot is shown in **eFigure 1**. Therefore, no statistical evidence for the presence of publication bias was found.

Subgroup analyses and meta-regression

eTable 3 shows the subgroup analyses. The effectiveness of exercise-based interventions to reduce depressive symptoms was higher (p≤0.001) in interventions with objective verification, and it was lower in the elderly, selective prevention, low intensity of exercise, and larger sample size (>200).

Unadjusted meta-regressions using standard errors by the Knapp & Hartung method, showed that selective prevention, larger sample size and RCTs using a subsample were statistically (p<0.05) associated with lower effectiveness to reduce depressive symptoms; and this was higher in interventions with a group format and RCTs with a waiting list as comparator (**Table 2**). When adjusted for risk of bias in meta-regression models, group format lost statistical significance; and after adjustment for sample size, only country (Asia) and selective prevention were associated with lower effectiveness (**Table 2**).

[Table 2 here]

A final meta-regression model including only two moderators explained 100% of the heterogeneity (**table 3**), and its goodness of fit was good (see **eFigure 2**). Larger sample size (β (log)=0.29 [0.16 to 0.42]; *P*<0.001) and country (Asia) (β=0.39 [95%CI, 0.05 to 0.73]; *P*=0.027) were significantly associated with lower effectiveness, although the latter did not reach significance when multiplicity adjustment (Higgins & Thompson permutation test) was performed.

[Table 3 here]

Quality of Evidence

The initial grading of the quality of evidence was high since we included only RCTs. Although the pooled effect size including only RCTs with an overall low risk of bias was slightly higher than the rest of the RCTs, there were very few RCTs with a low risk of bias, and therefore we reduced the rating from high to moderate. The heterogeneity was moderate, and although this was explained entirely (by 100%) through meta-regression, we reduced the rating from moderate to low. Indirectness was low since the target

population, the interventions and our outcome did not differ from those of primary interest. There was no statistical evidence of publication bias. We included a sufficient number of studies, and the total number of participants in our study allowed adequate precision. In summary, the quality of evidence according to GRADE was low.

DISCUSSION

Exercise-based interventions had a small effect on the reduction of depressive symptoms in non-depressed subjects and this result was robust in sensitivity analysis. Most of these interventions were aerobic, of moderate intensity, with 2 to 4 weekly sessions of 60 minutes or less, supervised and objectively verified. These findings were derived from 14 RCTs (18 comparisons) including 1737 adult and elderly participants from 8 countries and 4 continents. We found no publication bias but there was moderate heterogeneity, 100% of which was explained by only two moderators. There were only 2 RCTs with a low risk of bias and 3 with longer follow-up. Multivariate meta-regression showed that larger sample size, selective prevention and country (Asia) were associated with lower effectiveness. Finally, the strength of evidence, according to GRADE, was low.

Strengths

To the best of our knowledge, this is the first meta-analysis to evaluate the effectiveness of exercise-based interventions for the reduction of depressive symptoms in non-depressed people through RCTs conducted in all types of adult and elderly populations. Our strict inclusion criteria, analyzing only RCTs with participants free of depression at baseline, allowed us to clearly distinguish prevention from treatment. Our meta-analysis included a reasonable number of RCTs representing a large population of individuals with different characteristics and from different settings, which supports its external validity. We used multiple complementary electronic databases, 56 SR/MA and supplementary hand searching. The variety of databases utilized, combined with the broad range of search terms and no restriction on study publication language, contributed to a highly sensitive search. We applied rigorous methodology (PRISMA, GRADE) to the SR/MA process and the evaluation of the strength of evidence. We also performed sensitivity analyses, which support the robustness of the pooled SMDs in different setups (analyses and evaluation times) or when only RCTs with a low risk of bias were included. Finally, subgroup analyses and meta-regression allowed the identification of possible sources of heterogeneity, and multivariate meta-regression let us adjust for confounding biases and multiple comparisons and relativize our results according to the risk of bias and sample size.

Limitations

There are several limitations. First, from a qualitative perspective only 2 RCTs had an overall low risk of bias; therefore further RCTs of higher quality are needed. Second, only 3 RCTs had a longer follow-up (12 to 24 months); consequently, firm conclusions about long-term effectiveness cannot be drawn from our

study. Third, applying our selection criteria, we were unable to include RCTs with children and young adolescents because they had some exclusion criteria (e.g., did not discard participants with depression at baseline) and therefore conclusions cannot be inferred from this population. Fourth, it is difficult to draw conclusions about certain subgroups of interest (e.g., strength programs), due to the low number of RCTs included. Fifth, only one RCT measured incidence of new cases of depression and used a standardized diagnostic interview. As we mentioned previously in the introduction, the reduction of depression symptoms in non-depressed people is also included in the conceptual framework of depression prevention and has a positive and relevant effect on quality of life and cost.[48] However standardized diagnostic interviews generally has greater validity than symptom scales and the endpoint of preventive intervention is the reduction of the occurrence of new cases of depression. Therefore, further RCTs that assess the incidence of new cases of depression through standardized diagnostic interviews are also needed.

Comparison with Existing Literature

In our SR/MA, exercise-based group interventions were more effective than individual interventions in unadjusted meta-regression, although this did not reach statistical significance in the multivariate meta-regression. It has been argued that exercise-based group interventions could reduce depressive symptoms, in addition to physical activity itself, through social support and social learning, although there is no difference between individual and group intervention to prevent depression in the case of psychological interventions.[49] We found a trend toward greater effectiveness when physical activity was objectively verified versus subjective verification, but it was not statistically significant. Social desirability bias may cause participants to respond to physical activity verification questionnaires too optimistically, and variability in mood may influence the ability to accurately respond to self-report questionnaires.[50] The NICE (National Institute for Health and Clinical Excellence) in the UK, in the 2018 update,[51] recommends physical activity programs in less severe depression (including people with subthreshold depression) detailing the type of physical activity: delivery in groups (usually 8 people per group) by a competent practitioner, 45 minutes of aerobic exercises of moderate intensity, twice a week for 4 to 6 weeks , then weekly for a further 6 weeks of structured exercise. However, from the results of our SA/MA, no conclusions can be drawn about the characteristics of exercise-based interventions associated with the reduction of depressive symptoms in non-depressed people.

The effectiveness of yoga to reduce depressive symptoms could be mediated, in addition to physical exercise, by the person involved in relaxation, mindfulness and meditation.[52] However, regarding the yoga interventions included in our SR/MA,[46,47] we did not find significant differences in the pooled effect size in bivariate and multivariate meta-regression (Table 2). Attention control is used to achieve some degree of masking of the participants in RTCs. For example, Lewis et al.[38] used “general wellness topics support contact by telephone” as a control group versus an exercise-based intervention. Perhaps this type of control group could have a very small preventive effect for depression and therefore this could reduce the effectiveness of exercise in reducing depressive symptoms. Nevertheless, this was not found in our

analyses. Waiting list may be a nocebo condition in RCTs focused on anxiety and depression,[53] which would overestimate the effect of the interventions. In fact, we found higher effectiveness with waiting list as a comparator, but adjusted analysis by multivariate meta-regression cancelled out any statistical significance.

The Physical Activity Guidelines Advisory Committee (PAGAC) in the United States, in its report of February 2018,[12] affirms that physical activity reduces the risk of experiencing depression and depressive symptoms in individuals with and without major depression across the lifespan (PAGAC grade: Strong). This evidence was extracted from 38 SR/MA. Among the limitations of this overview are the overlap of primary studies in more than one SR/MA, which could contribute to obtaining biased estimates of effectiveness,[54] the inclusion of many low quality trials, the difficulty to separate depressed and non-depressed persons within the same primary study and the difficulty to separate the effect of exercise when it is in combination with other potentially effective treatments to prevent depression. In addition to our updated search for primary RCTs, the references of these 38 SR/MA and 18 more were also evaluated. Including only RCTs with non-depressed participants and exclusive exercise-based interventions, we could only incorporate 14 RCTs and, although we found a small preventive effect, the strength of evidence was low according to GRADE.

Practical implications

Physical activity may protect against depression, and/or depression may result in decreased physical activity.[55] Implementing regular exercise is difficult for most people and is even more challenging for those with major depression because of their symptoms of low energy and motivation.[56] From our study we can say that encouraging or prescribing regular exercise could be useful for the reduction of depressive symptoms in non-depressed people; although, due to the low quality of the evidence, the strength of this recommendation initially would be weak. However, balance between desirable and adverse effects, values and preferences of patients and providers, and costs-effectiveness analyses[33] were not included as outcomes in our systematic review, so any attempt to establish the strength of the recommendation would have a high uncertainty. Otherwise, exercise would have the advantage of acting preventively or as treatment for other mental (anxiety, insomnia, or dementia) and physical diseases (cardiovascular, diabetes, cancer, etc.).[12,57]

The effect size of psychological and exercise-based interventions for the reduction of depressive symptoms in non-depressed people could be similar.[10,11] Nonetheless, from our study we can conclude that so far the quality of evidence is lower for exercise-based interventions than for psychological interventions.[11] Further RCTs including exercise-based interventions with a low risk of bias, larger samples and longer follow-up are needed as well as others directly comparing psychological versus exercise-based interventions. Finally, studies to establish the optimal type, intensity, frequency and duration of the exercise-based interventions are also required.

The effectiveness of the programs for the primary prevention of depression might be small. However, if such programs were scaled to a large part of the population, their impact in terms of increased health, quality of life, and cost reduction would be relevant.[5-6] If people had two alternatives with similar effectiveness, psychological and exercise-based interventions, the impact would be even greater, since those who are little motivated by psychological programs could be more motivated by exercise programs and vice versa. Clinicians could encourage and advise their patients towards either intervention or both, and massive prevention programs in schools,[58] workplaces[59] and through information and communication technologies[29,60] might also be implemented.

AUTHORS' CONTRIBUTIONS

JAB, SCC and PMP designed the study and the other authors collaborated on the design. ASC, BRM, DB, SCC and PMP participated in the selection of studies and data extraction. SCC, PMP and JAB assessed the risk of bias. JAB performed the statistical analysis and drafted the manuscript and all authors discussed and approved the final version. JAB is the guarantor.

FUNDING STATEMENT

This study is supported by the Spanish Ministry of Health, Carlos III Institute of Health, and the European Regional Development Fund «Una manera de hacer Europa» (grant FIS references: PI15/00401; PI15/01021; PI15/01035) and the Andalusian Council of Health (grant reference: AP-0095-2016); Biomedical Research Networking Center on Frailty and Healthy Aging (CIBERFES) and FEDER funds from the European Union (CB16/10/00477); as well as by the Prevention and Health Promotion Research Network 'redIAPP' (RD12/0005 and RD16/0007).

COMPETING INTERESTS STATEMENT

The authors all declare they have no competing interests. The funders had no direct role in the design of the study and they will have no input on the interpretation or publication of study results.

REFERENCES

1. World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva: 2017.
2. GBD. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018 ;392(10159):1789-1858
3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:e442.
4. Chisholm D, Sanderson K, Ayuso-Mateos JL, et al. Reducing the global burden of depression: population-level analysis of intervention cost effectiveness in 14 world regions. Br J Psychiatry. 2004;184: 393-403.
5. Cuijpers P, Beekman AT, Reynolds C. Preventing depression: a global priority. JAMA 2012;307:1033–4. doi:10.1001/jama.2012.271.
6. Muñoz RF, Cuijpers P, Smit F, et al. Prevention of major depression. Annu Rev Clin Psychol. 2010;6:181-212.
7. Mrazek PJ & Haggerty R. Reducing risks for mental disorders. Washington D. C. EE. UU: National Academy Press: 1994.
8. Bellón JÁ, Luna JD, King M, et al. Predicting the onset of major depression in primary care: international validation of a risk prediction algorithm from Spain. Psychol Med. 2011;41:2075-88.
9. Cuijpers P & Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. Acta Psychiatr Scand. 2004;109: 325–331
10. Bellón JA, Moreno-Peral P, Motrico E, et al. Effectiveness of psychological and/or educational interventions to prevent the onset of episodes of depression: A systematic review of systematic reviews and meta-analyses. Prev Med. 2015;76 Suppl:S22-32.
11. van Zoonen K, Buntrock C, Ebert DD, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. Int J Epidemiol. 2014;43(2):318-29. doi: 10.1093/ije/dyt175.
12. 2018 Physical Activity Guidelines Advisory Committee. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: U.S. Department of Health and Human Services, 2018.
13. Schuch FB, Vancampfort D, Firth J, et al. Physical activity and incident depression: a meta-analysis of prospective cohort studies. Am J Psychiatry. 2018;175(7):631-648. doi:10.1176/appi.ajp.2018.17111194
14. Herring MP, Puetz TW, O'Connor PJ, et al. Effect of exercise training on depressive symptoms among patients with a chronic illness: a systematic review and meta-analysis of randomized controlled trials. Arch Intern Med. 2012;172(2):101-11.
15. Conn VS. Depressive symptom outcomes of physical activity interventions: meta- analysis findings. Ann Behav Med. 2010;39(2):128-38.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Ann Intern Med 2009;151(4): 264-9.
17. Piantadosi S. Clinical Trials: A Methodological Perspective, 2nd Edition. New Jersey, NJ: John Wiley & Sons; 2005.
18. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100:126–31.
19. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Updated March 2011. http://handbook.cochrane.org/
20. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd Edition. Hillsdale, NJ: Erlbaum; 1988.
21. Rücker G, Christopher J, Cates CJ and Schwarzer G. Methods for including information from multi-arm trials in pairwise meta-analysis. Res Synth Methods. 2017;8(4):392-403.

22. Cooper H, Hedges LV, Valentine JC. The handbook of research synthesis and meta- analysis. 2nd Edition. Russel Sage Foundation, New York: 2009.
23. Higgins JP, Thompon SG, Deeks JJ, et al. Measuring inconsistency in meta- analyses. BMJ. 2003;327:557-60.
24. Stern JACM, Egger M, Smith GD. Investigating and dealing with publication and other bias. In: Systematic Reviews in Health Care: Meta-Analysis in context, 2nd edition, ed Egger M, Smith GD, Altman DG. London, BMJ Books: 2001:189-208.
25. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000;56:455-63.
26. Begg CB, Mazumdar M. Operating Characteristics of a Rank Correlation Test for Publication Bias. Biometrics 1994;50:1088. doi:10.2307/2533446
27. Moreno-Peral P, Conejo-Cerón S, Rubio-Valera M, et al. Effectiveness of Psychological and/or Educational Interventions in the Prevention of Anxiety: A Systematic Review, Meta-analysis, and Meta-regression. JAMA Psychiatry. 2017;74(10):1021-1029.
28. Rigabert A, Motrico E, Moreno-Peral P, et al. Effectiveness of online interventions in preventing depression: a protocol for systematic review and meta-analysis of randomised controlled trials. BMJ Open. 2018 Nov 28;8(11):e022012.
29. Rigabert A, Motrico E, Moreno-Peral P, et al. Effectiveness of online psychological and psychoeducational interventions to prevent depression: Systematic review and meta-analysis of randomized controlled trials. Clinical Psychology Review. 2020 Oct 24;82:101931.
30. D'Agostino RB, Belanger AJ, D'Agostino RB Jr. A suggestion for using powerful and informative tests of normality. American Statistician 1990;44: 316-321.
31. Knapp G, Hartung J. Improved tests for a random-effects meta-regression with a single covariate. Stat Med. 2003; 22(17):2693-710.
32. Higgins JPT, Thompson SG. Controlling the risk of spurious findings from meta- regression. Stat Med. 2004; 23(11):1663-82.
33. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401–6.
34. Brenes GA, Williamson JD, Messier SP, et al. Treatment of minor depression in older adults: a pilot study comparing sertraline and exercise. Aging Ment Health. 2007;11(1):61-68. doi: 10.1080/13607860600736372
35. Chen HM, Tsai CM, Wu YC, Lin KC, Lin CC. Randomised controlled trial on the effectiveness of home-based walking exercise on anxiety, depression and cancer-related symptoms in patients with lung cancer. Br J Cancer. 2015;112(3):438-445. doi: 10.1038/bjc.2014.612.
36. de Zeeuw ELEJ, Tak ECPM, Dusseldorp E, Hendriksen IJM. Workplace exercise intervention to prevent depression: A pilot randomized controlled trial. Ment Health Phys Act. 2010;3(2):72-77. doi: 10.1016/j.mhpa.2010.09.002
37. DiLorenzo TM, Bargman EP, Stucky-Ropp R, Brassington GS, Frensch PA, LaFontaine T. Long-term effects of aerobic exercise on psychological outcomes. Prev Med. 1999;28(1):75-85. doi: 10.1006/pmed.1998.0385
38. Lewis BA, Gjerdingen DK, Avery MD, et al. A randomized trial examining a physical activity intervention for the prevention of postpartum depression: The healthy mom trial. Ment Health Phys Act. 2014;7(1):42-49. doi: 10.1016/j.mhpa.2013.11.002
39. Mohammadi F, Malakooti J, Babapoor J, Mohammad-Alizadeh-Charandabi S. The effect of a home-based exercise intervention on postnatal depression and fatigue: A randomized controlled trial. Int J Nurs Pract. 2015;21(5):478-485. doi: 10.1111/ijn.12259.
40. Osei-Tutu KB, Campagna PD. The effects of short- vs. long-bout exercise on mood, VO2max, and percent body fat. Prev Med. 2005;40(1):92-98. doi: 10.1016/j.ypmed.2004.05.005
41. Pakkala I, Read S, Leinonen R, Hirvensalo M, Lintunen T, Rantanen T. The effects of physical activity counseling on mood among 75- to 81-year-old people: a randomized controlled trial. Prev Med. 2008;46(5):412-418. doi: 10.1016/j.ypmed.2007.11.002
42. Penninx BW, Rejeski WJ, Pandya J, et al. Exercise and depressive symptoms: a comparison of aerobic and resistance exercise effects on emotional and physical function in older persons with high and low depressive symptomatology. J Gerontol B Psychol Sci Soc Sci. 2002;57(2):124-132. doi: 10.1093/geronb/57.2.p124

43. Sadeghi-Bahman D, Razazian N, Farnia V, Alikhani M, Tatari F, Brand S. Compared to an Active Control Condition, in Persons With Multiple Sclerosis Two Different Types of Exercise Training Improved Sleep and Depression, but Not Fatigue, Paresthesia, and Intolerance of Uncertainty. Mult Scler Relat Disord 2019;36:101356. doi: 10.1016/j.msard.2019.07.032.
44. Taglietti M, Facci LM, Trelha CS, et al. Effectiveness of aquatic exercises compared to patient-education on health status in individuals with knee osteoarthritis: a randomized controlled trial. Clin Rehabil. 2018;32(6):766-776. doi: 10.1177/0269215517754240.
45. Vargas-Terrones M, Barakat R, Santacruz B, Fernandez-Buhigas I, Mottola MF. Physical exercise programme during pregnancy decreases perinatal depression risk: a randomised controlled trial. Br J Sports Med. 2019;53(6):348-353. doi: 10.1136/bjsports-2017-098926.
46. Williams K, Abildso C, Steinberg L, et al. Evaluation of the effectiveness and efficacy of Iyengar yoga therapy on chronic low back pain. Spine. 2009;34(19):2066-2076. doi: 10.1097/BRS.0b013e3181b315cc.
47. Woolery A, Myers H, Sternlieb B, Zeltzer L. A yoga intervention for young adults with elevated symptoms of depression. Altern Ther Health Med. 2004;10(2):60-63.
48. Lynch FL, Hornbrook M, Clarke GN, et al. Cost-effectiveness of an intervention to prevent depression in at-risk teens. Arch Gen Psychiatry. 2005;62(11):1241-1248.
49. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry. 2008;165(10):1272-1280.
50. Rosenbaum S, Morell R, Abdel-Baki A, et al. Assessing physical activity in people with mental illness: 23-country reliability and validity of the simple physical activity questionnaire (SIMPAQ). BMC Psychiatry. 2020;20(1):108.
51. Depression in adults: treatment and management. NICE guideline: short version. Draft for second consultation, May 2018: https://www.nice.org.uk/guidance/GID-CGWAVE0725/documents/short-version-of-draft-guideline
52. de Manincor M, Bensoussan A, Smith C, et al. Establishing key components of yoga interventions for reducing depression and anxiety, and improving well-being: a Delphi method study. BMC Complement Altern Med 2015;15:85.
53. Furukawa TA, Noma H, Caldwell DM, et al. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatr Scand. 2014;130(3):181-92.
54. Pieper D, Buechter R, Jerinic P, Eikermann M. Overviews of reviews often have limited rigor: a systematic review. J. Clin. Epidemiol. 2012;65:1267–73.
55. Choi, Chen CY, Stein MB et al. Assessment of Bidirectional Relationships Between Physical Activity and Depression Among Adults: A 2-Sample Mendelian Randomization Study. JAMA Psychiatry. 2019;76(4):399-408. doi:10.1001/jamapsychiatry.2018.4175
56. Köhler-Forsberg O, Cusin C, Nierenberg AA. Evolving Issues in the Treatment of Depression. JAMA. 2019;321(24):2401-2402
57. Pedersen BK, Saltin B. Exercise as medicine – evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015: (Suppl. 3) 25: 1–72
58. Werner-Seidler A, Perry Y, Calear AL, et al. School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. Clin Psychol Review 2017;51:30–47.
59. Bellón JÁ, Conejo-Cerón S, Cortés-Abela C, et al. Effectiveness of psychological and educational interventions for the prevention of depression in the workplace: A systematic review and meta-analysis. Scand J Work Environ Health. 2019 Jul 1;45(4):324-332.
60. Jahangiry L, Farhangi MA, Shab-Bidar S, Rezaei F, Pashaei T. Web-based physical activity interventions: a systematic review and meta-analysis of randomized controlled trials. Public Health 2017;152:36-46.

Table 1: Effectiveness of exercise-based interventions in reducing depressive symptoms in non-depressed people.

Effectiveness	Number of comparisons	SMD (95% C.I.)	*P* value	I^2 (95% C.I.)
Primary analysis (#)	18	-0.34 (-0.51 to -0.17)	<.001	54 % (22% to 73%)
Sensitivity analyses				
At first evaluation post-intervention	18	-0.32 (-0.49 to -0.15)	<.001	54 % (23% to 73%)
At last evaluation post-intervention	18	-0.35 (-0.52 to -0.18)	<.001	55 % (23% to 73%)
Hedges' g	18	-0.33 (-0.49 to -0.17)	<.001	54% (22% to 73%)
Profile likelihood method*	18	-0.33 (-0.51 to -0.17)	<.001	45% (22% to 73%)
[a] Pakkala et al., 2008 excluded	17	-0.37 (-0.52 to -0.23)	<.001	26% (0% to 59%)
[b] Including only RCTs with low risk of bias	2	-0.55 (-0.87 to -0.22)	0.001	0% (#)
[c] Woolery et al., 2004 excluded	17	-0.30 (-0.45 to -0.14)	<.001	45% (2% to 69%)

*Abbreviations: **SMD**, Standardized Mean Difference. (#) taking the different post-intervention evaluations as an average.*

**Between-studies variance estimate (τ^2): 0.042 (95% CI: 0.000 to 0.157)*

[a] The randomized controlled trial that most increased heterogeneity

 # It is not possible to calculate the confidence interval because degrees of freedom (n-1) must be at least 2.

[b] Low risk of bias criteria for inclusion (randomized controlled trials that scored low risk of bias in sequence generation, allocation concealment, blinding of outcome assessors, incomplete outcome data addressed and selective reporting): Brenes et al, 2007 and Lewis et al., 2014.

[c] This randomized controlled trial might be an outlier.

Table 2: Coefficient statistics of unadjusted and adjusted meta-regression on the association between reduction of depressive symptoms (SMD) and other covariates.

Independent variables	Unadjusted Coefficient*	P	Adjusted for risk of bias	P	Adjusted for sample size	P	Adjusted for risk of bias and sample size	P
Participant characteristics								
Country (Asia)	.239 (-.178 to .654)	.242	.230 (-.193 to .652)	.265	.389 (.051 to .726)	.027	.380 (.039 t0 .722)	.032 ‡[a]
Sex (women)	-.039 (-.465 to .387)	.850	-.094 (-.532 to .344)	.655	-.018 (-.371 to .335)	.914	-.047 (-.429 to .333)	.792
Age (elderly)	.261 (-.059 to .581)	.104	.251 (-.072 to .574)	.118	.010 (-237 to .437)	.538	.104 (-.244 to .453)	.532
Chronic (Yes)	.111 (-.254 to .475)	.529	.130 (-.239 to .499)	.464	.077 (-.210 to .364)	.576	.102 (-.199 to .404)	.478
Prevention (Selective)	.604 (.211 to .997)	.005	.610 (.175 to 1.056)	.009	.442 (.033 to .852)	.036	.460 (.005 to .916)	.048 ‡[b]
Intervention characteristics								
Type of exercise (aerobic)	-.040 (-.399 to .320)	.818	-.046 (-.412 to .319)	.790	-.146 (-.449 to 156)	.319	-.156 (-.469 to .157)	.304
Walking (Yes)	-.171 (-.761 to .419)	.548	-.142 (-.740 to .456)	.621	.036 (-.513 to .586)	.889	.036 (-.527 to .560)	.892
Yoga (Yes)	-.412 (-1.00 to .178)	.158	-.461 (-1.057 to .134)	.120	-.256 (-.761 to .248)	.296	-.315 (-.852 to .223)	.230
Type of verification (Objective)	-.272 (-.592 to .049)	.091	-.238 (-.592 to .116)	.173	-.225 (-.458 to .008)	.057	-.235 (-.509 to .038)	.086
Supervised exercise (Yes)	-.221 (-.562 to .119)	.187	-.181 (-.581 to .218)	.349	-.130 (-.417 to .158)	.352	-.115 (-.473 to .243)	.503
Format (group)	-.345 (-.667 to -.023)	.037	-.325 (-.661 to .010)	.056	-.196 (-.526 to .134)	.225	-.193 (-.538 to .153)	.251
Duration of intervention (up to 12 weeks)	-.252 (-.588 to .085)	.133	-.217 (-.595 to .161)	.241	-.016 (-.407 to .375)	.932	.026 (-.404 to .456)	.899
Frequency of sessions (2-4/week)	-.070 (-.456 to .317)	.707	.054 (-424 to .532)	.812	.049 (-.275 to .373)	.752	.214 (-.219 to .647)	.307
Duration of sessions (up to 60 minutes)	-.163 (-.504 to .178)	.325	-.191 (-.528 to .146)	.245	-.011 (-.350 to .328)	.945	-.033 (-.394 to .328)	.846
Volume (up to 150 minutes/week)	-.240 (-.566 to .086)	.138	-.225 (-.556 to .106)	.169	-.032 (-.403 to .338)	.856	-.033 (-.418 to .352)	.857
Intensity (low)	.252 (-.083 to .587)	.130	.209 (-.164 to .583)	.251	.184 (-.075 to .443)	.151	.176 (-.134 to .487)	.244

Methodological characteristics								
Independent variables	**Unadjusted Coefficient***	**P**	**Adjusted for risk of bias**	**P**	**Adjusted for sample size**	**P**	**Adjusted for risk of bias and sample size**	**P**
Sample size (ln)	.220 (.074 to .366)	.006	.205 (.045 to .365)	.015 ‡[c]	----------	----	----------	----
Follow-up months (ln)	.182 (-.002 to .367)	.053	.179 (-.005 to .364)	.056	-.187 (-.637 to .262) †	.388	-150 (-.674 to .374) †	.548
Risk of bias (sqrt)	-.173 (-.534 to .189)	.326	----------	----	-.093 (-.388 to .201)	.509	----------	----
Depression exclusion at baseline (standardized diagnostic interview)	-.053 (-.544 to .439)	.823	-.170 (-.694 to .353)	.499	-.033 (-.444 to .377)	.864	-.107 (-.574 to .359)	.628
Subsample (#)	.320 (.019 to .621)	.038	.364 (.090 to .638)	.013	.125 (-.255 to.506)	.493	.206 (-.209 to .621)	.305
Outcome measure (standardized diagnostic interview)	-.233 (-.872 to .406)	.451	-.527 (-1.198 to .145)	.115	-.270 (-.724 to .184)	.225	-.449 (-.937 to .038)	.068
Type of outcome (secondary)	.093 (-.729 to .914)	.814	.096 (-.731 to .923)	.808	.293 (-.431 to 1.018)	.402	.281 (-.459 to 1.02)	.429
Comparator (waiting list)	-1.30 (-2.41 to .186)	.025	-1.234 (-2.376 to -.092)	.036	-1.029 (-2.109 to .050)	.060	-1.01 (-2.11 to .098)	.071

(*)*The coefficient means the change of the dependent variable (SMD: standardized mean difference of depressive symptoms between intervention and control groups) with each unit increase of the independent variables. A negative coefficient increases the preventive effect (reduction of symptoms) and a positive coefficient the opposite.* ***(ln)****: Neperian logarithm;* ***(sqrt)****: square root.* ***(#)*** *Studies that included depressed and non-depressed participants at baseline but give separate outcomes for non-depressed;* ***(‡)*** *Higgins & Thompson permutation test to calculate P values considering multiplicity adjustment (Monte Carlo approach with 20000 permutations): (****a****)Country (Asia) p= .105 (95%CI: .100 to .109); (****b****)Prevention (selective) p=.120 (95%CI: .115 to .125); (****c****)Sample size (lg) p= .043 (95%CI: .041 to .047) .* ***(†)*** *the variables follow-up (ln) and sample size (ln) had a high correlation (r =0.91) and as a consequence including both variables the probability of collinearity of the meta-regression model was very high, and so the estimation of coefficients would be biased.*

Table 3: Final meta-regression model

Final Model [a]	β (95% CI) [b]	*P* Value	*P* Value (95% CI) [c]
Sample size (ln) [d]	0.293 (0.164 to 0.422)	<0.000	0.0011 (0.0007 to 0.0017)
Country (Asia)	0.389 (0.051 to 0.727)	0.027	0.0667 (0.0632 to 0.0702)

[a] Model $F_{2,15}$=11.85; *P*=0.0008; I^2 residual = 0%; Adjusted R^2 = 100%.
[b] Knapp & Hartung method for estimation of Standard Error and 95% Confidence Intervals (CI).
[c] Higgins & Thompson permutation test to calculate *P* values considering multiplicity adjustment (Monte Carlo approach with 20000 permutations).
[d] Neperian logarithm transformation.

Figure 1: PRISMA Flowchart

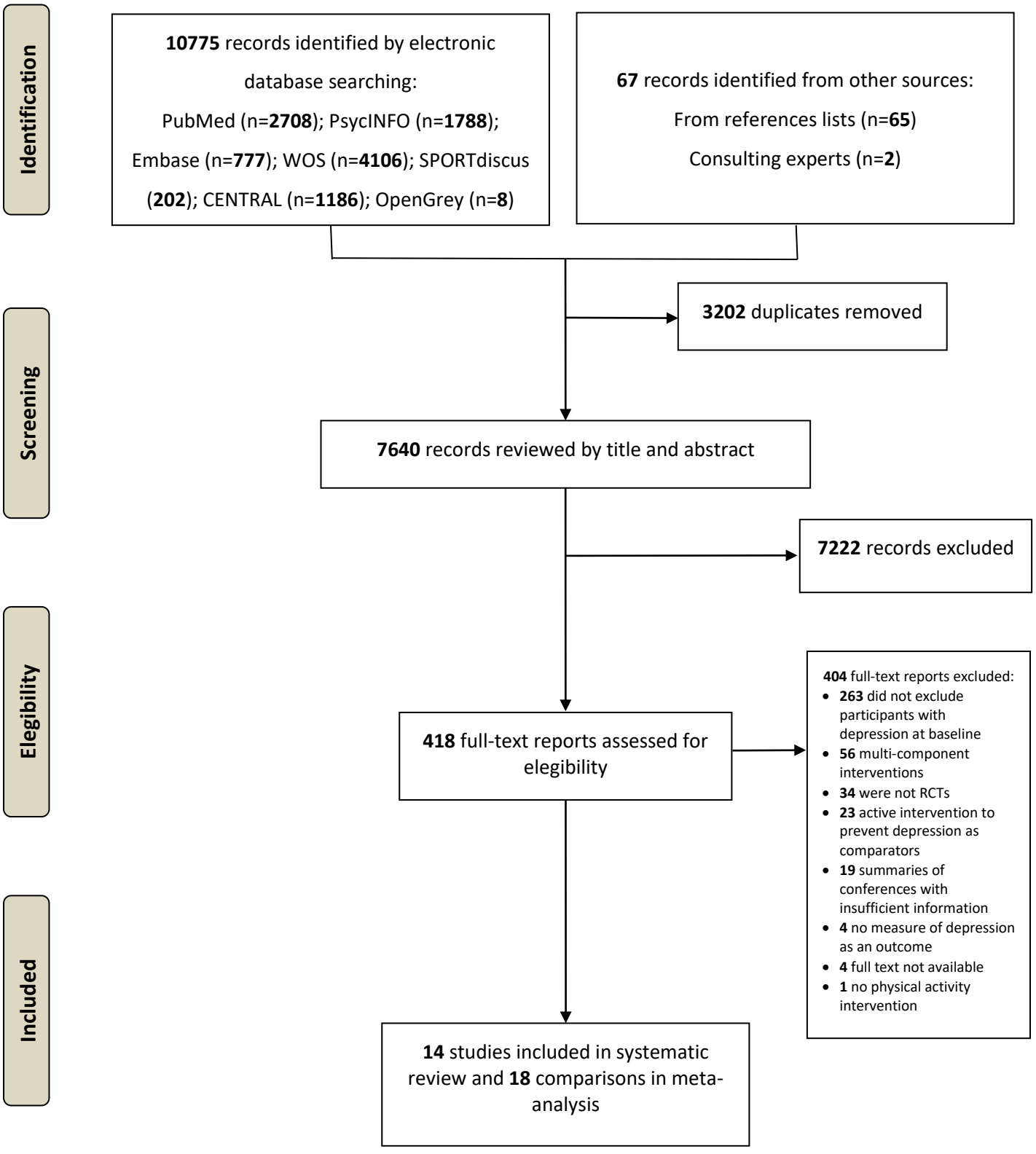
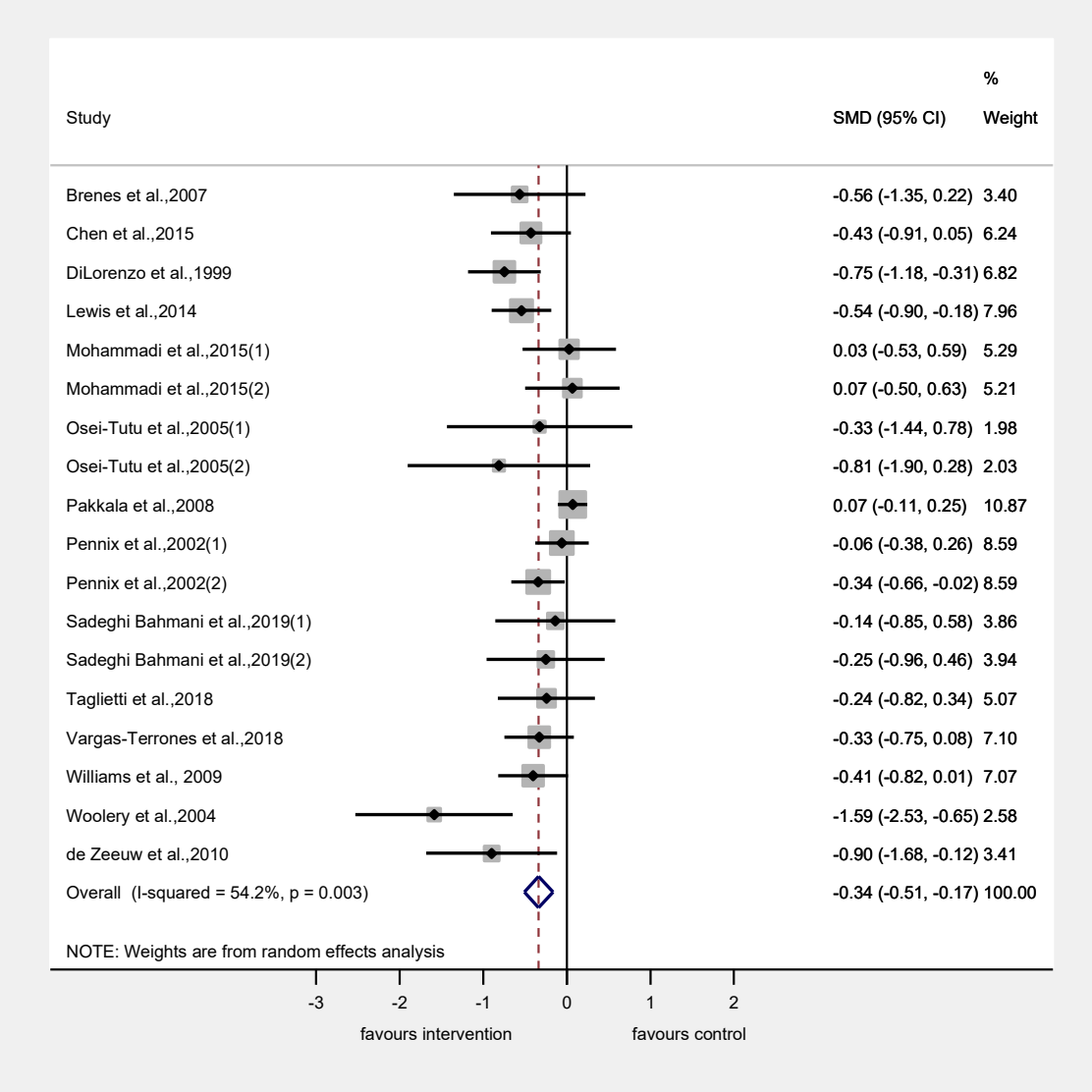


Figure 2: Forest plot



SMS: Standardized Mean Difference

SUPPLEMENT

Index

Appendix A. References of Reviewed Systematic Reviews and Meta-analyses

1. Brown HE, Gilson ND, Burton NW, Brown WJ. Does physical activity impact on presenteeism and other indicators of workplace well-being? Sports Med. 2011;41(3):249-262. doi:10.2165/11539180-000000000-00000.
2. Potter R, Ellard D, Rees K, Thorogood M. A systematic review of the effects of physical activity on physical functioning, quality of life and depression in older people with dementia. Int J Geriatr Psychiatry. 2011;26(10):1000-1011. doi:10.1002/gps.2641.
3. Das JK, Salam RA, Lassi ZS, et al. Interventions for adolescent mental health: an overview of systematic reviews. J Adolesc Health. 2016;59(4S):S49-S60. doi:10.1016/j.jadohealth.2016.06.020.
4. Wegner M, Helmich I, Machado S, Nardi AE, Arias-Carrion O, Budde H. Effects of exercise on anxiety and depression disorders: review of meta-analyses and neurobiological mechanisms. CNS Neurol Disord Drug Targets. 2014;13(6):1002-1014.
5. Rosenbaum S, Vancampfort D, Steel Z, Newby J, Ward PB, Stubbs B. Physical activity in the treatment of Post-traumatic stress disorder: a systematic review and meta-analysis. Psychiatry Res. 2015;230(2):130-136. doi:10.1016/j.psychres.2015.10.017.
6. Hall KS, Hoerster KD, Yancy WS. Post-traumatic stress disorder, physical activity, and eating behaviors. Epidemiol Rev. 2015;37:103-115. doi:10.1093/epirev/mxu011.
7. Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. Am J Prev Med. 2013;45(5):649-657. doi:10.1016/j.amepre.2013.08.001.
8. Zhai L, Zhang Y, Zhang D. Sedentary behaviour and the risk of depression: a meta-analysis. Br J Sports Med. 2015;49(11):705-709. doi:10.1136/bjsports-2014-093613.
9. Park SH, Han KS, Kang CB. Effects of exercise programs on depressive symptoms, quality of life, and self-esteem in older people: a systematic review of randomized controlled trials. Appl Nurs Res. 2014;27(4):219-226. doi:10.1016/j.apnr.2014.01.004.
10. Cooney GM, Dwan K, Greig CA, et al. Exercise for depression. Cochrane Database Syst Rev. 2013;(9):Cd004366. doi:10.1002/14651858.CD004366.pub6.
11. Cramer H, Lauche R, Langhorst J, Dobos G. Yoga for depression: a systematic review and meta-analysis. Depress Anxiety. 2013;30(11):1068-1083. doi:10.1002/da.22166.
12. Robertson R, Robertson A, Jepson R, Maxwell M. Walking for depression or depressive symptoms: a systematic review and meta-analysis. Ment Health Phys Act. 2012;5(1):66-75.
13. de Souza Moura AM, Lamego MK, Paes F, et al. Effects of aerobic exercise on anxiety disorders: a systematic review. CNS Neurol Disord Drug Targets. 2015;14(9):1184-1193. doi:10.2174/1871527315666151111121259.
14. Nystrom MB, Neely G, Hassmen P, Carlbring P. Treating major depression with physical activity: a systematic overview with recommendations. Cogn Behav Ther. 2015;44(4):341-352. doi:10.1080/16506073.2015.1015440.
15. Rebar AL, Stanton R, Geard D, Short C, Duncan MJ, Vandelanotte C. A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. Health Psychol Rev. 2015;9(3):366-378. doi:10.1080/17437199.2015.1022901.
16. Yan S, Jin Y, Oh Y, Choi Y. Effect of exercise on depression in university students: a meta-analysis of randomized controlled trials. J Sports Med Phys Fitness. 2016;56(6):811-816.
17. Josefsson T, Lindwall M, Archer T. Physical exercise intervention in depressive disorders: meta-analysis and systematic review. Scand J Med Sci Sports. 2014;24(2):259-272. doi:10.1111/sms.12050.
18. Schuch FB, Vancampfort D, Richards J, Rosenbaum S, Ward PB, Stubbs B. Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. J Psychiatr Res. 2016b;77:42-51. doi:10.1016/j.jpsychires.2016.02.023.
19. Farah, WH, Alsawas, M, Mainou, M, et al. Non-pharmacological treatment of depression: a systematic review and evidence map. Evid Based Med. 2016;21(6):214-221.
20. Mura G, Carta MG. Physical activity in depressed elderly. A systematic review. Clin Pract Epidemiol Ment Health. 2013;9:125-135. doi:10.2174/1745017901309010125.

21. Schuch FB, Vancampfort D, Rosenbaum S, et al. Exercise for depression in older adults: a meta-analysis of randomized controlled trials adjusting for publication bias. Rev Bras Psiquiatr. 2016c;38(3):247-254. doi:10.1590/1516-4446-2016-1915.
22. Rhyner KT, Watts A. Exercise and depressive symptoms in older adults: a systematic meta-analytic review. J Aging Phys Act. 2016;24(2):234-246. doi:10.1123/japa.2015-0146.
23. Lindheimer JB, O'Connor PJ, Dishman RK. Quantifying the placebo effect in psychological outcomes of exercise training: a meta-analysis of randomized trials. Sports Med. 2015;45(5):693-711. doi:10.1007/s40279-015-0303-1.
24. Cramer H, Anheyer D, Lauche R, Dobos G. A systematic review of yoga for major depressive disorder. J Affect Disord. 2017;213:70-77. doi:10.1016/j.jad.2017.02.006.
25. Bridges L, Sharma M. The efficacy of yoga as a form of treatment for depression. J Evid Based Complementary Altern Med. 2017;2156587217715927. doi:10.1177/2156587217715927.
26. Liu X, Clark J, Siskind D, et al. A systematic review and meta-analysis of the effects of Qigong and Tai Chi for depressive symptoms. Complement Ther Med. 2015;23(4):516-534. doi:10.1016/j.ctim.2015.05.001.
27. Sarris J, Moylan S, Camfield DA, et al. Complementary medicine, exercise, meditation, diet, and lifestyle modification for anxiety disorders: a review of current evidence. Evid Based Complement Alternat Med. 2012. 2012:809653. doi:10.1155/2012/809653.
28. Yin J, Dishman RK. The effect of Tai Chi and Qigong practice on depression and anxiety symptoms: a systematic review and meta-regression analysis of randomized controlled trials. Database of Abstracts of Reviews of Effects. 2014;(2):135-146.
29. Wang F, Lee Ek, Wu T, et al. The effects of Tai Chi on depression, anxiety, and psychological well-being: a systematic review and meta-analysis. Int J Behav Med. 2014;21(4):605-617.
30. Meekums B, Karkou V, Nelson EA. Dance movement therapy for depression. Cochrane Database Syst Rev. 2015;(2):CD009895. doi:10.1002/14651858.CD009895.pub2.
31. Loi SM, Dow B, Ames D, et al. Physical activity in caregivers: What are the psychological benefits? Arch Gerontol Geriatr. 2014;59(2):204-210. doi:10.1016/j.archger.2014.04.001.
32. Abraha I, Rimland JM, Trotta FM, et al. Systematic review of systematic reviews of non-pharmacological interventions to treat behavioural disturbances in older patients with dementia. The SENATOR-OnTop series. BMJ Open. 2017;7(3):e012759. doi:10.1136/bmjopen-2016-012759.
33. Barreto Pde S, Demougeot L, Pillard F, Lapeyre-Mestre M, Rolland Y. Exercise training for managing behavioral and psychological symptoms in people with dementia: A systematic review and meta-analysis. Ageing Res Rev. 2015;24(Pt B):274-285. doi:10.1016/j.arr.2015.09.001.
34. Adamson BC, Ensari I, Motl RW. Effect of exercise on depressive symptoms in adults with neurologic disorders: a systematic review and meta-analysis. Arch Phys Med Rehabil. 2015;96(7):1329-1338. doi:10.1016/j.apmr.2015.01.005.
35. Eng JJ, Reime B. Exercise for depressive symptoms in stroke patients: a systematic review and meta-analysis. Clin Rehabil. 2014;28(8):731-739. doi:10.1177/0269215514523631.
36. Radovic S, Gordon MS, Melvin GA. Should we recommend exercise to adolescents with depressive symptoms? A meta-analysis. J Paediatr Child Health. 2017;53(3):214-220. doi:10.1111/jpc.13426.
37. Brown H, Pearson N, Braithwaite R, Brown W, Biddle S. Physical activity interventions and depression in children and adolescents: a systematic review and meta-analysis. Sports Med. 2013;43:195-206. doi:10.1007/s40279-012-0015-8.
38. Hoare E, Skouteris H, Fuller-Tyszkiewicz M, Millar L, Allender S. Associations between obesogenic risk factors and depression among adolescents: a systematic review. Obes Rev. 2014;15(1):40-51. doi:10.1111/obr.12069.
39. Hoare E, Milton K, Foster C, Allender S. The associations between sedentary behaviour and mental health among adolescents: a systematic review. Int J Behav Nutr Phys Act. 2016;13(1):108. doi:https://doi.org/10.1186/s12966-016-0432-4.
40. Korczak DJ, Madigan S, Colasanto M. Children's physical activity and depression: a meta-analysis. Pediatrics. 2017;139(4):1-14.

41. Carter T, Morres ID, Meade O, Callaghan P. The effect of exercise on depressive symptoms in adolescents: a systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry. 2016;55(7):580-590. doi:10.1016/j.jaac.2016.04.016.
42. Schuch FB, Deslandes AC, Stubbs B, Gosmann NP, Silva CT, Fleck MP. Neurobiological effects of exercise on major depressive disorder: a systematic review. Neurosci Biobehav Rev. 2016a;61:1-11. doi:10.1016/j.neubiorev.2015.11.012.
43. Pedersen BK, Saltin B. Exercise as medicine – evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015: (Suppl. 3) 25: 1–72.
44. Knapen J, Vancampfort D, Moriën Y, Marchal Y. Exercise therapy improves both mental and physical health in patients with major depression. Disabil Rehabil. 2015;37(16):1490-5.
45. Stubbs B, Vancampfort D, Rosenbaum S, Ward PB, Richards J, Ussher M, Schuch FB. Challenges Establishing the Efficacy of Exercise as an Antidepressant Treatment: A Systematic Review and Meta-Analysis of Control Group Responses in Exercise Randomised Controlled Trials. Sports Med. 2016;46(5):699-713.
46. Herring MP, Puetz TW, O'Connor PJ, Dishman RK. Effect of exercise training on depressive symptoms among patients with a chronic illness: a systematic review and meta-analysis of randomized controlled trials. Arch Intern Med. 2012;172(2):101-11.
47. Conn VS. Depressive symptom outcomes of physical activity interventions: meta-analysis findings. Ann Behav Med. 2010;39(2):128-38.
48. Ekeland E, Heian F, Hagen KB. Can exercise improve self-esteem in children and young people? A systematic review of randomised controlled trials. Br J Sports Med 2005;39(11):792-8.
49. Carter T, Bastounis A, Guo B, Morrell CJ. The effectiveness of exercise-based interventions for preventing or treating postpartum depression: a systematic review and meta-analysis. Archives of Women's Mental Health 2019;22(1): 37-53
50. Chan JSY, Liu G, Liang D, Deng K, Wu J, Yan JH. Special Issue- Therapeutic Benefits of Physical Activity for Mood: A systematic review on the Effects of Exercise Intensity, Duration and Modality. The Journal of Psychology 2019;153(1):102-125.
51. Nakamura A, van der Waerden J, Melchior M, Bolze C, El-Khoury F, Pryor L. Physical activity during pregnancy and postpartum depression: A systematic review and meta-analysis. Journal of Affective Disorders 2019;246:29-41
52. Pascoe M C, Parker AG. Physical activity and exercise as an universal depression prevention in young people: a narrative review. Early Interv Psychiatry. 2019;13(4):733-739
53. Davenport MH, McCurdy AP, Mottola MF, Skow RJ, Meah VL, Poitras VJ, Jaramillo Garcia A, Gray CE, Barrowman N, Riske L, Sobierajski F, James M, Nagpal T, Marchand AA, Nuspl M, Slater LG, Barakat R, Adamo KB, Davies GA, Ruchat SM. Impact of prenatal exercise on both prenatal and postnatal anxiety and depressive symptoms: a systematic review and meta-analysis.Br J Sports Med. 2018 Nov;52(21):1376-1385
54. O'Connor E, Senger CA, Henninger ML, Coppola E, Gaynes BN. Interventions to Prevent Perinatal Depression: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2019;321(6):588-601.
55. Gordon BR, McDowell CP, Hallgren M, Meyer JD, Lyons M, HerringMP. Association of efficacy of resistance exercise training with depressive symptoms. JAMA Psychiatry. 2018;75(6):566-576.
56. Poyatos-León R, García-Hermoso A, Sanabria-Martínez G, Álvarez-Bueno C, Cavero-Redondo I, Martínez-Vizcaíno V. Effects of exercise-based interventions on postpartum depression: A meta-analysis of randomized controlled trials. Birth. 2017 Sep;44(3):200-208.

Appendix B. Search Strategies

PubMed (MEDLINE, CINHAL): 2708 (25/05/2020)

("depressive disorder" OR "major depressive disorder" OR depress* OR "depression") AND ("physical activity" OR "exercise" OR "fitness" OR "Sport" OR "leisure activities") AND (prevent* OR incidence) AND ("effectiveness" OR "trial" OR "controlled trial" OR "randomi*" OR "intervention" OR "efficacy")

The Cochrane Central Register of Controlled Trials (CENTRAL): 1186 (25/05/2020)

("depression":ti,ab,kw or "depressive disorder":ti,ab,kw or "major depressive disorder":ti,ab,kw or depress*:ti,ab,kw) AND (physical activity:ti,ab,kw or exercise:ti,ab,kw or fitness:ti,ab,kw or Sport:ti,ab,kw or leisure activities:ti,ab,kw) AND (prevent*:ti,ab,kw or incidence:ti,ab,kw) AND (effectiveness:ti,ab,kw or trial:ti,ab,kw or controlled trial:ti,ab,kw or randomi*:ti,ab,kw or intervention:ti,ab,kw or efficacy:ti,ab,kw)

Embase: 777 (25/05/2020)

('depression'/exp OR 'depression' OR 'depressive disorder'/exp OR 'depressive disorder' OR 'major depressive disorder'/exp OR 'major depressive disorder' OR 'depress*') AND ('physical activity'/exp OR 'physical activity' OR 'exercise'/exp OR 'exercise' OR 'fitness'/exp OR 'fitness' OR 'Sport'/exp OR 'sport' OR leisure activities/exp OR 'leisure activities') AND (prevent* OR incidence) AND ('effectiveness' OR 'trial' OR 'controlled trial' OR 'randomi*' OR 'intervention' OR 'efficacy')

Web of Science: 4106 (25/05/2020)

TS= ((depressive disorder OR (depressive AND disorder) OR depression OR major depressive disorder)) AND TS= ((physical activity OR exercise OR fitness OR Sport OR leisure activities)) AND TS= ((prevent* OR incidence)) AND TS= ((effectiveness OR trial OR (controlled AND trial) OR randomi* OR intervention OR efficacy))

PsycINFO: 1788 (25/05/2020)

(TX depression OR TX depressive disorder OR (TX depressive AND TX disorder) OR TX depress* OR TX major depressive disorder) AND (TX physical activity OR TX exercise OR TX fitness OR TX Sport OR TX leisure activities) AND (prevent* OR incidence) AND (effectiveness OR trial OR (controlled AND trial) OR randomi* OR intervention OR efficacy)

Open Grey: 8 (25/05/2020)

(depression OR depressive disorder OR major depressive disorder OR depress*) AND (prevent*) AND (physical activity OR exercise OR fitness OR Sport OR leisure activities)

SPORTdiscus: **202 (25/05/2020)**

("depressive disorder" OR " major depressive disorder" OR depress* OR "depression") AND ("physical activity" OR "exercise" OR "fitness" OR "Sport" OR "leisure activities") AND (prevent* OR incidence) AND ("effectiveness" OR "trial" OR "controlled trial" OR "randomi*" OR "intervention" OR "efficacy")

eTable 1. Characteristics of Included Randomized Controlled Trials of Depression Prevention

Author / Year / Country	Target population / Type of prevention [a]	Depression exclusion at baseline	Sample (intervention / control)	Conditions intervention – control	Characteristics of exercise-based intervention: a) intervention duration b) session duration c) format d) frequency e) intensity f) type g) supervision	Verification of physical activity (objective / subjective)	Type of outcome on depression (primary / secondary)	Follow-up	Depression Outcomes (standardized interview / symptoms scale)
· **Brenes et al.** · **2007** · **United States**	· Older adults · Mean age: 73.7; standard deviation: 6.8 · Indicated	· No major depression DSM-IV (PHQ-9 algorithm)	26 (14/12)	**1**. Aerobic and resistance training[d] **2**. Usual care	a) 16 weeks b) 60 minutes c) Group d) 3 sessions/week e) Moderate f) Warm up, aerobic phase, resistance training phase, aerobic phase, cool down g) Supervised	Objective: distance walked in 6 minutes, time to walk 4 meters, and time to sit and stand from a chair 5 times	Primary	4 months	· Symptoms scale (HADS-D and GDS)
· **Chen et al.** · **2015** · **Taiwan**	· Patients with lung cancer · 37–88 years (mean age: 64.16; standard deviation: 10.89) · Selective	· No depression (HADS-D ≥8)	84 (40/44)[c]	**1**.Walking **2.** Usual care	a) 12 weeks b) 40 minutes c) Individually d) 3 sessions/week e) Moderate f) Walking g) Unsupervised	Objective: heart rate monitor + training diary Subjective: Borg Rating of Perceived Exertion scale	Primary	6 months	· Symptoms scale (HADS-D)

· **de Zeeuw et al.** · **2010** · **Netherlands**	· Employees of a company · Mean age: 41.2 years · Indicated	· No depression (PHQ-9 ≥10)	30 (15/15)	**1**. Fitness program **2**. Usual care	a) 10 weeks b) 50 minutes c) Group d) 2 sessions/week e) Moderate f) Cardiovascular and stretching exercises/ training/cycling/jogging/ walking/climbing/sit-ups/relaxation g) Supervised	Objective: heart rate monitored	Primary	2.5 months	· Symptoms scale (PHQ-9)
· **DiLorenzo et al.** · **1999** · **Colombia**	· Healthy adults · 18-39 years · Mean age: 31.5 · Universal	· No depression (BDI >19)	111 (82/29)	**1**. Aerobic fitness **2**. No intervention	a) 12 weeks b) 24 minutes c) Group d) 4 sessions/week e) Moderate f) Bicycle g) Supervised	Objective: heart rate monitored and bicycle ergometer test (not validated)	Primary	12 months	· Symptoms scale (BDI)
· **Lewis et al.** · **2014** · **United States**	· Healthy postpartum women · Mean age: 31.54; standard deviation: 4.95 · Selective	· No major depression DSM -IV (SCID)	130 (66/64)	**1.** Exercise intervention **2**. General wellness topics support contact by telephone	a) 24 weeks b) 30 minutes c) Individually d) 5 sessions/week+11 telephone sessions e) Moderate to vigorous f) Varied exercises (types of exercise they preferred) g) Unsupervised	Objective: device that measures movement and intensity of physical activity (ActiGraph) Subjective: 7-Day Physical Activity Recall Interview	Primary	6 months	· Incidence (SCID-I) · Symptoms scale (PHQ-9; EPDS)

· **Mohammadi et al.** · **2015** · **Iran**	· Pregnant women at 26-32 weeks of pregnancy · 25.3 years · Selective	· No depression (EPDS ≥15)[b]	127 (43/42/42)	**1**. Home-based antenatal exercise **2**. Home-based antenatal plus postnatal exercise **3**. Usual care	a) 11 weeks (antenatal) and 15 weeks (antenatal + postnatal) b) 25 minutes c) Individually d) 3 sessions/week e) Low f) Stretching and breathing practices g) Unsupervised	Subjective: telephone calls	Primary	5 months	· Symptoms scale (EPDS)
· **Osei-tutu et al.** · **2005** · **Canada**	· Healthy sedentary volunteers · 20-40 years · Mean age: 34 · Universal	· No depression (POMS)	40 (15/15/10)	**1**. 30 min walking **2**. 3 x 10 min walking with minimum 2h rest intervals **3**. Non-exercise control	a) 8 weeks b) 30 minutes c) Individually d) 5 sessions/week e) Moderate f) Walking g) Unsupervised	Objective: heart rate monitored	Primary	2 months	· Symptoms scale (POMS)
· **Pakkala et al.** · **2008** · **Finland**	· Elderly · 75-81 years · Mean age: 77.6 · Selective	· No depression · Sub-sample excluding people with CES-D≥16	501 (253/248)	1. Physical activity counseling 2. Usual care	a) Not specified b) 60 minutes c) Individually d) 7 sessions/week e) Low f) Home calisthenics exercised, walking, performing every day activities and recommendation to do inexpensive exercise classes organized by the municipality g) Unsupervised	Subjective: standardized physical activity questionnaire [Grimby, 1986]	Primary	24 months	· Symptoms scale (CES-D)

· **Penninx et al.** · **2002** · **United States**	· Adults with knee osteoarthritis · 68.8 years (standard deviation: 5.6) · Selective	· No depression · Sub-sample excluding people with CES-D short form ≥5	340 (115/112/113)	**1.** Aerobic exercise program **2.** Resistance exercise program **3.** Health education related to arthritis management	a) 72 weeks b) 60 minutes c) Group d) 3 sessions/week e) Moderate f) Walking g) Supervised ---------- a) 72 weeks b) 60 minutes c) Group d) 3 sessions/week e) Moderate f) Upper and lower body exercises using dumbbells and cuff weights g) Supervised	Objective: 6-minute walking speed test	Primary	18 months	· Symptoms scale (CES-D)
· **Sadeghi-Bahmani et al.** · **2019** · **Iran**	· Women with multiple sclerosis · 18-65 years (mean age: 37.36 years) · Selective	· No major depressive disorders (MINI)	83 (27/30/26)	**1**. Coordinative training **2**. Endurance training **3.** Active control condition	a) 8 weeks b) 30-45 minutes c) Group d) 3 sessions/week e) Moderate f) Warming up, coordinative training (e.g. balancing on small bar, balancing balls, mirroring and imitating instructors' movements), cooling down g) Supervised ---------- a) 8 weeks b) 30-45 minutes c) Group d) 3 sessions/week e) Moderate f) Warming-up, stretching, exercises on treadmill/ bicycles/walking/jogging, cooling down	Subjective	Primary	2 months	· Symptoms scale (BDI-FS)

					g) Supervised				
· **Taglietti et al.** · **2018** · **Brazil**	· Adults with knee osteoarthritis · 68.3 years (standard deviation: 4.8) · Selective	· No depression · Sub-sample excluding people with GDS ≥ 6	47 (27/20)[c]	**1.** Aquatic mixed exercise **2.** Educational program ('how to deal with chronic pain')	a) 8 weeks b) 60 minutes c) Individually d) 2 sessions/week e) Low-moderate f) Warm-up, stretching the leg muscles, knee and hip isometric and dynamic exercises with elastic bands, aerobic exercises, step training and proprioceptive exercises, cool down with massage and relaxation g) Supervised	Subjective	Secondary	3 months	· Symptoms scale (GDS)
· **Vargas-Terrones et al.** · **2018** · **Spain**	· Pregnant women <16 weeks pregnant · 32.8 years (standard deviation: 3.95) · Selective	· No depression · Sub-sample excluding people with CES-D ≥16	100 (56/44)[c]	1. Specific mixed exercise program 2. Usual care	a) 29 weeks b) 60 minutes c) Group d) 3 sessions/week e) Moderate (intensity of 55%- 60% of heart rate reserve) f) Warm-up, aerobic exercise, muscle strengthening exercises, coordination and balance, pelvic floor exercises, stretching and relaxation. g) Supervised	Objective: Polar FT7 heart rate monitor Subjective: Borg Rating of Perceived Exertion scale	Primary	8.5 months	· Symptoms scale (CES-D)

· **Williams et al.** · **2009** **United States**	· Adults with chronic low back pain · 18-70 years · Mean age:48 Selective	No depression (BDI-II ≥20)	90 (43/47)	**1**. Iyengar yoga **2**. Usual care	a) 24 weeks b) 90 min classes and 30 minutes at home c) Group (classes) and individually (at home) d) Classes: 2 sessions/week; at home: 3 sessions /week e) Low f) Yoga exercises g) Supervised	Subjective: reports on duration and frequency of the home practice	Primary	12 months	· Symptoms scale (BDI-II)
· **Woolery et al.** · **2004** **United States**	· Adults · 18-29 years · Mean age 21.5 (sd: 3.23) Indicated	No depression (BDI >15)	28 (13/15)	1. Iyengar yoga 2. Waiting list	a) 5 weeks b) 60 minutes c) Group d) 2 sessions/week e) Low f) Yoga g) Supervised	Subjective	Primary	1,25 months	· Symptoms scale (BDI)

BDI= Beck Depression Inventory; BDI-II: Beck Depression Inventory version 2; BDI-FS: Beck Depression Inventory Fast Screen; DSM-IV= Diagnostic and Statistical Manual of Mental Disorders; SCID= Structured Clinical Interview for DSM; MINI= Mini-International Neuropsychiatric Interview; PHQ-9= Patient Health Questionnaire-9; EPDS= Edinburgh Postnatal Depression Scale; CES-D= Center for Epidemiologic Studies of Depression; HADS= Hospital Anxiety and Depression Scale; GDS= Geriatric Depression Scale; POMS= Profile of Mood States.

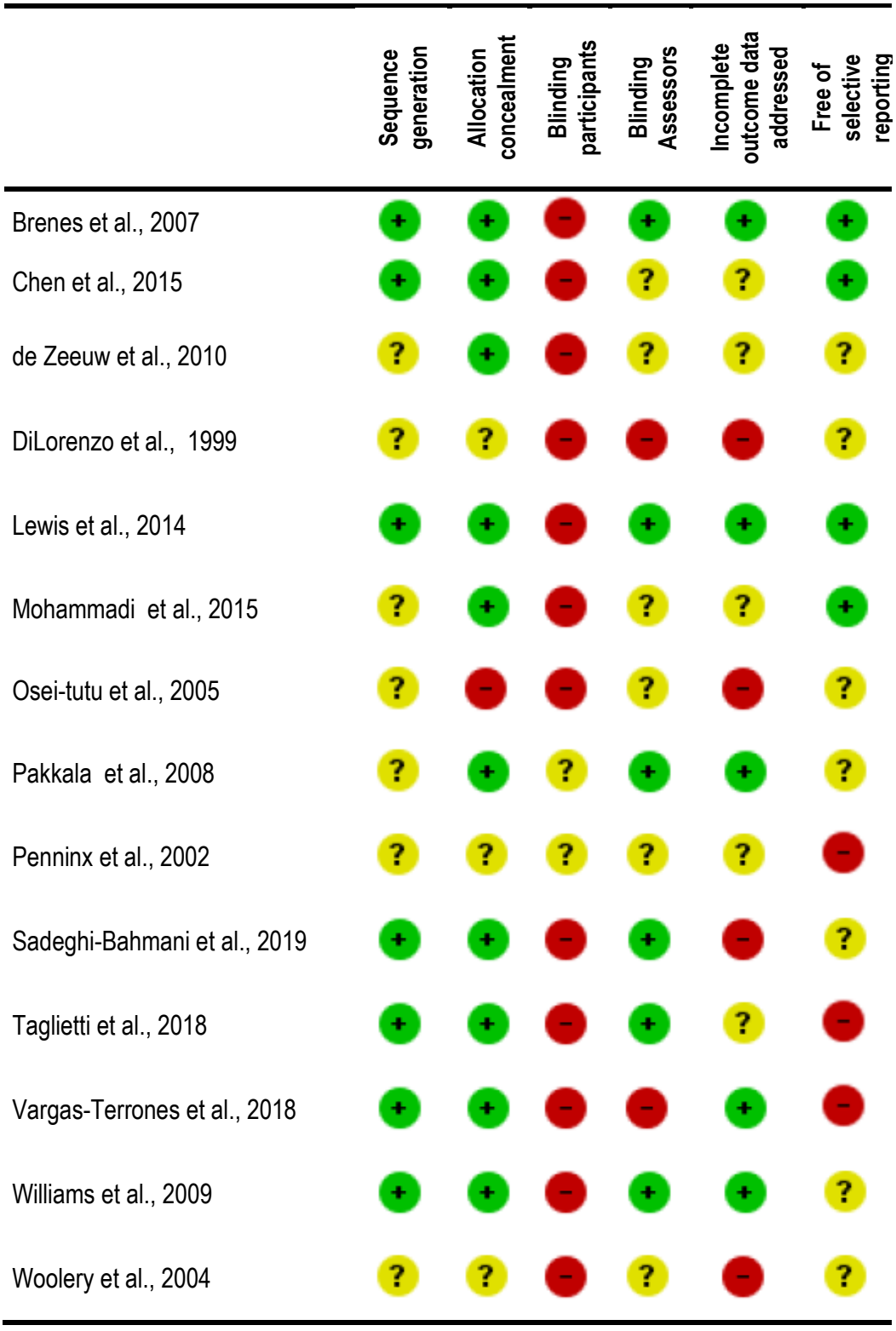
[a] Type of prevention: Indicated: patients with subthreshold depression; Selective: patients with a risk factor for depression; Universal: general population.

[b] Validated cut-off antenatal depression 15 or more (Matthey, S., Henshaw, C., Elliott, S., Barnett, B. Variability in use of cut-off scores and formats on the Edinburgh Postnatal Depression Scale – implications for clinical and research practice. Arch Womens Ment Health (2006) 9: 309–315)

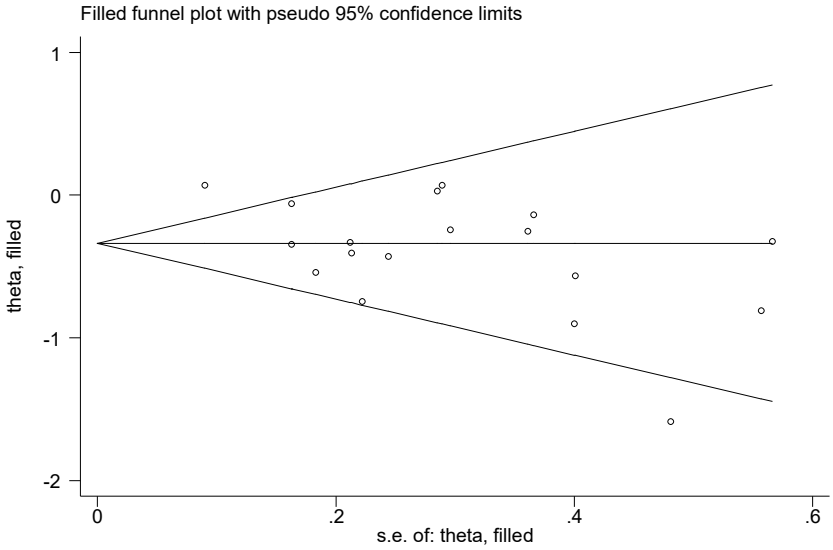
[c] Data on non-depressed people were provided by the authors.

[d] Medication group was excluded in this meta-analysis.

eTable 2. Risk of Bias



eFigure 1. Funnel Plot



eTable 3: Subgroup Analysis

Subgroup analyses	N	SMD	95% CI	*P* [a]	I^2	Between-group heterogeneity [b]
Participant characteristics						
Country						
North America	8	-0.434	-0.673 to -0.196	<0.001	41%	Q=13.09; d.f.(Q)=3;
Europe	3	-0.270	-0.746 to 0.205	0.228	75%	*P*=0.004
Asia	5	-0.161	-0.422 to 0.101	0.001	0%	
South America	2	-0.533	-1.019 to -0.047	0.032	46%	
Sex						
Only women	4	-0.395	-0.634 to -0.156	0.001	0%	Q=2.40; d.f.(Q)=1;
Women and men	14	-0.342	-0.551 to -0.133	0.001	61%	*P*=0.121
Age						
Elderly	6	-0.172	-0.378 to 0.034	0.102	46%	Q=11.14; d.f.(Q)=1;
Adult	12	-0.436	-0.645 to -0.227	<0.001	34%	P=0.001
Chronic disease						
Yes	7	-0.264	-0.429 to -0.098	0.002	0%	Q=0.37; d.f.(Q)=1
No	11	-0.427	-0.716 to -0.138	0.004	71%	0.546
Type of prevention						
Universal	3	-0.705	-1.084 to -0.325	<0.001	0%	Q=16.97; d.f.(Q)=2;
Selective	12	-0.203	-0.350 to 0.057	0.007	35%	*P*<0.001
Indicated	3	-0.968	-1.527 to -0.409	0.001	26%	
Intervention characteristics						
Type of exercise						
Aerobic	8	-0.387	-0.680 to -0.094	0.010	71%	Q=1.74; d.f.(Q)=3;
Strength	1	-0.344	-0.663 to -0.025	0.035	n.a.	*P*=0.629
Stretching	4	-0.373	-0.918 to -0.171	0.179	71%	
Mixed	5	-0.302	-0.566 to -0.038	0.025	0%	
Walking						
Yes	3	-0.470	-0.877 to -0.062	0.024	0%	Q=1.50; d.f.(Q)=1;
No	15	-0.323	-0.510 to -0.137	0.001	60%	*P*=0.221
Yoga						
Yes	2	-0.918	-2.068 to 0.232	0.118	80%	Q=4.00; d.f.(Q)=1;
No	16	-0.288	-0.452 to -0.124	0.001	47%	*P*=0.045
Type of verification						
Objective	10	-0.412	-0.568 to -0.256	<0.001	10%	Q=11.38; d.f.(Q)=1;
Subjective	8	-0.200	-0.471 to 0.070	0.146	56%	*P*=0.001
Supervised exercise						
Yes	11	-0.410	-0.604 to -0.216	<0.001	36%	Q=7.89; d.f.(Q)=1;
No	7	-0.204	-0.479 to 0.070	0.144	56%	*P*=0.005
Format						
Individually	8	-0.203	-0.446 to 0.039	0.100	49%	Q= 7.82; d.f.(Q)=1;
Group	10	-0.431	-0.642 to -0.219	<0.001	42%	*P*=0.005
Duration of intervention						
Up to 12 weeks	10	-0.480	-0.740 to -0.221	<0.001	33%	Q=7.64; d.f.(Q)=1;
>12 weeks	8	-0.227	-0.424 to -0.031	0.023	57%	*P*=0.006
Frequency of sessions						
2-4 sessions / week	13	-0.349	-0.536 to -0.162	<0.001	39%	Q=4.34; d.f.(Q)=1;
5-7 sessions / week	5	-0.316	-0.687 to 0.055	0.095	70%	*P*=0.037
Duration of sessions						
<60 minutes	10	-0.411	-0.585 to -0.238	<0.001	7%	Q=8.25 d.f.(Q)=1;
≥60 minutes	8	-0.277	-0.543 to -0.012	0.041	67%	*P*=0.004
Volume						
Up to 150 minutes/week	12	-0.437	-0.658 to -0.215	<0.001	35%	Q=9.79; d.f.(Q)=1;
>150 minutes/week	6	-0.195	-0.400 to 0.011	0.063	52%	*P*=0.002

Intensity						
Low	6	-0.221	-0.557 to 0.115	0.198	68%	Q=11.08; d.f.(Q)=1;
Moderate	12	-0.387	-0.526 to -0.248	<0.001	0%	*P*=0.001

Methodological characteristics

Subgroup analyses	N	SMD	95% CI	*P* [a]	I^2	Between-group heterogeneity [b]
Depression exclusion at baseline						
Symptom scale	15	-0.336	-0.529 to -0.143	0.001	57%	Q=2.10; d.f.(Q)=1;
Standardized diagnostic interview	3	-0.426	-0.718 to -0.134	0.004	0%	*P*=0.147
Outcome measure						
Symptom scale	17	-0.318	-0.494 to -0.141	<0.001	53%	Q=3.32; d.f.(Q)=1;
Standardized diagnostic interview	1	-0.543	-0.902 to -0.184	0.003	n.a.	*P*=0.069
Type of outcome						
Primary	17	-0.344	-0.523 to -0.165	<0.001	57%	Q=0.01; d.f.(Q)=1;
Secondary	1	-0.244	-0.824 to 0.336	0.410	n.a.	p=0.943
Comparator						
Usual care	11	-0.327	-0.562 to 0.093	0.006	58%	Q=9.27; d.f.(Q)=2;
Active control	6	-0.282	-0.453 to -0.111	0.001	0%	*P*=0.010
Waiting list	1	-1.589	-2.532 to -0.646	<0.001	n.a.	
Sample size						
<100	12	-0.372	-0.587 to -0.156	0.001	22%	Q=16.28; d.f.(Q)=2;
100-200	3	-0.535	-0.765 to -0.304	<0.001	0%	*P*<0.001
>200	3	-0.084	-0.328 to 0.159	0.497	60%	
Subsample						
No	13	-0.454	-0.651 to -0.258	<0.001	28%	Q=13.47 d.f.(Q)=1;
Yes [c]	5	-0.132	-0.326 to 0.062	0.184	44%	P<0.001
Risk of bias (qualitative)						
Low	2	-0.547	-0.873 to 0.220	0.001	0%	Q=4.17; d.f.(Q)=1;
Moderate-high	16	-0.310	-0.491 to -0.128	0.001	54%	*P*=0.041
Risk of bias (quantitative) (range 0-12)						
Low (scored 0-3)	5	-0.290	-0.613 to 0.033	0.079	69%	Q= 6.53; d.f.(Q)=2;
Moderate (scored 4-5)	5	-0.161	-0.422 to 0.101	0.228	0%	*P*=0.038
High (scored ≥6)	8	-0.508	-0.792 to -0.225	<0.001	54%	
Follow up						
<6 months	10	-0.377	-0.667 to -0.086	0.011	35%	Q=8.34; d.f.(Q)=2;
6 and <12 months	4	-0.437	-0.643 to -0.231	<0.001	0%	*P*=0.015
12-24 months	4	-0.233	-0.558 to 0.092	0.160	79%	

Abbreviations: ***N****: number of comparisons;* ***SMD****: standardized mean difference*

[a] *Significance tests in which for each subgroup the null hypothesis is that SMD=0*

[b] *Q values represent the comparison of subgroup means based on a chi-square distribution in which the null hypothesis is that the effect size is the same for all subgroups*

[c] *Studies that include depressed and non-depressed participants at baseline but give separate outcomes for non-depressed*

d.f.*= degree of freedom*

n.a.*= not applicable.*

eFigure 2. Normal Probability Plot of Standardized Shrunken Residuals of the final meta-regression model