Exercise Intervention for Late-Life Depression: A Meta-Analysis

Sivan Klil-Drori, MD[a,*]; Adi J. Klil-Drori, MD, MSc[b]; Shamira Pira, MD, MSc[c]; and Soham Rej, MD, MSc[c]

ABSTRACT

Objective: To quantify the association between physical exercise intervention (PEI) and reduction in depressive symptoms in older adults.

Data Sources: MEDLINE, PsycINFO, and EMBASE were searched from inception through December 2018 with no language restrictions using keywords related to exercise, depression, elderly adults, and randomized controlled trials.

Study Selection: Randomized controlled trials comparing a sedentary control group, with no physically active intervention, to a supervised, moderate-to-vigorous PEI with participants aged ≥ 60 years and having a primary outcome of depressive symptoms were included.

Data Extraction: Data on pre- and post-intervention scores on scales measuring depressive symptoms were extracted using a standard form. Random-effects models were used to pool standardized mean differences (Hedges *g*) in depressive symptoms across studies.

Data Synthesis: Nine studies involving 1,308 participants were included; mean participant age was 82 years. Moderate-to-vigorous PEI was associated with a medium effect size of 0.64 (95% CI, 0.27 to 1.01; $z = 3.38$; $P < .001$) in reducing depressive symptoms. However, there was considerable heterogeneity ($T^2 = 0.22$, $Q = 36.34$, $P < .0001$; $I^2 = 78.0\%$) in the effect of PEI across included studies. Age > 80 years, Mini-Mental State Examination (MMSE) score < 23, and no depressive symptoms at baseline contributed to heterogeneity. Fitness metrics and adherence to exercise were inconsistently reported, and 5 of 9 studies were deemed at high risk of bias.

Conclusions: A moderate reduction in depressive symptoms was seen with PEI among older adults. Nevertheless, more work is needed to support PEI for late-life depression in adults over age 80 years or with MMSE scores < 23 suggestive of cognitive decline.

J Clin Psychiatry 2020;81(1):19r12877

To cite: Klil-Drori S, Klil-Drori AJ, Pira S, et al. Exercise intervention for late-life depression: a meta-analysis. *J Clin Psychiatry*. 2020;81(1):19r12877.

To share: https://doi.org/10.4088/JCP.19r12877

[a]Rotman Research Institute, Baycrest Health Sciences, University of Toronto, Toronto, Ontario, Canada

[b]Sunnybrook Health Sciences, University of Toronto, Toronto, Ontario, Canada

[c]Geri-PARTy Research Group and McGill Meditation and Mind-body Medicine Research Clinic (MMMM-RC), Department of Psychiatry, Jewish General Hospital, McGill University, Montreal, Quebec, Canada

**Corresponding author:* Sivan Klil-Drori, MD, Department of Neurology, Rotman Research Institute, Baycrest Health Sciences Centre, Toronto, ON, Canada (sklil-drori@research.baycrest.org).

Depression has a considerable societal toll: the World Health Organization estimated that 7.5% of years lived with disability are caused by depression.[1] Depressive symptoms are manifest in 27% of the adult population, and although they do not follow a clear pattern across age groups, they are most common in adults aged 80 years or older.[2] Nevertheless, although late-life depression (LLD; ie, depression in older adults) is widespread, it is often overlooked or inadequately treated.[3] Current mainstream treatment options for LLD have important limitations. Although antidepressants are superior to placebo for the acute treatment phase, the response rate to these medications in individuals with LLD is lower than in younger-aged individuals, and relapse rate is higher.[4] Individual psychotherapy (eg, one-on-one cognitive-behavioral therapy) has also been shown to be effective for LLD but is costly and may not be readily accessible.[5]

Physical exercise intervention (PEI) is a promising treatment modality for LLD. It can help with physical comorbidities often seen in older age, with benefits extending to important domains of physical function such as the risk of falling and activities of daily living.[6] However, although physical activity has been recommended for older adults since at least 2008,[7] no decrease in LLD rates was noted over this period.[8] Thus, assessing the strength of the evidence supporting PEI specifically for LLD and identifying possible knowledge gaps may better the care of LLD. We identified several meta-analyses[9–15] that estimated the effect of psychosocial interventions on depression in older adults. A major limitation of all of these reviews is the inclusion of studies in which the intervention was mind-body exercise (eg, yoga, tai chi, qigong) or educational sessions (eg, on depression or diet)—all of which are of light absolute intensity—alongside studies of higher-intensity exercise interventions. Importantly, current recommendations for older adults are for at least 150 minutes per week of moderate- to vigorous-intensity aerobic activity.[7] Therefore, the isolated effect of guideline-concordant PEI on depressive symptoms remains uncertain. Further, grouping several types of interventions limits the discussion on the role of specific components of PEI that define its dose, such as intensity, duration, frequency, and target population.

The key objectives of this systematic review and meta-analysis were to synthesize the evidence from randomized controlled trials (RCTs) that assessed PEI to improve depressive symptoms in patients aged 60 years or older and to indicate possible gaps in the literature regarding this evidence.

Clinical Points

- Evidence for the effect of exercise on depressive symptoms in older adults is scant and heterogeneous. This meta-analysis focused on trials with a specific exercise dose and assessed sources of heterogeneity.
- The meta-analysis confirms a moderate effect for this narrow dose definition of exercise. Nevertheless, age over 80 years and cognitive impairment at baseline may reduce the effectiveness of exercise.

METHODS

Data Sources and Searches

This systematic review was conducted in accordance with the PRISMA guidelines.[16] Details of the protocol were registered on PROSPERO (available at http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018103292). Studies published through December 2018 were identified through EMBASE RCTs (1947–2018), PubMed MEDLINE (1946–2018), PubMed MEDLINE RCTs (1946–2018), and PsycINFO (1967–2018) using the Ovid search engine. There were no language restrictions. The search strategy was as follows: *exercise* OR *cardio* OR *aerobic* OR *anaerobic* OR *resistance* OR *physical activity* AND *depression* OR *depressive disorder* AND *aged* OR *geriatrics* OR *older adult* or *elderly* OR *late life* or *gerontology* OR *old people* OR *older people* AND *randomized control trial* OR *controlled clinical trial* OR *randomized* OR *placebo* OR *drug therapy* OR *randomly* OR *trial* OR *groups* and NOT *animals/not humans*. Supplementary searches within previous systematic reviews[9,10,12,14,15,17] and references of included studies were conducted manually.

Study Selection and Inclusion Criteria

We included published clinical trials with randomized allocation to moderate-to-vigorous PEI and a nonactive control condition, ie, not participating in a physically active intervention. These studies were required to use a validated measure of depressive symptoms as the primary outcome,[18] which had been assessed both at baseline and during the post-intervention period. Reduction of depressive symptoms with exercise was shown to be higher in studies whose primary outcome was depression compared with studies that assessed depressive symptoms as a secondary outcome.[17] Patients in included studies had to either (1) be aged exclusively 60 years or older or (2) have a mean age of 65 years or older. For reliable reporting of adherence to protocol, the exercise intervention had to be supervised by study personnel such as an exercise trainer (as opposed to an "exercise prescription" without observation of patients performing exercise) and be of either aerobic (eg, running, jogging, cycling, swimming) or nonaerobic (eg, strength, core, resistance) exercise types. Resistance and balance training are part of the multicomponent exercise recommendation[7] for older adults and were thus permitted. Requirements for the exercise intervention in included studies were modified from the American[7] and World Health Organization (WHO)[19] guidelines to meet clinical trials limitations. The following components were minimum requirements: (1) absolute intensity of moderate to vigorous, (2) frequency ≥ 2 sessions per week, (3) duration ≥ 20 minutes per session, and (4) an overall length of ≥ 8 weeks. Studies with a population with an important comorbid condition in all participants (eg, only patients with Alzheimer's disease, cardiovascular disease, specific types of cancer), with an intervention that was neither aerobic nor anaerobic (eg, yoga, tai chi, qigong), with an unsupervised exercise intervention, or with an unclear dose of exercise were excluded. Three independent reviewers (S.K.-D., A.J.K.-D., S.P.) completed abstract screening. Full-text screening was performed by 2 investigators (S.K.-D., A.J.K.-D.), and disagreements were resolved by consensus. Full-text studies were excluded due to wrong exercise intervention (type, frequency, or intensity not in accordance with our criteria modified from WHO recommendations; 3 studies), participants aged < 60 years (2 studies), and wrong primary outcome, ie, not depressive symptoms (8 studies). Figure 1 depicts the study selection process.

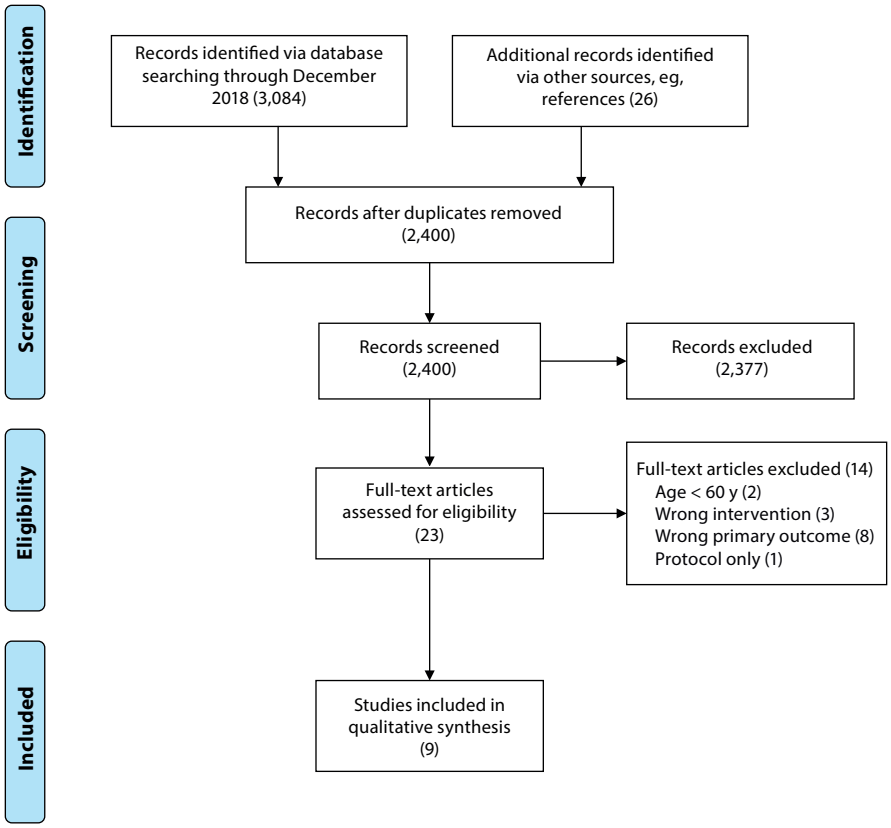
Data Extraction

Data from included studies were extracted into a standard form. We calculated mean age for study participants and noted place of residence (community or nursing home), sex, cognitive function using the Mini-Mental State Examination (MMSE),[20] and the metric used to evaluate depressive symptoms. Finally, we calculated the proportion of study participants with depression at baseline, defined as a Geriatric Depression Scale (GDS)[21] score ≥ 5 (on the 15-item GDS [GDS-15]) or ≥ 10 (on the 30-item GDS [GDS-30]). Alternatively, a Beck Depression Inventory (BDI) score[22] ≥ 14 or Hamilton Depression Rating Scale (HDRS)[23] score ≥ 17 were used. These values represent at least minor depression. When these exact scores were not used as inclusion criteria for the study, we standardized the baseline scores and used *z*-score distribution to compute the percentage of subjects above the cutoff score. Missing information was sought from study investigators.

Risk-of-Bias Assessment

Two reviewers (S.K.-D., A.J.K.-D.) independently assessed each trial for risk of bias using 7 domains: random sequence allocation, allocation concealment, blinding of participants/personnel, blinding of assessors, incomplete outcome data, selective reporting, and other; analyses were done using Review Manager (RevMan) Version 5.3 (2014; The Nordic Cochrane Centre, The Cochrane Collaboration; Copenhagen, Denmark). Each domain was ranked as low, high, or unclear risk. Disagreements between reviewers were resolved by consensus. For the overall risk-of-bias judgment, we did not include the domain of participant blinding, which is unrealistic for PEI. If no other domain

Figure 1. PRISMA Flow Diagram of Study Selection



was deemed high risk, the study was defined as low risk; conversely, high-risk studies had at least 1such domain.

Effect Size Calculation

To calculate the standardized mean differences (SMDs, expressed as the Hedges *g*) for each study, we subtracted the mean difference in the control group from the mean difference in the exercise group and divided the difference by the pooled standard deviation.[24] Study-specific effect sizes were extracted through a uniform system. We used data on pre- and post-intervention depressive scores in the exercise and control groups only and assessed the between-group difference in mean scores. In the single study[25] that had a pharmacotherapy group and a usual-care control group, we used data only from the latter. Two studies had 2 intervention groups: one had 2 aerobic groups whose effect sizes were pooled herein,[26] while the other[27] had aerobic (traditional walking) and meditation (Buddhism walking meditation) groups; we included only the aerobic group in our analyses because of lack of clarity about the added effect of mind-body interventions. The sample size used corresponded to the number of subjects with available scores in the experimental and control groups regardless of randomization. Participants without posttreatment depressive scores were not included in our analyses.

Data Synthesis and Analyses

DerSimonian and Laird random-effects models were used to aggregate study-specific SMDs and calculate a pooled point estimate with a corresponding 95% confidence interval (CI).[28] The between-study variance was calculated with T^2 using the formula

$$T^2 = (Q - df)/C$$

in which Q is the Cochrane heterogeneity statistic (total variance expressed as χ^2), df is degrees of freedom, and C is a scaling factor. The total proportion of variance due to heterogeneity was calculated using the I^2 metric. Finally, to assess publication bias, we constructed a funnel plot and subjectively evaluated it for asymmetry. Due to the small number of studies, we refrained from empirical evaluations of asymmetry, which would have been underpowered.[29] We conducted 3 exploratory analyses to examine sources of heterogeneity to our pooled estimate. First, we stratified studies by mean age lower or higher than 80 years, which was akin to the weighted mean age for all participants. Second, we restricted the analysis to studies with participants with intact cognitive function. Thus, only studies that excluded participants with MMSE score < 23 or dementia were included. Last, we restricted our analysis to studies in which

Table 1. Study and Participant Characteristics of Included Studies

Study	N	Control	Age, Mean, y	Residence	Sex	MMSE Score, Mean[a]	Measure	Depression[b]
Antunes et al, 2004[30]	46	Waiting list	67.0	NA	Men	≥ 24	GDS-30	33.5
Belvederi Murri et al, 2015[26]	121	Antidepressant	75.2	Community	Mixed	26.9	HDRS	100.0
Brenes et al, 2007[25]	37	Antidepressant; usual care	74.5	NA	Mixed	28.8	GDS-15	73.6
Lok et al, 2017[31]	80	Waiting list	> 65.0	Nursing home	Mixed	NA	BDI	80.0
Prakhinkit et al, 2014[27]	45	Walking; sedentary	76.6	Community	Women	NA	GDS-30	99.9
Sims et al, 2006[32]	32	Educational sessions	74.3	Community	Mixed	No dementia	GDS-30	100.0
Singh et al, 1997[33]	32	Educational sessions	70.9	Community	Mixed	28.5	BDI	99.5
Underwood et al, 2013[34]	891	Educational sessions	86.5	Nursing home	Mixed	18.4	GDS-15	48.9
Verrusio et al, 2014[35]	24	Antidepressant	75.5	NA	Mixed	NA	GDS-15	100.0

[a]MMSE administered at baseline or stated exclusion criteria for cognitive function.
[b]Percent with at least minor depression at baseline as assessed by a depressive symptom score.
Abbreviations: BDI = Beck Depression Inventory, GDS-15 = 15-item Geriatric Depression Scale, GDS-30 = 30-item GDS, HDRS = Hamilton Depression Rating Scale, MMSE = Mini-Mental State Examination, NA = not available.

Table 2. Exercise Intervention Characteristics for Included Studies

Study	Type of Intervention	Frequency/wk	Session Length, min	Intensity	Intervention, wk
Antunes et al, 2004[30]	Aerobic	3	20–60	NA	24
Belvederi Murri et al, 2015[26]	Aerobic	3	60	Moderate-vigorous	24
Brenes et al, 2007[25]	Both	3	60	NA	16
Lok et al, 2017[31]	Aerobic	4	40–70	NA	10
Prakhinkit et al, 2014[27]	Aerobic	3	20–30	Moderate	12
Sims et al, 2006[32]	Resistance	3	By repetitions	Moderate	10
Singh et al, 1997[33]	Resistance	3	45	Moderate-vigorous	10
Underwood et al, 2013[34]	Both	2	45	Moderate	51
Verrusio et al, 2014[35]	Aerobic	2	60	Moderate	24

Abbreviation: NA = not available.

> 50% patients had depression at baseline. Analyses were done with RevMan 5.3.

RESULTS

Study Populations

Nine RCTs involving 1,308 study participants met study inclusion criteria (Figure 1). Table 1 details the study and participant characteristics of all included RCTs.[25–27,30–35] Eight studies[25–27,30–33,35] were of modest sample size (24–121 subjects) and included participants with mean ages from 67 to 77 years, whereas the remaining study[34] was larger and had participants with a mean age of 86 years; the weighted mean age for all included study participants was 82 years. We also noticed various levels of the participants' cognitive function: 5 studies[25,26,30,32,33] excluded participants with cognitive decline, while 3[27,31,35] provided partial or no information on this variable. The remaining study[34] included nursing home residents with a mean MMSE[20] score of 18.4, which is lower than the 5th percentile for community-residing adults of similar age.[36] Further, this score indicates that about 75% of study participants would not have been included in the studies with an inclusion criterion of MMSE score ≥ 23. Depressive symptoms were assessed using the GDS in 6 studies; the BDI was used in 2 studies, and the HDRS was used in the remaining study. Last, 7 studies had a preponderance of patients with at least minor depression at baseline, while 2 studies[30,34] included a majority of patients with no depression.

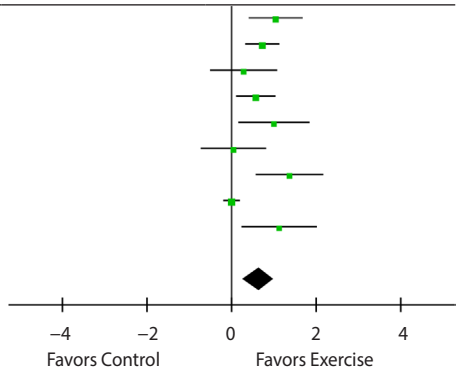
Exercise Interventions

Table 2 lists the characteristics of the physical exercise interventions in included studies. Most of the studies used an aerobic or mixed aerobic and anaerobic intervention with a frequency of 2 to 4 sessions per week. Three studies[27,30,31] used an escalating duration per session, while a single study[32] used a resistance exercise intervention for which session duration was defined by repetitions. Mean intervention length was 20.1 ± 13.2 weeks, and only 1 study[34] was longer than 24 weeks. Only 5 studies[25–27,33,34] had baseline and post-intervention assessments of physical fitness. Four used the Short Physical Performance Battery (SPPB)[37] or its components (eg, 6-minute walk distance, timed chair stand). Of these, 2 studies[27,33] reported a significant improvement in some SPPB metrics and 1[25] reported no change, while the remaining study[34] reported a strong trend toward reduced post-intervention physical fitness according to reduction in SPPB score. Another study[26] used peak oxygen consumption and reported mild improvement with exercise. Last, 5 studies[25,27,30,31,35] reported no measure of adherence to the exercise intervention other than attrition. Four studies[26,32–34] had explicit session attendance rates. In 2 small studies with 30 exercise sessions over 10 weeks, median participation was 93%[33] and > 63%,[32] amounting to exercise frequencies of 3 and 2 sessions per week, respectively. In another study,[26] there was a 70% attendance in 72 sessions. In contrast, a larger study[34] with a 51-week exercise intervention reported a median participation of only 52% among study completers. This adherence amounts to a median frequency of 1 session

Figure 2. Forest Plot of Study-Specific and Overall Standardized Mean Differences in Depressive Symptoms Between Exercise Intervention and Nonactive Control Condition

Study	Exercise Mean	Exercise SD	Exercise Total	Control Mean	Control SD	Control Total	Weight	Standardized Mean Difference IV, Random, 95% CI
Antunes et al, 2004[30]	5.74	5.22	23	0.69	4.25	23	11.2%	1.04 (0.42 to 1.66)
Belvederi Murri et al, 2015[26]	13.83	7.1196	79	8.7	6.809	42	13.8%	0.73 (0.34 to 1.11)
Brenes et al, 2007[25]	2.5	3.4	14	1.6	2.6	12	9.5%	0.28 (−0.49 to 1.06)
Lok et al, 2017[31]	5.67	7.03	40	2.16	4.92	40	13.1%	0.57 (0.13 to 1.02)
Prakhinkit et al, 2014[27]	1.8	1.31	13	0.7	0.74	13	9.0%	1.00 (0.18 to 1.82)
Sims et al, 2006[32]	0.41	4.49	13	0.22	3.9	14	9.7%	0.04 (−0.71 to 0.80)
Singh et al, 1997[33]	11.5	6.13	17	4.6	2.96	15	9.4%	1.37 (0.59 to 2.15)
Underwood et al, 2013[34]	0.1	3.25	224	0.1	3.25	260	15.7%	0.00 (−0.18 to 0.18)
Verrusio et al, 2014[35]	3	2.27	12	0.4	2.19	12	8.6%	1.13 (0.25 to 2.00)
Total (95% CI)			**435**			**431**	**100.0%**	**0.64 (0.27 to 1.01)**

Heterogeneity: $T^2 = 0.22$; $\chi^2_8 = 36.34$, $P < .0001$; $I^2 = 78\%$
overall effect: $z = 3.38$ ($P = .0007$)



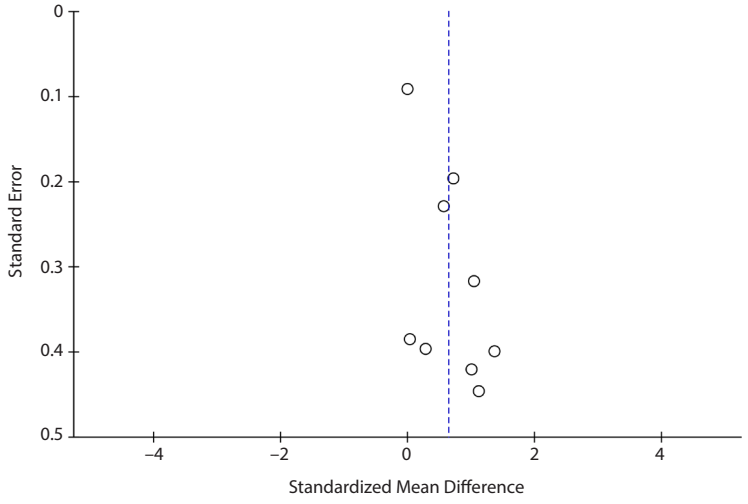
Abbreviation: IV = inverse variance.

per week among completers and possibly lower among participants with shorter follow-up.

Standardized Mean Differences, Heterogeneity, and Publication Bias

A forest plot of the distribution of SMDs is depicted in Figure 2. We used data from 435 and 431 subjects in the exercise and control groups, respectively, which account for 73.9% of study participants; the remainder had no posttreatment scores and represent attrition, which varied across studies. Thus, we could include only participants with pre- and posttreatment scores in our summary results. Eight of 9 included studies reported an effect that was larger than zero, thus favoring the exercise intervention; the remaining study reported no effect. Moderate-to-vigorous PEI was associated with a moderate effect size of 0.64 (95% CI, 0.27 to 1.01; $z = 3.38$; $P < .001$) in reducing depressive symptoms. There was considerable heterogeneity ($T^2 = 0.22$, $Q = 36.34$, $P < .0001$; $I^2 = 78.0\%$) across included studies. Finally, a funnel plot (Figure 3) suggested missing studies in the bottom left side, which could be due to publication bias. However, given the small number of studies, other explanations (such as chance) cannot be ruled out.

Figure 3. Funnel Plot to Assess Publication Bias[a]



[a]The pooled standardized mean difference (SMD) is depicted with a dashed line. Each study-specific SMD is depicted with a circle. The y-axis represents the standard error of the SMD.

Exploratory Analyses

The results of our exploratory analyses are presented in Supplementary Figures 1, 2, and 3. For studies with a lower mean age ($k = 8$), the pooled effect was 0.74 (95% CI, 0.48 to 1.01; $z = 5.54$; $P < .0001$) and heterogeneity was not statistically significant ($T^2 = 0.04$, $Q = 9.67$, $P = .21$; $I^2 = 28.0\%$); the SMD for the study with a higher mean age was 0.00 (95% CI, −0.58 to 0.58). Restricting to studies that reported the exclusion of subjects with dementia ($k = 5$) led to a pooled effect of 0.71 (95% CI, 0.31 to 1.11; $z = 3.46$; $P < .001$) and nonsignificant heterogeneity ($T^2 = 0.10$, $Q = 7.99$, $P = .09$; $I^2 = 49.9\%$). Finally, restricting to studies with a majority of patients with depression at baseline ($k = 7$) led to an effect size of 0.70 (95% CI, 0.40 to 0.99; $z = 4.76$; $P < .0001$) and nonsignificant heterogeneity ($T^2 = 0.04$, $Q = 8.57$, $P = .20$; $I^2 = 30.0\%$).

Risk-of-Bias Assessment

Study-specific and domain-specific risk-of-bias assessments are presented in Supplementary Figures 4 and 5, respectively. Excepting blinding of participants (which was absent in all studies), 4 studies were deemed at low risk, whereas 5 were deemed at high risk of bias. Included studies had a median of 2 high-risk domains with a range of 1

to 4. No domain was particularly concerning. However, this assessment is limited by the deposition of the experimental protocol a priori in a public domain in only 1 instance.[34] Thus, comparisons with prespecified outcome analysis were by and large not available.

DISCUSSION

This systematic review and meta-analysis consolidates previous work by revealing a moderate decrease in depressive symptoms among older adults enrolled in a supervised, moderate-to-vigorous PEI. Nevertheless, our study was among the first to indicate age > 80 years, cognitive decline (MMSE score < 23), and no depression at baseline as likely sources for the observed heterogeneity of this effect.

The considerable heterogeneity we observed is striking in the face of the restrictive approach we used to select RCTs with an adequate exercise intervention. As the T^2 statistic indicates, between-study variance contributed substantially to this heterogeneity. Possible sources of heterogeneity include population attributes such as age, nursing home residence versus community living, independence in activities of daily living, baseline physical fitness, and cognitive function. While these attributes are presumably in considerable correlation, some may be more important than others as effect modifiers of physical exercise. The number of included studies alongside inconsistent reporting did not allow us to perform meta-regression to try to assess the effect modification of exercise on LLD in a continuous fashion. Instead, we used further restriction in exploratory analyses that revealed nonsignificant heterogeneity when only studies with participants under 80 years of age, intact cognition, and some depression at baseline were included. Of special note, in all of these analyses, the T^2 (between-study heterogeneity) was much lower.

Cognitive decline is age related. It is thus biologically plausible that exercise is relatively less beneficial in reducing depressive symptoms in older adults if they have age-related cognitive decline. Physical exercise is thought to enhance adult hippocampal neurogenesis,[38] a process that is dysregulated in depression.[39] Physical exercise very likely strengthens adult hippocampal neurogenesis through increased angiogenesis.[40] However, exercise-induced angiogenesis is negatively correlated with age.[40] Thus, an important pathway of benefit with physical exercise is impaired with increasing age.

Attendance is an important marker of exercise adherence. However, exercise adherence was not uniformly reported in included RCTs and differed among those that did document it. Interestingly, adherence has been prospectively shown to be associated with improvement in physical fitness among participants in an exercise intervention.[41] Thus, high adherence to exercise may be a prerequisite for the effectiveness of such an intervention in reducing depressive symptoms. We note that the observed dose of exercise (given specified adherence) did not meet the recommended dose of exercise in many instances.

Comparison With Previously Published Meta-Analyses

As noted previously, this meta-analysis (9 studies) had relatively narrow inclusion criteria with respect to PEI. The age cutoff used to define LLD is usually 65 years.[42] One systematic review (11 studies)[11] focused on women aged 40 years or over and thus did not meet this definition. Another meta-analysis (18 studies)[13] on older adults addressed the effect of yoga alone on depressive symptoms and is thus not comparable to ours. Another systematic review (11 studies)[14] pooled data from studies of various non-pharmacologic interventions (eg, cognitive-behavioral therapy, light therapy, problem solving, exercise), did not restrict the review to RCTs, and included other empirical studies. Our approach for inclusion was restricted to studies with a clearly defined dose of physical exercise. One larger systematic review (41 studies)[9] used a broad definition of intervention with any type of movement adopted from a previous meta-analysis[10] and included studies with mind-body interventions (eg, yoga, tai chi, meditation) and several unpublished studies. It is thus challenging to compare the authors' finding that age > 75 years was not an effect moderator of exercise and depressive symptoms.[9] Another moderately sized (19 studies) meta-analysis[12] included studies with psychosocial experimental arms such as group support, skills training, and social interventions. Finally, a smaller meta-analysis (8 studies)[15] used a definition of exercise that was akin to the one used herein but included studies with unsupervised interventions and studies whose populations had a major chronic disease (Alzheimer's disease and stroke). Of importance, the authors found a significant effect for exercise on depressive symptoms only without a major comorbid condition, whereas no effect was observed in patients who had such a condition.[15] Accordingly, we decided to exclude studies enrolling such patients from our meta-analysis.

We identified 3 prominent meta-analyses[17,43,44] that estimated the effect of exercise and depressive symptoms without an age cutoff. While these meta-analyses did not address LLD per se, they reported effect sizes of rather similar magnitude to one another and to the one reported herein: one (39 studies)[43] noted a pooled SMD of 0.62 (95% CI, 0.42 to 0.81), another separate meta-analysis (33 studies)[17] reported 0.66 (95% CI, 0.48 to 0.83), and yet another (23 studies)[44] reported 0.68 (95% CI, 0.44 to 0.92). Interestingly, effect moderation by allocation concealment was reported.[17] In our study, allocation concealment was judged as present in only 2 included studies; it was unclear in 5 other studies and missing in 2.

We have not found previous analyses of preexisting cognitive impairment and the effect of exercise on depressive symptoms. Nonetheless, 2 prominent narrative reviews[45,46] about late-life depression stressed the importance of this comorbid condition and stated that more research is required to examine optimal therapy for depression with concomitant cognitive decline.

Depression severity at baseline was not found in a previous meta-analysis[10] as an effect modifier of exercise

and depression. Authors did not discuss the difference in heterogeneity and removed studies with significant heterogeneity even from their primary analysis. Conversely, a meta-analysis not limited to LLD[17] did report effect modification by baseline depression, at least in a univariate analysis. Thus, our finding that at least minor depression at baseline is associated with the heterogeneity of the effect size of PEI on depressive symptoms joins the current literature and strengthens the importance of addressing the severity of depression in future studies.

Limitations of This Meta-Analysis

While several criteria were used to exclude RCTs with an inadequate dose of exercise, included RCTs still had diverse methodology. The components of exercise dose—namely, type, duration, intensity, frequency, and length—varied among studies, as did adherence to assigned intervention. Also, the population in each study was not homogeneous in terms of cognitive and physical functioning. Because cognitive function and physical limitations could affect the ability to perform exercise, it might have been beneficial to assess effect modification with such variables using meta-regression analyses. Instead, we used stratification and restriction, as the number of included studies (ie, < 10) did not permit meta-regression.[29] Former systematic reviews[9,10,14,15,17,44] have included both randomized and non-randomized study designs with active and non-active control groups and with assorted types and doses of exercise interventions and mixed population characteristics; for example, including and pooling together various types and intensities of exercise interventions: low intensity as well as high intensity; aerobic and anaerobic types of exercise as well as yoga and tai chi; and populations with depression as well as without depressive symptoms. Power notwithstanding, our restrictive approach to study selection was meant to facilitate the interpretation of our findings as relates to older adults and the strength of evidence. The small number of studies and their corresponding modest sample sizes, however, may have introduced a bias in sampling error, especially given our use of the Hedges *g* as the point estimate.[47] Nevertheless, our summary result falls in the range of previous meta-analyses.[17,43,44] Thus, the small number of studies in our meta-analysis did not appear to cause a sampling error.

CONCLUSION

Physical exercise intervention was found to have a moderate effect on reducing LLD symptoms, consistent with the effect seen in younger individuals. Nevertheless, there is large heterogeneity among studies, probably owing to methodological differences. Given that age > 80 years, MMSE score < 23, and depression at baseline appeared to contribute to this heterogeneity, we suggest that more work is needed to optimize physical exercise interventions to the oldest old and to cognitively impaired individuals. Whether physical exercise could decrease depressive symptoms in subjects who do not fulfill criteria for minor depression awaits further confirmation. In all, a uniform exercise intervention is not likely to benefit all older adults with LLD, and tailoring physical exercise to specific subpopulations is probably required.

Submitted: April 16, 2019; accepted August 8, 2019.

Published online: January 21, 2020.

Author contributions: Dr S. Klil-Drori: concept and design, data extraction, interpretation of data, preparation of manuscript; Dr A. J. Klil-Drori: data extraction, data analysis, interpretation of data; Dr Pira: data extraction; Dr Rej: concept and design, interpretation of data.

Potential conflicts of interest: **Dr A. J. Klil-Drori** has received a research grant from Sanofi Canada outside the submitted work. **Dr Rej** has received a clinician-scientist award from Fonds de Recherche Santé Québec (FRQS) and a research grant from Satellite Healthcare (dialysis company) for an unrelated project. **Dr S. Klil-Drori** was funded in part by a grant from Satellite Healthcare. **Dr Pira** reports no conflict of interests.

Funding/support: No specific funding was received for this study.

Supplementary material: Available at PSYCHIATRIST.COM.

REFERENCES

1. Global Burden of Disease Study 2013 Collaborators. 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. *Lancet*. 2015;386(9995):743–800.
2. Wang J, Wu X, Lai W, et al. Prevalence of depression and depressive symptoms among outpatients: a systematic review and meta-analysis. *BMJ Open*. 2017;7(8):e017173.
3. Taylor WD. Clinical practice: depression in the elderly. *N Engl J Med*. 2014;371(13):1228–1236.
4. Haigh EAP, Bogucki OE, Sigmon ST, et al. Depression among older adults: a 20-year update on five common myths and misconceptions. *Am J Geriatr Psychiatry*. 2018;26(1):107–122.
5. Bosmans JE, van Schaik DJ, Heymans MW, et al. Cost-effectiveness of interpersonal psychotherapy for elderly primary care patients with major depression. *Int J Technol Assess Health Care*. 2007;23(4):480–487.
6. 2018 Physical Activity Guidelines Advisory Committee. Scientific Report, Part F, Chapter 9: Older Adults. Office of Disease Prevention and Health Promotion website. https://health.gov/paguidelines/second-edition/report/pdf/15_F-9_Older_Adults.pdf. 2018. Cited December 16, 2018.
7. Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. *JAMA*. 2018;320(19):2020–2028.
8. Blazer DG. Depression in late life: review and commentary. *J Gerontol A Biol Sci Med Sci*. 2003;58(3):249–265.
9. 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.
10. Bridle C, Spanjers K, Patel S, et al. Effect of exercise on depression severity in older people: systematic review and meta-analysis of randomised controlled trials. *Br J Psychiatry*. 2012;201(3):180–185.
11. Pérez-López FR, Martínez-Domínguez SJ, Lajusticia H, et al; Health Outcomes Systematic Analyses Project. Effects of programmed exercise on depressive symptoms in midlife and older women: a meta-analysis of randomized controlled trials. *Maturitas*. 2017;106:38–47.
12. Forsman AK, Schierenbeck I, Wahlbeck K. Psychosocial interventions for the prevention of depression in older adults: systematic review and meta-analysis. *J Aging Health*. 2011;23(3):387–416.
13. Patel NK, Newstead AH, Ferrer RL. The effects of yoga on physical functioning and health related quality of life in older adults: a systematic review and meta-analysis. *J Altern Complement Med*. 2012;18(10):902–917.
14. Holvast F, Massoudi B, Oude Voshaar RC, et al. Non-pharmacological treatment for depressed older patients in primary care: a systematic review and meta-analysis. *PLoS One*. 2017;12(9):e0184666.
15. 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. *Br J Psychiatry*. 2016;38(3):247–254.
16. Moher D, Liberati A, Tetzlaff J, et al; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269, W64.
17. Gordon BR, McDowell CP, Hallgren M, et al. Association of efficacy of resistance exercise

training with depressive symptoms: meta-analysis and meta-regression analysis of randomized clinical trials. *JAMA Psychiatry*. 2018;75(6):566–576.
18. Freemantle N. Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? *BMJ*. 2001;322(7292):989–991.
19. Global Recommendations on Physical Activity for Health: 65 Years and Above. World Health Organization website. https://www.who.int/dietphysicalactivity/physical-activity-recommendations-65years.pdf?ua=1. 2011. Cited December 17, 2018.
20. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–198.
21. Yesavage JA. Geriatric Depression Scale. *Psychopharmacol Bull*. 1988;24(4):709–711.
22. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4(6):561–571.
23. Endicott J, Cohen J, Nee J, et al. Hamilton Depression Rating Scale: extracted from Regular and Change Versions of the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry*. 1981;38(1):98–103.
24. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. Orlando, FL: Academies Press; 1985.
25. 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.
26. Belvederi Murri M, Amore M, Menchetti M, et al; Safety and Efficacy of Exercise for Depression in Seniors (SEEDS) Study Group. Physical exercise for late-life major depression. *Br J Psychiatry*. 2015;207(3):235–242.
27. Prakhinkit S, Suppapitiporn S, Tanaka H, et al. Effects of Buddhism walking meditation on depression, functional fitness, and endothelium-dependent vasodilation in depressed elderly. *J Altern Complement Med*. 2014;20(5):411–416.
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188.
29. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol*. 2000;53(11):1119–1129.
30. Antunes HK, Stella SG, Santos RF, et al. Depression, anxiety and quality of life scores in seniors after an endurance exercise program. *Br J Psychiatry*. 2005;27(4):266–271.
31. Lok N, Lok S, Canbaz M. The effect of physical activity on depressive symptoms and quality of life among elderly nursing home residents: randomized controlled trial. *Arch Gerontol Geriatr*. 2017;70:92–98.
32. Sims J, Hill K, Davidson S, et al. Exploring the feasibility of a community-based strength training program for older people with depressive symptoms and its impact on depressive symptoms. *BMC Geriatr*. 2006;6(1):18.
33. Singh NA, Clements KM, Fiatarone MA. A randomized controlled trial of progressive resistance training in depressed elders. *J Gerontol A Biol Sci Med Sci*. 1997;52(1):M27–M35.
34. Underwood M, Lamb SE, Eldridge S, et al. Exercise for depression in elderly residents of care homes: a cluster-randomised controlled trial. *Lancet*. 2013;382(9886):41–49.
35. Verrusio W, Andreozzi P, Marigliano B, et al. Exercise training and music therapy in elderly with depressive syndrome: a pilot study. *Complement Ther Med*. 2014;22(4):614–620.
36. Kvitting AS, Fällman K, Wressle E, et al. Age-normative MMSE data for older persons aged 85 to 93 in a longitudinal Swedish cohort. *J Am Geriatr Soc*. 2019;67(3):534–538.
37. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*. 1994;49(2):M85–M94.
38. Anacker C, Hen R. Adult hippocampal neurogenesis and cognitive flexibility—linking memory and mood. *Nat Rev Neurosci*. 2017;18(6):335–346.
39. Toda T, Parylak SL, Linker SB, et al. The role of adult hippocampal neurogenesis in brain health and disease. *Mol Psychiatry*. 2019;24(1):67–87.
40. Maass A, Düzel S, Goerke M, et al. Vascular hippocampal plasticity after aerobic exercise in older adults. *Mol Psychiatry*. 2015;20(5):585–593.
41. Hall SE, Lawal OA, Clark CM, et al. Novel approach to characterize heterogeneity in an aerobic exercise intervention. *Med Sci Sports Exerc*. 2019;51(7):1506–1516.
42. Alexopoulos GS. Depression in the elderly. *Lancet*. 2005;365(9475):1961–1970.
43. Cooney GM, Dwan K, Greig CA, et al. Exercise for depression. *Cochrane Database Syst Rev*. 2013;9(9):CD004366.
44. Kvam S, Kleppe CL, Nordhus IH, et al. Exercise as a treatment for depression: a meta-analysis. *J Affect Disord*. 2016;202:67–86.
45. Landi F, Abbatecola AM, Provinciali M, et al. Moving against frailty: does physical activity matter? *Biogerontology*. 2010;11(5):537–545.
46. Unützer J. Clinical practice: late-life depression. *N Engl J Med*. 2007;357(22):2269–2276.
47. Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. *PLoS One*. 2018;13(9):e0204056.

Editor's Note: We encourage authors to submit papers for consideration as a part of our Focus on Geriatric Psychiatry section. Please contact Jordan F. Karp, MD, at jkarp@psychiatrist.com, or Gary W. Small, MD, at gsmall@psychiatrist.com.

See supplementary material for this article at PSYCHIATRIST.COM.

THE JOURNAL OF
CLINICAL PSYCHIATRY
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL PSYCHOPHARMACOLOGY

Supplementary Material

Article Title: Exercise Intervention for Late-Life Depression: A Meta-Analysis

Author(s): Sivan Klil-Drori, MD; Adi J. Klil-Drori, MD, MSc; Shamira Pira, MD, MSc; and Soham Rej, MD, MSc

DOI Number: https://doi.org/10.4088/JCP.19r12877

List of Supplementary Material for the article

1. Figure 1 — Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with age <80)
2. Figure 2 — Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with no dementia)
3. Figure 3 — Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with majority depression at baseline)
4. Figure 4 — Study-specific risk of bias assessment
5. Figure 5 — Domain-specific risk of bias assessment

Disclaimer

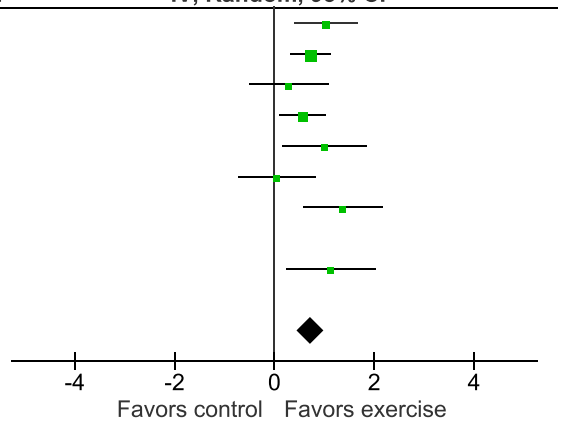
This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplementary Figure 1. Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with age <80)

	Exercise			Control				Std. Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Antunes 2004	5.74	5.22	23	0.69	4.25	23	13.0%	1.04 [0.42, 1.66]
Belvederi Murri 2015	13.83	7.1196	79	8.7	6.809	42	23.3%	0.73 [0.34, 1.11]
Brenes 2007	2.5	3.4	14	1.6	2.6	12	9.2%	0.28 [-0.49, 1.06]
Lok 2017	5.67	7.03	40	2.16	4.92	40	19.8%	0.57 [0.13, 1.02]
Prakhinkit 2014	1.8	1.31	13	0.7	0.74	13	8.4%	1.00 [0.18, 1.82]
Sims 2006	0.41	4.49	13	0.22	3.9	14	9.6%	0.04 [-0.71, 0.80]
Singh 1997	11.5	6.13	17	4.6	2.96	15	9.1%	1.37 [0.59, 2.15]
Underwood 2013	0.1	3.25	224	0.1	3.25	260	0.0%	0.00 [-0.18, 0.18]
Verrusio 2014	3	2.27	12	0.4	2.19	12	7.6%	1.13 [0.25, 2.00]
Total (95% CI)			**211**			**171**	**100.0%**	**0.74 [0.48, 1.01]**

Heterogeneity: Tau² = 0.04; Chi² = 9.67, df = 7 (P = 0.21); I² = 28%

Test for overall effect: Z = 5.54 (P < 0.00001)

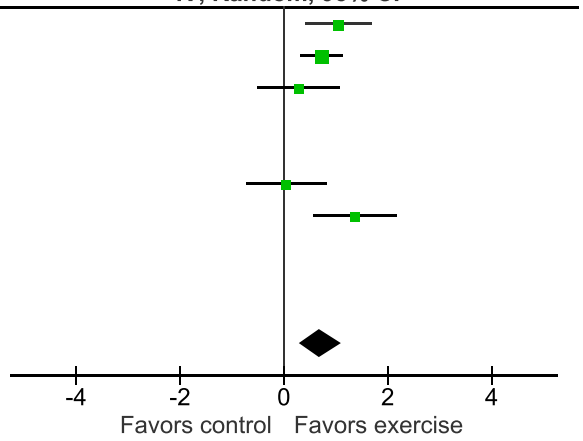


Supplementary Figure 2. Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with no dementia)

	Exercise			Control				Std. Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Antunes 2004	5.74	5.22	23	0.69	4.25	23	20.9%	1.04 [0.42, 1.66]
Belvederi Murri 2015	13.83	7.1196	79	8.7	6.809	42	29.9%	0.73 [0.34, 1.11]
Brenes 2007	2.5	3.4	14	1.6	2.6	12	16.3%	0.28 [-0.49, 1.06]
Lok 2017	5.67	7.03	40	2.16	4.92	40	0.0%	0.57 [0.13, 1.02]
Prakhinkit 2014	1.8	1.31	13	0.7	0.74	13	0.0%	1.00 [0.18, 1.82]
Sims 2006	0.41	4.49	13	0.22	3.9	14	16.8%	0.04 [-0.71, 0.80]
Singh 1997	11.5	6.13	17	4.6	2.96	15	16.1%	1.37 [0.59, 2.15]
Underwood 2013	0.1	3.25	224	0.1	3.25	260	0.0%	0.00 [-0.18, 0.18]
Verrusio 2014	3	2.27	12	0.4	2.19	12	0.0%	1.13 [0.25, 2.00]
Total (95% CI)			**146**			**106**	**100.0%**	**0.71 [0.31, 1.11]**

Heterogeneity: Tau² = 0.10; Chi² = 7.99, df = 4 (P = 0.09); I² = 50%

Test for overall effect: Z = 3.46 (P = 0.0005)

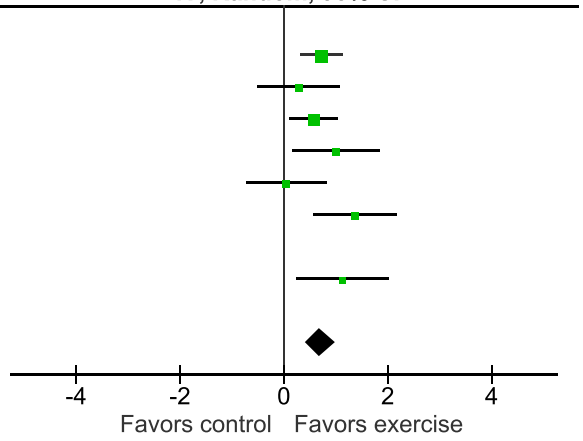


Supplementary Figure 3. Forest plot of study-specific and overall standardized mean differences in depressive symptoms between exercise intervention and nonactive control condition (only studies with majority depression at baseline)

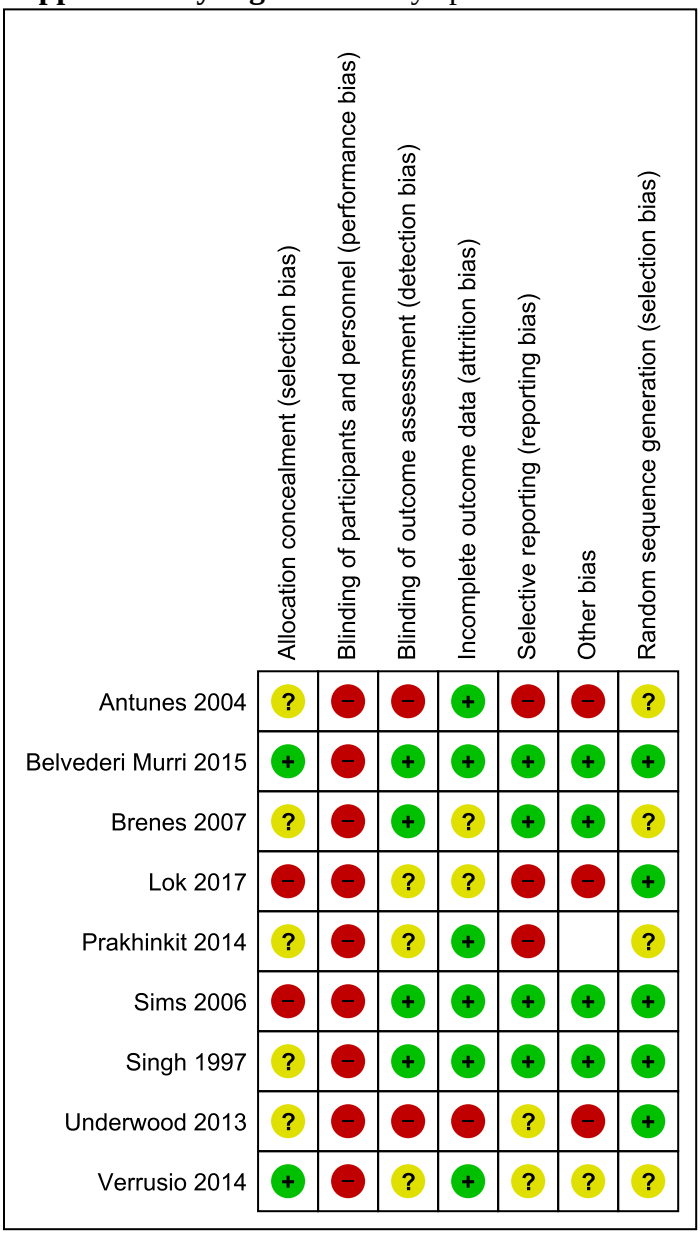
	Exercise			Control				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Antunes 2004	5.74	5.22	23	0.69	4.25	23	0.0%	1.04 [0.42, 1.66]
Belvederi Murri 2015	13.83	7.1196	79	8.7	6.809	42	26.2%	0.73 [0.34, 1.11]
Brenes 2007	2.5	3.4	14	1.6	2.6	12	10.8%	0.28 [-0.49, 1.06]
Lok 2017	5.67	7.03	40	2.16	4.92	40	22.5%	0.57 [0.13, 1.02]
Prakhinkit 2014	1.8	1.31	13	0.7	0.74	13	9.8%	1.00 [0.18, 1.82]
Sims 2006	0.41	4.49	13	0.22	3.9	14	11.2%	0.04 [-0.71, 0.80]
Singh 1997	11.5	6.13	17	4.6	2.96	15	10.6%	1.37 [0.59, 2.15]
Underwood 2013	0.1	3.25	224	0.1	3.25	260	0.0%	0.00 [-0.18, 0.18]
Verrusio 2014	3	2.27	12	0.4	2.19	12	8.9%	1.13 [0.25, 2.00]
Total (95% CI)			**188**			**148**	**100.0%**	**0.70 [0.41, 0.99]**

Heterogeneity: Tau² = 0.04; Chi² = 8.57, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 4.76 (P < 0.00001)



Supplementary Figure 4. Study-specific risk of bias assessment.



Supplementary Figure 5. Domain-specific risk of bias assessment.