

Table 1

Meta-analyses on the effects of exercise on Depression. RCT=Randomized controlled trial; SMD=Standardized mean difference; g=Hedges' g. MDD=Major depressive disorder; HDRS=Hamilton Depression Rating Scale; BDI=Beck Depression inventory; GDS=Geriatric Depression Scale; CES-D=Center for Epidemiologic Studies Depression Scale; Bold=Significant effects ($p < 0.05$). *Adequately concealed random allocation, blinded outcome assessment, intention-to-treat-analysis. ** PEDro score ≥ 6 .

AUTHORS, YEAR	STUDY POPULATION	TYPES OF EXERCISE	CONTROLS	NUMBER OF STUDIES (K)	NUMBER OF PARTICIPANTS (N)	EFFECTS (95% CI)
AEROBIC AND/OR RESISTANCE EXERCISE						
Lawlor & Hopper (2001)	Non-clinical population (k = 9): clinical interview or cut-off on depression-scales	Walking, running, resistance	No treatment (N = 432)	14	620	No treatment: SMD = -1.1 (-1.5, -0.6)
	Clinical population with depression (k = 5)		Cognitive therapy (N = 188)			Cognitive therapy: SMD = -0.3 (-0.7, 0.1)
Rethorst et al. (2009)	Clinical population (k = 17)	Aerobic, resistance, combined	Psychotherapy (k = 4)	58	2.982	g = -0.80 (-0.92, 0.67)
	Non-clinical population (k = 41)		Antidepressants (k = 3)			Psychotherapy: g = -0.26 (-0.63, 0.12)
			N/A (k = 51)			Antidepressants: g = 0.02 (-0.15, 0.18)
						Effect in clinical population significantly larger.
Krogh et al. (2011)	Clinical depression	Aerobic, resistance, combined	No treatment (k = 6)	13	584	SMD = -0.40 (-0.66, -0.14)
			Control intervention (k = 7; health education, meditation, occupational therapy, low intensity exercise)			High-quality*: SMD = -0.19 (-0.70, 0.31)
Silveira et al. (2013)	Clinical population	Aerobic, resistance	N/A	10	758	SMD = -0.61 (-0.88, -0.33)
Cooney et al. (2013)	Clinical diagnosis of depression or cut-off on scales	Various	No treatment, waitlist, placebo intervention, usual care	35	1.356	SMD = -0.62 (-0.81, -0.42) High-quality*: SMD = -0.18 (-0.47, 0.11)
Josefsson et al. (2014)	Clinical population (k = 2)	Aerobic, resistance	No treatment, treatment as usual, low intensity exercise, health education	13	720	g = -0.77 (-1.14, -0.41)
	Non-clinical population (k = 10)					vs. no treatment: g = -0.97 (-1.40, -0.54)
	Both (k = 3)					High-quality*: g = -0.43 (-1.07, 0.21)
Schuch et al. (2016)	Clinical MDD (k = 9)	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	No intervention, usual care, waitlist. Active controls (i.e., stretching) excluded	25	1.487	SMD = -0.98 (-1.28, -0.68)
	MDD and dysthymia (k = 2)					Larger effects in MDD and exercise at moderate or vigorous intensities.
	Mixed or non-clinical (k = 14)					Larger effects in MDD and exercise at moderate or vigorous intensities.
	MDD and dysthymia (k = 2)					
	Mixed or non-clinical (k = 14)					
Kvam et al. (2016)	Diagnosis of unipolar depression, small proportion with minor depression/dysthymia	Aerobic, resistance, combined	Placebo treatment, usual care, no treatment, waitlist, pharmacotherapy, psychotherapy, alternative interventions. Active controls excluded	23	977	g = -0.68 (-0.92, -0.44) vs. psychotherapy: g = -0.22 (-0.65, 0.21) vs. antidepressant: g = -0.08 (-0.33, 0.18) High-quality* & vs. control: g = -0.26 (-0.61, 0.08)
Nebiker et al. (2018)	Diagnosis of depression or depression scale above threshold	Aerobic/endurance,	Placebo treatment, usual care, no treatment, pharmacotherapy, cognitive therapy. Active controls excluded	27	1.452	Endurance: SMD = -0.79 (90% CI: -1.10, -0.48); p < 0.00001
		resistance/neuromuscular				Neuromuscular: SMD = -1.14 (90% CI: -1.50, -0.78); p < 0.00001 Low risk of bias**: Moderate effects size (endurance), large effects (neuromuscular).
Morres et al. (2019)	Clinical population, diagnosis of MDD	Aerobic	Any treatment (e.g., psychotherapy, medication) or condition (waitlist, no treatment). Active controls excluded	11	455	g = -0.79 (-1.01, -0.57) Low risk of bias** : g = -0.70 (-0.94, -0.45)
Carneiro et al. (2020)	Clinical diagnosis of depression or scores above threshold	Resistance with only minimal or no warm-up	No additional physical activity	4	253	Several combinations (2 primary outcomes in 2 studies (HDRS/BDI and HDRS/GDS; 2 studies used HDRS & CES-D): HDRS only: g = -0.281 (-1.05, 0.49) BDI/GDS: g = -0.602 (-1.39, 0.18) High heterogeneity reported.
MIND-BODY EXERCISE						
Cramer et al. (2013)	Diagnosis of depression or elevated scores in scales	Yoga	TAU, relaxation, aerobic exercise	12	619	vs. TAU: SMD = -0.69 (-0.99, -0.39) vs. relaxation: SMD = -0.62 (-1.03, -0.22) vs. aerobic exercise: SMD = -0.99 (-0.99, -0.18) Low methodological quality of included studies, only 2 studies with aerobic exercise as comparator.
Guo et al. (2019)	MDD	Qigong	Waitlist, usual care, pharmacological therapy, walking	7	382	g = -0.64 (-0.92, -0.35)
Brinsley et al. (2020)	Depressive symptoms in patients with mental disorders (depression, anxiety, panic, trauma and stress related, psychotic, substance use disorders)	Yoga	TAU, waitlist, attention control	13	632	SMD = -0.41 (-0.65, -0.17)
Seshadri et al. (2021)	MDD (outpatients)	Exercise (k = 15), yoga (k = 7), tai chi (k = 3)	Any control condition (including other treatment interventions or TAU)	25	2.083	g = -0.63 (-0.76, -0.50)
						Exercise: g = -0.74 (-1.40, -0.06)
						Tai chi: g = -0.28 (-0.68, 0.12)
						Yoga: g = -0.69 (-1.02, -0.36) Only studies with lowest risk of bias (k = 6): g = -0.41 (-0.66, -0.14)

Table 2

Meta-analyses on various or biological markers; VO₂max=Maximal oxygen uptake; VO₂peak=Peak oxygen uptake; g=Hedges' g. MDD=Major depressive disorder; TAU=Treatment as usual; BDNF=Brain-derived neurotrophic factor. Bold=Significant effects (p<0.05). *Quality-criteria not specified.

AUTHORS, YEAR	DOMAIN	STUDY POPULATION	TYPES OF EXERCISE	CONTROLS	NUMBER OF STUDIES (K)	NUMBER OF PARTICIPANTS (N)	EFFECTS (95% CI)
Stubbs et al. (2016a)	Cardiorespiratory fitness (VO ₂ max or VO ₂ peak)	Clinical diagnosis of MDD and cut-off in scales	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	Not specified	7	498	G = 0.64 (0.32, 0.96)
							MDD only: g = 0.41 (0.18, 0.64)
							High-quality studies*: g = 0.60 (0.19, 1.00)
Schuch et al. (2016)	Quality of life (validated instruments)	MDD and cut-off in scales, dysthymia	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	Non-active	6	198	Overall QoL: SMD = 0.39 (0.05, 0.73)
							Physical domain: SMD = 0.53 (0.22, 0.84)
							Psychological domain: SMD = 0.54 (0.22, 0.86)
							Social relationships domain: 0.29 (-0.13, 0.71)
							Environmental domain: 0.37 (-0.12, 0.85)
Stubbs et al. (2016b)	Dropout rates in exercise trials	Clinical diagnosis of MDD and cut-off in scales	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	N/A	40	1.720	18.1% (15.0, 21.8%)
							MDD only: 17.2% (13.5, 21.7%) (n.s. vs. all)
							Higher baseline depressive symptoms Higher rate (p = 0.04)
							Supervision by physiotherapists lower rate (p = 0.008)
							Supervision by exercise physiologists lower rate (p = 0.01)
Sun et al. (2018)	Cognitive symptoms (global cognition, heterogenous measures)	MDD	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	Various, including active controls (low intensity exercise); k = 3 studies: uncontrolled	9	642	Global cognition:
							All interventions: SMD = 0.08 (-0.08, 0.24)
							Low intensity exercise: p = 0.048
							Mind-body exercise: p = 0.048
Dinoff et al. (2018)	Resting BDNF (serum or plasma)	MDD	Aerobic, combination with resistance (k = 1)	Not specified	6	176	SMD = 0.43 (-0.06, 0.92) Significant heterogeneity.
Kurebayashi et al. (2018)	Resting BDNF (serum or plasma)	MDD	Aerobic, combination with resistance (k = 1)	TAU, stretching, relaxation	5	199	SMD = 0.05 (-0.23, 0.32)
Beserra et al. (2018)	Cortisol levels (different measures at different times of the day)	Clinical diagnosis of MDD and cut-off in scales	Aerobic, resistance, mind-body, stretching	Non-active	5	209	SMD = -0.65 (-1.30, 0.01)
							Larger and significant effects for aerobic exercise, very large reduction if exercise took place 5 times/week.
Brupbacher et al. (2021)	Sleep quality	Unipolar depression (clinical, cut-off in scales)	Aerobic, resistance, mind-body exercises (network meta-analysis)	Passive control conditions	17	1.645	Moderate aerobic: SMD = -0.31 (-0.62, 0.00)
							Mind-body: SMD = -0.44 (-0.65, -0.24)
							Vigorous aerobic: SMD = -0.50 (-0.81, -0.19)
							Light resistance: SMD = -0.63 (-1.12, -0.14)
							Vigorous resistance: SMD = -1.09 (-1.64, -0.55)

Table 3 - Part 1

Meta-analyses on the effects of exercise on depression in special populations. SMD=Standardized mean difference; g=Hedges' g; EPDS=Edinburgh Postnatal Depression Scale; OR=Odds ratio; TAU=Treatment as usual; MCI=Mild cognitive impairment; QoL=Quality of life; HDRS=Hamilton Depression Rating Scale; BDI=Beck Depression inventory; GDS=Geriatric Depression Scale. *Score >= 5 on the Delphi list. **Inverse variant weighted method. Bold=Significant effects (p<0.05).

AUTHORS, YEAR	STUDY POPULATION	TYPES OF EXERCISE	CONTROLS	NUMBER OF STUDIES (K)	NUMBER OF PARTICIPANTS (N)	EFFECTS (95% CI)
CHILDREN AND ADOLESCENTS						
Brown et al. (2013)	Children 5-11 years.	Interventions to promote or increase physical activity (aerobic exercise, educational lessons, yoga)	Non-physically active controls	9	581	g = -0.26 (-0.43, -0.08)
	Adolescents 12-19 years. Mostly at-risk groups for depression.					
Carter et al. (2016)	Adolescents 13-17 years.	Exercise (aerobic, resistance)	No treatment, waitlist, TAU, psychosocial or educational intervention, attention control (equivalent contact)	8	1.449	SMD = -0.48 (-0.87, -0.10) High-quality studies*: SMD = -0.51 (-0.86, 0.05) Clinical samples: SMD = -0.43 (-0.84, -0.02)
	Volunteers, at-risk population, clinical population with diagnosis of depression.					
Bailey et al. (2017)	Adolescents and young adults 12-25 years.	Any bodily movement produced by skeletal muscles that results in energy expenditure above resting levels. Most trials used aerobic exercise (k = 12)	No treatment, waitlist, attention control	16	771	SMD = -0.82 (-1.02, -0.61) Clinical samples: SMD = -0.72 (-1.15, -0.30) Overall quality of evidence was considered as low to very low.
	Clinical and non-clinical samples; score above cut-off, clinical diagnosis.					
Axelsdottir et al. (2020)	Children and adolescents 6-18 years.	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	Non-exercise interventions, no treatment	4	159	SMD = -0.59 (-1.08, -0.10) Overall quality of evidence was considered as low to very low.
	Score above cut-off (school-samples), k = 1 clinical sample with depression					
PERIPARTUM PERIOD						
Pritchett et al. (2017)	Postpartum population with possible depression, score above cut-off (EPDS) or clinical diagnosis.	Interventions designed to increase aerobic exercise, combination with co-intervention possible (education, diet or counselling).	No treatment or TAU	13	1.734	SMD = -0.44 (-0.75, -0.12) Exercise-only: SMD = -0.56 (-1.13, 0.01) With co-interventions: SMD = -0.35 (-0.66, -0.04)
McCurdy et al. (2017)	Postpartum population with (i) clinical diagnosis of depression or score above cut-off (treatment);	Aerobic exercise (primarily), aerobic/resistance combined, yoga, stretching program. Combination with co-intervention possible.	TAU, waitlist, attention control	16	1.327	SMD = -0.48 (-0.73, -0.22) Treatment trials (i): SMD = -0.48 (-0.73, -0.22) Prevention trials (ii): SMD = -0.22 (-0.36, -0.08)
	(ii) general postpartum population without depression screening (prevention)					
Poyatos-Leon et al. (2017)	Healthy pregnant or postpartum women; with or without depressive symptoms (treatment and prevention)	Wide variety from aerobic to stretching, pilates and yoga	Usual care	12	932	Effect-size** = -0.41 (-0.54, -0.28) Women with postpartum depression: Effect-size = -0.67 (-0.90, -0.40) No depressive symptoms: Effect-size = -0.29 (-0.45, -0.12)
Davenport et al. (2018)	Healthy pregnant women	Prenatal exercise: "Any bodily movement produced by skeletal muscles that results in energy expenditure above resting levels". Combination with co-intervention possible.	No exercise or different frequency, intensity, volume or type of exercise.	13	1.076	Prenatal: Depressive Symptoms: SMD = -0.38 (-0.51, -0.25) Odds of depression: OR = 0.33 (0.21, 0.53) Significant only for exercise-only interventions.
						Postnatal: Depressive symptoms: SMD = 0.05 (-0.02, 0.12)
						Odds of depression: OR = 0.91 (0.61, 1.36)

Table 3 - Part 2

Meta-analyses on the effects of exercise on depression in special populations. SMD=Standardized mean difference; g=Hedges' g; EPDS=Edinburgh Postnatal Depression Scale; OR=Odds ratio; TAU=Treatment as usual; MCI=Mild cognitive impairment; QoL=Quality of life; HDRS=Hamilton Depression Rating Scale; BDI=Beck Depression inventory; GDS=Geriatric Depression Scale. *Score >= 5 on the Delphi list. **Inverse variant weighted method. Bold=Significant effects (p<0.05).

AUTHORS, YEAR	STUDY POPULATION	TYPES OF EXERCISE	CONTROLS	NUMBER OF STUDIES (K)	NUMBER OF PARTICIPANTS (N)	EFFECTS (95% CI)
Peripartum Period (Part 2)						Depressive symptoms:
Carter et al. (2019)	Pregnant and postpartum women at-risk for depression or healthy.	Aerobic, resistance, coaching, motivational approach.	TAU, no intervention, active comparators.	17	1.887	SMD = -0.64 (-0.96, -0.33) Studies with low risk of bias: SMD = -0.30 (-0.45, -0.15) Larger effects in samples with depressive symptoms. Quality of evidence considered low.
Nakamura et al. (2019)	Healthy pregnant women. Primary outcome: postpartum depression scores.	Interventional (k = 6) studies: yoga, aerobic exercise, stretching & resistance. observational (k = 11) studies: leisure time physical activity or all movements contributing to energy expenditure.	Not specified	17	93.676	Overall: SMD = -0.22 (-0.42, -0.01) Interventional studies: SMD = -0.58 (-1.09, -0.08) Effect stronger for studies with moderate-high intensity. (SMD = -0.70 (-1.19, -0.22)) Observational studies: SMD = -0.07 (-0.20, 0.06)
OLDER ADULTS						
Bridle et al. (2012)	Adults ≥ 60 years with clinical diagnosis of depression or score above cut-off.	Mixed exercise (endurance and resistance)	TAU, waitlist, educational program, attention control	7	519	SMD = -0.34 (-0.52, -0.17) Clinical diagnosis: SMD = -0.38 (-0.67, -0.10) Score above cut-off: SMD = -0.34 (-0.62, -0.06)
Rhymer & Watts (2016)	Adults ≥ 60 years with clinical diagnosis of depression (k = 15) or measures of depression scales (no cut-off required, other diagnoses such as diabetes, dementia, osteoarthritis...)	Aerobic, resistance, combined, mind-body	TAU, waitlist, educational program, attention control	41	Not provided	SMD = -0.57 (-0.78, -0.36) Larger effect-size in studies with clinical diagnosis of depression. No difference between age groups or between different types of exercise.
Leng et al. (2018)	Adults with diagnosis of MCI or dementia and measures of depressive symptoms.	Aerobic, resistance, mind-body including combinations (most studies)	TAU, active control (i.e. stretching), social activities	21	2.589 (MCI: 1'160, dementia: 1'429)	Depressive symptoms: SMD = -0.23 (-0.39, -0.07) QoL: SMD = 0.23 (0.01, 0.46) No significant effects for anxiety or apathy.
Klii-Drori et al. (2020)	Adults ≥ 60 years with depression-scales as primary outcome measure (GDS, BDI, HDRS)	Aerobic or resistance at moderate to vigorous intensity, ≥ 2 sessions per week of ≥ 20minutes and for ≥ 8 weeks.	TAU, antidepressants, waitlist, educational program	9	1.308	SMD = -0.64 (-1.01, -0.27) Considerable heterogeneity in the effect of exercise, high risk of bias in 5 studies.

Table 5

Meta-analyses on the effects of exercise on patients with physical diseases (sorted by disease). Delta=Glass's Delta; SMD=Standardized mean difference; g=Hedges' g; TAU=Treatment as usual; BDI=Beck Depression inventory; d=Cohens' d; HIV=Human immunodeficiency virus. Bold=Significant effect (p<0.05). *Alzheimer, migraine, multiple sclerosis, Parkinson, spinal cord injury, stroke, traumatic brain injury.

AUTHORS, YEAR	STUDY POPULATION	TYPES OF EXERCISE	CONTROLS	NUMBER OF STUDIES (K)	NUMBER OF PARTICIPANTS (N)	EFFECTS (95% CI)
Herring et al. (2012)	Sedentary non-depressed adults with chronic illness	Any kind of exercise-training. No combination with other intervention (i.e., educational)	Non-active comparator	90	10.534	Delta = -0.30 (-0.36, -0.25)
Beland et al. (2020)	Adults with a major non-communicable chronic disease and depressive symptoms at baseline	Aerobic exercise	Non-active TAU	24	4.111	SMD = -0.50 (-0.76, -0.25) Cardiac patients only: SMD = -0.67 (-0.99, -0.35)
Eng & Reime (2014)	Adults with a diagnosis of stroke	Any physical activity that is planned, structured, repetitive and purposive to improve or maintain physical fitness or health	Non-active or active comparator (low intensity)	13	1.022	SMD = -0.13 (-0.26, -0.01)
Zheng et al. (2019)	Adults with acute myocardial infarction	Exercise based cardiac rehabilitation	TAU or no control (non-RCTs; k = 2)	6	703	SMD = -0.61 (-1.12, -0.09)
Samartzis et al. (2013)	Adults with chronic heart failure	Experimental exercise group not specified.	TAU	13	3.425	g = -0.32 (-0.52, -0.13) Trend for higher effects of SSRI than exercise.
Tu et al. (2014)	Adults with heart failure	Aerobic, resistance and combinations including mind-body exercise.	TAU	16	3.226	SMD = -0.38 (-0.55, -0.21) Blinded trials only: SMD = -0.14 (-0.22, -0.17)
Narita et al. (2019)	Adults or adolescents with diabetes	Aerobic or resistance for ≥4 weeks	TAU including low-intensity exercise or education	13	962	SMD = -0.59 (-0.93, -0.24) Mean difference in BDI = -2.90 (-4.53, -1.28)
Ensari et al. (2014)	Patients with multiple sclerosis	Exercise training	No-treatment controls	13	477	g = -0.36 (-0.54, -0.18)
Adamson et al. (2015)	Patients with neurologic disorders*	Aerobic, resistance, balance, mind-body and combinations	No intervention or low intensity exercise	23	1.324	g = -0.28 (-0.41, -0.15) Larger effect if public-health physical activity recommendations were met: g = -0.38
Song et al. (2018)	Hemodialysis-patients	Aerobic, resistance and combinations, pilates, stretching	TAU	8	368	SMD = -0.95 (-1.18, -0.73)
Ferreira et al. (2020)	Patients with chronic kidney-disease	Aerobic, resistance, combination, breathing exercise	Active and passive control groups	6	282	vs. active control: SMD = -0.66 (-1.00, -0.33) vs. passive control: Mean difference (BDI) = -6.95 (-8.76, -5.14)
Kelley et al. (2015)	Adults with arthritis	Aerobic, resistance or combination (≥4 weeks)	Non-active control	29	2.449	g = -0.42 (-0.58, -0.26) Number needed to treat = 7 (6, 11)
Craft et al. (2012)	Cancer survivors	Aerobic, resistance, combinations (≥4 weeks)	TAU	15	1.371	d = -0.22 (-0.43, -0.01)
Brown et al. (2012)	Cancer survivors	Aerobic, resistance, yoga	Non-active controls	37	2.929	d = -0.13 (-0.26, -0.01)
Heissel et al. (2019)	Adults living with HIV	Aerobic, aerobic and resistance, yoga	Non-active controls	9	395	SMD = -0.84 (-1.57, -0.11)